The birthweight effects of universal child benefits in pregnancy: quasi-experimental evidence from England and Wales

Mary Reader

CASE/222
February 2021

Centre for Analysis of Social Exclusion
London School of Economics
Houghton Street
London WC2A 2AE
CASE enquiries – tel: 020 7955 6679

Contents

Centre for Analysis of Social Exclusion

The Centre for Analysis of Social Exclusion (CASE) is a multi-disciplinary research centre based at the London School of Economics and Political Science (LSE), within the Suntory and Toyota International Centres for Economics and Related Disciplines (STICERD). Our focus is on exploration of different dimensions of social disadvantage, particularly from longitudinal and neighbourhood perspectives, and examination of the impact of public policy.

In addition to our discussion paper series (CASEpapers), we produce occasional summaries of our research in CASEbriefs, and reports from various conferences and activities in CASEreports. All these publications are available to download free from our website.

For further information on the work of the Centre, please contact the Centre Manager, Annie-Rose Nicholas, on:

Telephone: UK+20 7955 6679
Email: a.nicholas1@lse.ac.uk
Web site: http://sticerd.lse.ac.uk/case

Editorial note

Mary Reader is a Research Officer at CASE. The author would like to thank Dr Kitty Stewart for her invaluable support in the development of this research, as well as Dr Berkay Özcan, Dr Tania Burchardt, Dr Eleni Karagiannaki and Dr Giulia Giupponi for comments on earlier drafts, and Professor Daniel Sturm and Dr Janne Tukiainen for methodological advice. Thanks also to the Members of Parliament who were interviewed about the political context of the Health in Pregnancy Grant for this paper – Rt Hon Harriet Harman QC MP, Kate Green MP, Dr Dan Poulter MP and Rt Hon Ben Bradshaw MP.

Abstract

Over a decade ago, in April 2009, the UK Labour government introduced the Health in Pregnancy Grant (HPG), a cash transfer of the equivalent of child benefit over the third trimester (£190) as a lump sum to all pregnant women in the United Kingdom. As a labelled, universal and unconditional cash transfer with near-universal take-up, the HPG remains the only international example of paying the equivalent of child benefit during pregnancy to improve health outcomes at birth. The grant was designed to improve birthweight by helping mothers afford high-quality nutrition and reducing stress in the prenatal phase. In January 2011, the HPG was abolished by the Conservative-Liberal Democrat Coalition on grounds, in part, that it was a "gimmick" with little evidence of impact on birthweight. This CASEpaper quantitatively evaluates the impact of the HPG on birthweight in England and Wales. Using administrative birth registrations data, I implement a quasi-experimental regression discontinuity (RD) design based on an arbitrary eligibility rule for the HPG. I find that the HPG was responsible for an increase of 11g in birthweight on average and that effects were concentrated on the smallest babies. Increases in birthweight were largest for younger mothers aged 25 and under (29g average increase) and mothers living in areas with high levels of deprivation (20g average increase). While younger mothers experienced a reduction in the probability of low birthweight by 0.9 percentage points (12 percent in relative terms), low birthweight did not fall for the population as a whole. My findings suggest that paying the equivalent of universal child benefits in pregnancy as a labelled lump sum can disproportionately benefit disadvantaged groups such as younger mothers and lead to effect sizes that are larger than would be expected of more general windfall increases in income.

Key words: birthweight, cash transfers, child benefit, infant health
JEL number: I14, I38
Corresponding author: m.reader@lse.ac.uk

1. Introduction

The principle that the state should contribute towards the needs and costs of all children is well established internationally. In 2020, UNICEF found that 60 percent of countries with available information had some form of child benefit enshrined in legislation (UNICEF, 2020). Universal or quasi-universal child benefits – cash payments to all children on a regular basis without conditions attached – are an important plank of the social security system in many high-income countries, particularly in Europe. Prior to 2013, the UK was among the list of European countries (including Austria, Estonia, Finland, Germany) who paid fully universal child benefits, with many others paying variants on this (UNICEF, 2020). Typically, however, this kind of support has only applied to children at or after birth.

Meanwhile, cash transfers during pregnancy have also emerged as a popular mechanism for boosting health outcomes and human capital in many low- and middle-income countries (De Brauw and Peterman, 2020). However, they tend to be either means-tested (e.g., Mozambique, Nigeria, Indonesia, Jamaica) or conditional (e.g., Mexico, Brazil, Turkey, Peru, India) (Chersich et al, 2016).

Over a decade ago, the UK Labour government introduced a Health in Pregnancy Grant (HPG) – a single, tax-free, non-means-tested lump sum of £190 to all pregnant mothers from the twenty-fifth week of pregnancy (Roll, 2007). In doing so, it bucked the international trend by bringing together two disparate aspects of international social security policy design: universal child benefits to provide income support on one hand, and pregnancy cash transfers to improve health outcomes on the other. Other European countries had experimented with universal maternity grants but differed in important respects. France and Belgium have one-off 'birth payments', but their role is to provide income support around the birth of a child rather than to improve birth outcomes by boosting nutrition and reducing stress during pregnancy: they are either paid at birth (France) or at 28-32 weeks of pregnancy at the earliest (Belgium). The country that has come closest to the HPG is Finland, which has had a universal maternity package since 1949 that includes a choice between a 'baby box' of clothes and toys and a one-off €170 cash transfer from the twenty-second week of pregnancy. However, since the box is worth more than the transfer, 95 percent of first-time mothers choose the baby box instead of the cash grant (Koivu et al, 2021). As such, the Finnish maternity grant's take-up levels are significantly lower than the estimated take-up rates of the HPG (see Section 2.1).

The philosophy underpinning the HPG was that many health inequalities were established during pregnancy, and that birth was an arbitrary occasion to start contributing towards the costs of having a child. Paying the equivalent of child benefit in an easily accessible lump sum during the third trimester of pregnancy would help extend support for struggling families and narrow birth inequalities in birthweight and prematurity, the government argued. A government minister stated in committee that "In a way, we could call the measure "Child Benefit in Pregnancy," but we are calling it something else." (Public Bills Committee, 2008, January 10, c. 105). In "calling it something else", the UK government was drawing on European models of child benefit but reorientating it towards improved health at birth and being explicit about its labelling of the grant to that effect.

The HPG is, therefore, a unique international example of a universal, unconditional cash transfer during pregnancy with extremely high estimated take-up and, as such, represents a unique policy experiment of international significance. It was also short lived: the incoming Coalition government abolished the HPG in 2011 on grounds that it was ineffective.

The aims of the grant were threefold. First, the HPG was designed to subsidise better nutrition during pregnancy. Second, the HPG aimed to reduce stress during pregnancy. Both of these aims were intended to promote healthy birth outcomes and to reduce low birthweight and prematurity. Third, the HPG was intended to contribute towards goals beyond health: to prevent the formation of inequalities in the early stage of the life course, to reduce child poverty, and to promote maternal choice.

It was the HPG's record on the first of these aims – improvements in birth outcomes through nutrition – which was heavily contested by Opposition Members of Parliament (MPs), and which eventually led to the abolition of the HPG by the Coalition government in January 2011 (Edmonds and Kennedy, 2010). Charities and Opposition MPs argued that, since the grant was only paid in the final trimester of pregnancy, it was too late for nutritional improvements to have an impact on birthweight. In response, the Labour government underplayed the argument on the potential impact of the HPG on nutrition while maintaining that improvements in birthweight could still be achieved by reducing stress. The latter, they maintained, was a factor of particular concern in the later stages of pregnancy.

Clearly, therefore, the HPG was initially introduced on the basis that it would improve birthweight through a combination of improved nutrition or reduced stress. The Coalition government abolished the HPG on grounds that it would not improve birthweight because it was poorly designed for nutritional impact. The question as to whether the HPG was effective in boosting birthweights in England and Wales is therefore at the heart of this political debate.

Birthweight is widely acknowledged as an important determinant of a range of outcomes, including child and adult health, cognitive development, and educational and labour market inequality (Brooks-Gunn and Duncan, 1997;

Silverwood et al, 2013; Strully et al, 2010; Victora et al, 2008). The effectiveness of policies seeking to improve birthweight, therefore, is a matter of crucial empirical importance. Furthermore, this CASEpaper is to the author's knowledge the first study into the impact of the HPG in England and Wales – previous research concerned Scotland only (Leyland et al, 2017).

To answer this research question, I exploit an arbitrary eligibility rule that applied to the introduction of the HPG: all mothers with an expected date of delivery (EDD) on or after 6 April 2009 were eligible. This facilitates quasi-experimental conditions for the estimation of the impact of the HPG on birthweight, through a regression discontinuity (RD) design. To understand the political context of the HPG, I also conducted interviews with key Members of Parliament who were vocal about the HPG around its introduction and abolition.[1]

My regression discontinuity analysis suggests that the HPG was effective in increasing birthweights in England and Wales by approximately 11g on average (1.8 percent of a standard deviation), and that this effect was concentrated on the smallest babies. Though this positive treatment effect benefitted smaller babies the most, it did not translate into a significant fall in the overall probability of being born with a low birthweight. Birthweight gains were larger (at 29g) for younger mothers on average, however, and this group saw a statistically significant reduction in the probability of low birthweight by 0.9 percentage points (12 percent in relative terms). These effect sizes are comparable with, and indeed in some cases larger than, other quasi-experimental estimates of the impact of a $1000 income increase in the US (Hoynes, Miller and Simon, 2015; Mocan et al, 2015; Chung, Ha and Kim, 2016). This is impressive given that the financial value of the HPG was approximately a quarter of the size of the income increases in these studies. These results come within the context of a net increase in birthweight in the UK from 1986 to 2012 of approximately 40g (Ghosh et al, 2017). Note that the increasing trend in birthweight is not problematic for my research design, however, because the RD methodology examines birthweight within a small region of births around the cut-off eligibility date. Robustness checks support that there was no discontinuous treatment effect in other years, when the HPG had not been introduced.

The structure of the CASEpaper is as follows. Section 2 outlines the policy context and reviews the literature on the impact of increases in income on birthweight. Section 3 details my RD methodology. Section 4 presents the results of my analysis in three parts: the estimated treatment effect on birthweight and low birthweight; a heterogeneity analysis by birthweight, maternal age and English index of income deprivation; and a series of

[1] Interviews were conducted with Rt Hon Harriet Harman QC MP (2 April 2019), Kate Green MP (16 January 2019), Dr Dan Poulter MP (19 March 2019) and Rt Hon Ben Bradshaw MP (16 January 2019).

robustness checks to validate the RD design. Section 5 discusses these results and puts forward limitations of the analysis. Finally, Section 6 concludes and reflects on key policy implications.

2. Background

2.1 Policy specifics

Mothers of all babies due on or after 6 April 2009 were eligible for the HPG (Wright, 2009, February 8). In order to claim the grant, pregnant women were required to visit their midwife or doctor for an ante-natal check-up from the twenty-fifth week of pregnancy, when they would be invited to fill a simple application form for the HPG. After sending the form to HMRC in a free-post envelope, recipients received a letter confirming their application. They could expect to receive their money within seven days, through a direct payment into their bank or building society account (Directgov, 2010). In a podcast released by HMRC to promote the HPG, it was stated that women could spend the money "on whatever you like", including fresh fruit and vegetables, cots, nappies, buggies, and any other lump sums associated with having a healthy baby (Directgov, 2010, May 19). Harriet Harman, a senior member of the Cabinet when the HPG was introduced, described the HPG as "an expression of trust that the mother would do what was best… it was empowering women and expressing confidence in them" in an interview with the author.[2] The HPG hence built on the principle of child benefit – of giving parents (and mothers in particular, to whom the original family allowances were paid) the choice of how to use universal and unconditional payments to help with the costs of children.

The distinctiveness of the HPG, however, was that it extended the principles of child benefit and income support into pregnancy with the ambition of improving birth outcomes. In an interview, Ben Bradshaw, then Minister of State at the Department of Health, argued that there was "as strong, if not an even stronger argument, for extending that [support] to the prenatal period."[3] As the name of the grant made clear, the HPG was primarily an initiative to improve health outcomes, particularly birthweight. The Labour government envisioned two main mechanisms through which the HPG would improve birthweight: improved nutrition and reduced stress. In a standing committee session in 2008, Ben Bradshaw argued that the HPG would "address the serious problem of underweight babies in this country" (Public Bills Committee, 2008, January 10, c. 103).

In the media, nutrition was emphasised. The HPG was widely described as a grant to subsidise healthy eating (see, e.g., 'Eat well cash for mothers-to-be', BBC News, 2007, September 8; Revill, 2007, September 9), and HMRC's online information page about the HPG was illustrated with

[2] Interview with Rt Hon Harriet Harman QC MP (2 April 2019).
[3] Interview with Rt Hon Ben Bradshaw MP (16 January 2019).

drawings of fresh fruit and vegetables. Charities such as the National Childbirth Trust and Bliss, however, raised concerns that the grant was poorly designed for nutritional impact, as it was paid too late in pregnancy (Public Bills Committee, 2008, January 10, c.86). Meanwhile, many Opposition politicians dismissed the ability of the HPG to achieve any impact on birth outcomes, with one frontbench Liberal Democrat MP describing the grant as a "gimmick" (Public Bills Committee, 2008, January 24, c.460). Concerns were raised that the HPG could be squandered on "booze, fags, bingo or plasma screen televisions" (Public Bills Committee, 2008, January 10, cc 90–91; 103). The HPG was also portrayed as inefficient by virtue of its universality.

In committee, the Labour government conceded that the HPG was probably paid too late in pregnancy to significantly impact nutrition. However, they argued that paying the grant in the third trimester would instead alleviate stress – for example by helping women with the cost of large lump sums (such as cots or buggies) – and thus impact on birthweight through psychosocial means. The Labour government also defended the universality of the grant on the basis that it complemented pre-existing means-tested support for pregnant women, thus achieving progressive universalism by "delivering support for all pregnant women and more help for those who need it the most." (Public Bills Committee, 2008, January 10, c. 103-105). Ben Bradshaw MP, speaking on behalf of the government during the committee, described the HPG as a universal programme that would have progressive impacts: "everyone will get it. It will mean a lot more to poor people. It might not mean very much to the media commentators who were so sniffy about this when we announced it; but I have to say that £190 to some of the women in my constituency would make an awful lot of difference" (Public Bills Committee, 2008, January 10, c. 105).

There was, therefore, a discrepancy in the government's messaging about the purpose of the grant between public-facing statements, which tended to emphasise nutrition, and political discussions, which tended to emphasise stress. There was even at times a concession that the HPG was unlikely to significantly impact on birthweight. Labour MP Kate Green, a defender of the HPG's role in tackling child poverty in particular, stated in an interview with the author that the HPG "was a good public message: we were investing in newborns; we were investing in new mums… [but] arguably in terms of some of the outcomes – birthweight and so on – the design didn't really do that, it came too late in pregnancy. And everyone accepted and understood that."[4] These conflicting narratives generated confusion about the aims of the HPG, and scepticism of its impact on birthweight on both sides of the debate.

[4] Interview with Kate Green MP (16 January 2019).

Nevertheless, critiques of the HPG were successfully rebuffed, and the grant was rolled out across the UK from 6 April 2009. After the election of the Conservative-Liberal Democrat Coalition in 2010, however, the HPG fell under renewed scrutiny. In the political context of austerity in the wake of the financial crisis, the Coalition abolished the HPG.

2.2 Causal mechanisms

This CASEpaper tests the hypothesis that the HPG had an impact on birthweight. There are two main mechanisms through which this may have occurred: first, the linking to seeking antenatal health advice at 25 weeks, and second, the direct increase in income.

The hypothesis that the HPG improved birthweight is contested by the only existing study of the grant – Leyland et al (2017) – which concludes that the HPG had no impact on birthweight, prematurity or maternal health in Scotland. The study uses an interrupted time series analysis, which makes it difficult to control for contemporary events that coincided with the treatment period (April 2009 to January 2011). In particular, the paper does not account for the exogenous negative shock of the economic recession on birth outcomes during that period. The subsequent period of austerity may also have contributed towards this bias, though to a lesser degree due to the timings involved. The Coalition's Emergency Budget in May 2010 and Spending Review in October 2010 introduced a number of policies to reduce the deficit that reduced financial support for low-income families (Ridge, 2013; Stewart and Obolenskaya, 2016). Even if the HPG had been effective, the economic downturn from 2008-09 and subsequent public spending cuts may have introduced downward bias in Leyland et al's (2017) estimation of a treatment effect. In contrast, an RD methodology – as implemented in this paper – enables quasi-experimental conditions to be created such that potential outcomes are 'as if' randomised at the treatment cut-off (6 April 2009), and such exogenous shocks are controlled for.

This CASEpaper focuses exclusively on birthweight, as it was the main (and contested) policy objective of the HPG and it is the only relevant indicator available in the birth registrations data. While it has come under increasing scrutiny on the basis that its relative importance has declined in recent years (Goisis, Özcan and Myrskylä, 2016), and that its impact on health dissipates with age (Matsushima, 2018), it is widely acknowledged as one of the most important health outcomes in the life course. In terms of child health, low birthweight is associated with higher levels of infant mortality, particularly in the first 28 weeks after birth (Brooks-Gunn and Duncan, 1997). In adulthood, low birthweight can lead to "permanent impairment" (Victora et al, 2008, p. 340) in terms of shorter adult height, higher blood pressure, Type-2 diabetes and lower offspring birthweight (Silverwood et al, 2013; Victora et al, 2008). Low birthweight has also been associated with lower educational attainment and achievement, grade repetition, and lower adult income (Behrman and Rosenzweig, 2004; Black et al, 2007; Brooks-Gunn and Duncan, 1997; Victora et al, 2008). Birthweight therefore plays an important role in socio-economic stratification within the early part of the life course (Cramer, 1995) and the "the reproduction of inequality over generations" (Strully et al, 2010, p. 535).

One of the main reasons offered by the government for the grant being paid in the third trimester was so that it would be linked to seeking health advice at 25 weeks. Existing research on the HPG in Scotland suggests that it increased the odds of booking before 25 weeks by 10 percent, though it found no impact on birthweight (Leyland et al, 2017). If the HPG incentivised mothers to seek antenatal health advice at an earlier stage, it may have improved birthweight outcomes by increasing mothers' access to information about staying healthy during pregnancy, including on how to maintain a nutritious diet and on the risks of maternal smoking. A recent systematic review found that antenatal care is generally considered to be an important contributor to improved birth outcomes for disadvantaged and vulnerable mothers in high-income countries, though the evidence on specific interventions is less strong than the literature on income and birthweight (Hollowell et al, 2011). Mothers who are teenage, non-White mothers or living in deprived areas are among those who are more likely to engage with antenatal services at a later stage in the UK (Kapaya et al, 2015). These groups may have been particularly incentivised by the HPG to attend antenatal check-ups at an earlier stage in their pregnancy.

The HPG's direct windfall increase to income is also likely to be an important mechanism. Higher levels of income are associated with higher average birthweight and lower incidence of low and extremely low birthweight (Marmot, 2010). Most if not all of the determinants of birthweight – ethnicity (Kelly et al, 2008), gestational age at birth, maternal age, parental body size (including height, weight and BMI), maternal smoking, alcohol and drug consumption, stress, poor nutrition, occupation and socio-economic status (Cramer, 1995) – are patterned by income. There is also robust and wide-ranging evidence that increases in income during pregnancy contribute towards improved health at birth (Cooper and Stewart, 2017; Chersich et al, 2016). This includes permanent, marginal and windfall (one-off) increases in income (such as the Earned Income Tax Credit (EITC) in the US) (Cooper and Stewart, 2017; Hoynes, Miller and Simon, 2012; Almond, Hoynes and Schanzenbach, 2011; Strully et al, 2010). Among child health outcomes, birthweight and other neonatal outcomes appear the most sensitive to increases in income (Cooper and Stewart, 2017).

There are three predominant theories of how increases in income can increase birthweight: the investment model, the family stress model, and the behavioural model (Benzeval et al, 2014). In many cases, these models may interact with one another (Gregg, Waldfogel and Washbrook, 2006). While it is beyond the scope of this CASEpaper to investigate which causal mechanisms were most relevant for the impact of the HPG on birthweight, each were invoked by participants in the political debate surrounding the introduction and abolition of the HPG.

The investment model proposes that increases in income are spent directly on improved nutrition during pregnancy, which increases birthweight. Since fresh fruit and vegetables – a ready source of such nutrients – are relatively expensive (Food Foundation, 2019), the additional nutritional requirements during pregnancy (Williamson, 2006) have inevitable financial implications. Policies that directly increase income during pregnancy have hence been found to reduce food insufficiency (Loopstra and Tarasuk, 2013; Milligan and Stabile, 2011) and increase spending on fruit and vegetables (Gregg, Waldfogel and Washbrook, 2006).

Meanwhile, the family stress model suggests that increases in income improve birthweight by reducing stress both for pregnant women and their families. Stress is a robust predictor of gestational age at birth and low birthweight in itself and has been found to interact with poor diets and unhealthy behaviours such as smoking (Dunkel Schetter and Lobel, 2012). In particular, anxiety is a risk factor for premature delivery (a key cause of low birthweight) and depression is linked with low birthweight (Dunkel Schetter and Lobel, 2012; Glover and O'Connor, 2002). Policies that increase income during pregnancy can combat stress in several ways. For low-income women, increases in income directly reduce financial strain, which is one of the most acute stressors for this demographic (Bloom et al, 2013). Lump-sum increases in income, in particular, may be spent on 'self-care' investments (such as antenatal swimming classes or yoga).[5] A recent randomised control trial in the UK found that antenatal yoga classes, for example, significantly reduce stress and depression (Newham et al, 2014). Furthermore, financial abuse or neglect affects one in five women, and is associated with heightened antenatal stress (Bloom et al, 2013; Sharp-Jeffs, 2015). If cash transfers are paid directly to mothers, it may enhance maternal choice and agency and therefore ameliorate stress levels in such situations. Thus increases in income during pregnancy can have an impact on birthweight through psychosocial means. Windfall cash gains such as the EITC have been associated with a reduction in maternal depression (Boyd-Swan et al, 2016).

Finally, and relatedly, the behavioural model suggests that increases in income reduce unhealthy behaviours (such as maternal smoking, alcohol and drug consumption) and thereby improve birthweight. Approximately

[5] Public threads on Mumsnet about the HPG suggest that some women spent the £190 lump sum on 10-week courses of aqua natal swimming classes and antenatal yoga (see e.g., Mumsnet, 2009, April 17).

10.5 percent of women in England continue to smoke during pregnancy (BBC News, 2017, June 15). While evidence is mixed for adults in general, for pregnant women there is strong evidence that increases in income reduce the probability of engaging in unhealthy behaviours (Cooper and Stewart, 2017). In particular, several studies have found that the likelihood of maternal smoking decreases in response to increases in income (Averett and Wang, 2013; Cowan and Tefft, 2012; Gregg, Waldfogel and Washbrook, 2006; Strully et al, 2010). In large part, this is likely to be due to reductions in stress, in line with the family stress model.

These models of impact through income are important if we are to understand the potential causal mechanisms behind the findings of this CASEpaper. They were also central in political and parliamentary debates about the HPG: opponents of the HPG emphasised the investment model and criticised the HPG on the basis that it would fail to improve nutrition, since it was paid relatively late in pregnancy. Proponents of the HPG appeared to concede this point, instead relying on the family stress and behavioural models to justify why the grant could support healthy birth outcomes and why the grant was important more generally.

2.3 Regression discontinuity (RD) design

I analyse the impact of the HPG on birthweight in England and Wales using a quasi-experimental regression discontinuity (RD) design. The intuition behind an RD approach is that it helps address the problem of selection bias by mimicking the conditions of a randomised control trial. Simply comparing outcomes between mothers who received the HPG and mothers who did not would lead to selection bias, since there are likely to be systematic differences between women who were likely to take-up the HPG and those who were not. In any case, since take-up was not directly measured, this naïve approach is not possible. However, we do know that the grant had an arbitrary eligibility rule based on due dates: if your due date was 6 April 2009, you were eligible for the grant; if your due date was 5 April 2009, you were not. This creates a sudden (or discontinuous) jump in the probability of receiving the grant at this cut-off that is similar to being randomly allocated to a treatment or control group in an RCT. An RD design exploits this jump to evaluate the impact of the grant without being complicit in selection bias.

First introduced by Thistlethwaite and Campbell (1960), RD designs are applicable in cases in which treatment status is determined by an 'assignment' variable (known as the 'running variable'). Typically, RD designs feature a treatment that is conditional on achieving an arbitrary threshold (a 'cut-off' or 'discontinuity') in the running variable. RD designs create 'as if' randomised allocation into treatment and control groups and can lead to causal estimates that are as consistent as those from randomised experiments when robust bias corrected confidence intervals are used (Hyytinen et al, 2018). On the assumption that it is random whether an individual is located just below or just above the arbitrary cut-off, those just below (the control group) are a valid counterfactual for those just above (the treatment group). If this assumption is validated, then observed differences in outcomes between these groups represent a treatment effect.

There are two main approaches for RD analyses – parametric and non-parametric. Parametric approaches draw on a wide range of observations and make statistical assumptions (including normally distributed data) to impose a shape or functional form on the data, from which point estimates are made. Non-parametric approaches do not make such assumptions and tend to focus on a much smaller subset (or bandwidth) of the data, in this case around the cut-off.

Most of the early contributions to the RD literature relied on parametric methods. However, non-parametric methods have become extremely popular, and for many they are preferable to a global parametric approach (Hahn et al, 2001). Although a parametric approach is useful for providing an overall fit of the data and tends to have more statistical power, a non-parametric approach performs better in approximating the fit around the cut-off – the region in which principles for randomisation are most credible – and is less influenced by outliers.

Relying exclusively on either a parametric or non-parametric analysis is unwise (Lee and Lemieux, 2010; Jacob and Zhu, 2012). Specifically, non-parametric analyses can be problematic in the case of a discrete running variance, such as date of birth, because it becomes necessary to impose a functional form on the data (Lee and Card, 2008). Meanwhile, parametric analyses with higher order polynomials can lead to noisy RD estimates that are sensitive to the order of the polynomial and produce confidence intervals with poor coverage (Gelman and Imbens, 2018). In practice, the distinction between the two approaches has been overstated, and parametric methods can be viewed as equivalent to non-parametric methods, but with a larger bandwidth (Lee and Lemieux, 2010). In this CASEpaper I utilise both parametric and non-parametric methods.

3. Methodology

3.1 Analytical framework

The two birth outcomes explicitly referenced by proponents of the HPG were birthweight and prematurity. While a range of birth outcomes beyond birthweight are important (Conti et al, 2018), the wide-ranging evidence on the impact of birthweight on health in later life, combined with its completeness in the birth registrations microdata, make it a robust choice as the dependent variable. Unfortunately, the absence of a variable for gestational age in the birth registrations microdata prevented the use of prematurity as an alternative outcome.

The policy design of the HPG renders it a promising topic of analysis for a RD design. Receipt of the HPG was determined by an arbitrary cut-off based on expected date of delivery (EDD): all mothers of babies with an EDD on or after 6 April 2009 were eligible for the grant. Since EDDs, proxied by date of birth, are difficult if not impossible to control, the HPG is a particularly compelling case for an RD design. This is because date of birth is plausibly random.

As claiming the HPG was dependent on attending an ante-natal check-up from 25 weeks, it exhibited imperfect compliance. Unfortunately, no official take-up data was recorded, but HMRC expenditure data suggest high take-up rates of 92.9 percent. For the 2009–2010 financial year, HMRC's reported expenditure on HPG grants – excluding administrative costs – totalled £137.8 million in the UK as a whole (HMRC Departmental Accounts 2009–10, p. 55). Given that the grant was a flat-rate lump sum of £190 for all women, it can therefore be estimated that approximately 725,260 women received the grant in 2009-2010. This implies an extremely high take-up rate of the 781,000 maternities that were recorded in the UK in 2009, of approximately 92.9 percent. [6]

This high but imperfect compliance introduces 'fuzziness' in the discontinuity, as the probability of being treated increases sharply from 6 April 2009 but falls short of reaching unity. Somewhat unusually, in my setup the fuzziness of the RD is compounded by another source. Since the birth registrations data unfortunately do not include a due date variable, I use date of birth as a proxy for due date. This increases the fuzziness of the RD and, given the large sample size, can be expected to increase random measurement error. This is because some mothers will have received the HPG despite their baby being born before 6 April, while others will not have received it despite their baby being born afterwards. This measurement error or 'noise' is likely to reduce statistical significance and lead to underestimation of the real treatment effect.

The fact that official take-up rates were not measured renders it impossible to estimate either a 'treatment on the treated' (TOT) or 'local average treatment effect' (LATE) through a fuzzy RD approach. Instead, I estimate an 'intention to treat' (ITT) effect utilising the traditional sharp RD approach. As such, I estimate what would be referred to as the 'reduced form' equation within the instrumental variable approach, between an outcome (birthweight) and an instrumental variable (week of birth) which

[6] England and Wales: Births and maternities: ONS Births in England and Wales; Northern Ireland: Births and maternities: NI Statistics and Research Agency (NISRA) Registrar General Annual Report table 3, various years; Scotland: Births: National Records of Scotland Vital Events Reference Tables 2017, Sections 3 and 4.

serves as a proxy for an explanatory variable (receipt of the HPG). I use week of birth, centred around the cut-off, instead of date of birth in order to maximize sample size.[7]

My basic empirical specification is therefore the following:

$$Y_i = \beta_0 + \beta_1 HPG_i + \beta_2 WOB_i + \varepsilon_i$$

, where Y_i is the outcome of choice (birthweight in grams, or a dummy variable equal to one if a baby weighs less than 2500g and zero otherwise), $\beta_1$is the ITT effect of the HPG, HPG_i is a dummy variable equal to one if an individual was eligible and zero otherwise, WOB_i is week of birth, and i indexes individual births.

While arbitrary cut-offs were used both in the introduction and the abolition of the HPG (Howard, 2011, January 10), I focus exclusively on the introduction cut-off. The abolition of the HPG in January 2011 coincided with the restriction of the Sure Start Maternity Grant to the first child, which would make it more difficult to isolate the effect of the HPG's abolition.

3.2 Data and descriptive statistics

I use birth registrations microdata in England and Wales from 2006 to 2012. As an administrative source of data, it benefits from an extremely large sample size, with 1,402,979 observations (live births) between 6 April 2008 and 5 April 2010 (a two-year period around the treatment cut-off). Data was cleaned to drop duplicates, stillbirths (n=25,590) and implausible birthweight outliers (n=47,516).[8] The data enjoy a complete set of non-missing values for date of birth, birthweight and sex. Social class status was coded for a random sample of 10 percent of the population. The data include the postcode of mother's residence for all but 1,184 (0.08 percent) of the final sample, which was used to match in the English index of income deprivation at Lower Super Output Area (LSOA) level.[9]

---

[7] In the parametric models, in line with the empirical specification shown the running variable (date/week of birth) is included in the regression as a control, so using date of birth leads to a smaller number of observations from which to draw (see Appendix Table A14).

[8] After inspecting the birthweight distribution, outlier observations with a recorded birthweight less than 310 grams or larger than 5680 grams were dropped. This included missing values coded as '9999' or '9998'.

[9] Since the indices of multiple deprivation are not comparable across England and Wales (with the exception of a combined dataset for 2015-16), all analyses using deprivation data were for England only. Index of income deprivation scores from the nearest data point available to the introduction of the HPG, 2010, were matched into the births data at LSOA level for all observations with a non-missing postcode (99.92 percent of the sample).

Key covariates in the final sample are sex of the baby, maternal age, multiple birth status, English index of income deprivation score and combined occupational social class (for 10 percent of observations). Males tend to have slightly higher birthweights than females (Nascimento et al, 2017). While the evidence is mixed, maternal age has been found to have a (non-linear) relationship with birthweight (Goisis, Schneider and Myrskylä, 2018). Younger mothers tend to have smaller babies on average and a higher incidence of low birthweight (Ghosh et al, 2017). Multiple births are more prone to pre-term delivery and low birthweight (Blondel et al, 2002), and are particularly relevant as the HPG was administered per pregnancy, not per birth (WiredGov, 2009, January 20). Consequently, mothers experiencing multiple births did not receive as much money per child, and a higher multiple birth rate is likely to attenuate the impact of the HPG. Birthweight is also patterned by social class: at the introduction of the HPG in 2009, babies born to parents from routine, manual and other occupations were 13 percent more likely to be low birthweight than babies born to parents in professional, managerial, intermediate or small employer occupations (Stewart and Reader, 2021). The index of income deprivation – which measures the proportion of the population in a given area who experience deprivation relating to low income, whether due to unemployment or low pay (Smith et al, 2015) – is also associated with low birthweight (Dibben et al, 2006). These covariates are not used as controls, since the RD design renders that redundant and potential outcomes are quasi-randomised at the discontinuity (Lee and Lemieux, 2010). However, they are inspected using RD methods in Section 5.3 as a robustness check.

Table 1 shows summary statistics for the final sample of births from 6 April 2008 to 5 April 2010. Treatment and control groups were balanced in terms of sex composition, teenage and advanced maternal age pregnancies, multiple births and lone parents, and broadly balanced for maternal age (with the mean for the treatment group being 0.04 years higher than the control group). However, mean birthweight in the treatment group was 5 grams larger than in the control group, while low birthweight was 0.2 percentage points lower and extremely low birthweight was 0.1 percentage points lower.

Table 1 Summary statistics for the final sample

	Main sample (2008-2010)		Treatment group		Control group	
Variable	Mean	SD	Mean	SD	Mean	SD
Birthweight (grams)	3,336.46	598.52	3,339.14	597.53	3,333.75	599.51
Prop. low birthweight	0.071	0.257	0.070	0.255	0.072	0.258
Prop. extremely low birthweight	0.011	0.106	0.011	0.105	0.012	0.107
Maternal age (years)	29.45	6.07	29.43	6.08	29.47	6.05
Prop. teenage pregnancy	0.062	0.241	0.060	0.238	0.063	0.244
Prop. advanced maternal age	0.200	0.400	0.200	0.400	0.200	0.400
Prop. multiple births	0.031	0.173	0.031	0.174	0.031	0.172
Prop. male	0.512	0.500	0.512	0.500	0.512	0.500
Prop. lone parents	0.158	0.365	0.158	0.365	0.157	0.364
English index of income deprivation score	0.175	0.122	0.175	0.122	0.175	0.123
N	1,402,979		697,595		705,384	

3.3 Regression discontinuity (RD) design

- Parametric RD analysis

In the parametric part of the RD analysis, a wide range of observations (up to 24 weeks either side of the cut-off) is utilised. Various functional forms, including polynomials of different degrees and interaction terms, are fitted to a scatter plot of the data and an ITT effect is estimated. I evaluate the goodness of fit of these specifications by plotting the data and by minimising the Akaike information criterion (AIC) and Bayesian Information Criterion (BIC). Inference is conducted using robust bias-corrected standard errors that are clustered by week of birth.

- Non-parametric RD analysis

While the parametric analysis draws on a wide range of data to estimate the causal impact of the HPG, in the non-parametric analysis estimation and inference are restricted to a specified bandwidth *h* of observations close to the cut-off.

Non-parametric approaches can be problematic when using a discrete running variable such as date/week of birth, because the data is not continuous, and it would be easy to misinterpret discontinuities in the data (Lee and Card, 2008). However, in my case this is aided by the large sample size (n=1,402,979) and number of mass points (n=730) (Cattaneo et al, 2018b). I use standard errors that are clustered by week of birth, a technique proposed by Lee and Card (2008) to combat the discreteness problem when specification errors can be assumed to be identical regardless of treatment status. As in the case of Card and Shore-Sheppard's (2004) age-based RD, this assumption is likely to apply in the case of the HPG. Finally, I also make sure to compare non-parametric results with parametric ones.

The key decision in a non-parametric RD analysis is how to select the bandwidth around the cut-off, a decision informed by the traditional trade-off between bias and precision. A narrower bandwidth reduces bias but, as it reduces the number of observations, it increases the variance of the estimator. Conversely, a wider bandwidth is more prone to bias because other factors are more likely to be confounding the estimated effect, but the variance of the estimator is lower. Since it is inadvisable to select bandwidths *ad hoc*, I select my bandwidth in a data-driven way, by a cross-validation procedure that minimises the mean squared error (MSE) of the estimator (Cattaneo et al, 2018a; Imbens and Lemieux, 2008; Lee and Lemieux, 2010). In this minimisation, I use a linear polynomial in order to avoid overfitting the data, which can interfere with estimates at the cut-off (Cattaneo et al, 2018a). I also use a triangular kernel function, such that the weight of observations slopes away from the cut-off, since it is MSE-optimal and arguably best captures the spirit of regression discontinuity by attaching a greater weight to observations closer to the cut-off (Cattaneo et al, 2018a; Fan and Gijbels, 1996).

Since the MSE is the sum of the squared bias and squared variance, minimising the following objective function enables an optimal balance between bias and variance to be achieved:

$$CV(h) = \min(MSE) = \frac{1}{N}\sum_{i=1}^{N}\left(Y_i - \hat{Y}_i(WOB_i)\right)^2$$

, where CV is cross-validation, h is the bandwidth, MSE is the mean squared error, Y_i is the observed outcome, $\hat{Y}_i$ is the predicted outcome, WOB_i is week of birth, and i indexes an individual birth.

After a point estimate is made, non-standard inference procedures are necessary to establish statistical significance. OLS methods of establishing statistical significance are invalid in this context, since they do not account the possibility of misspecification error in the local polynomial fit (Cattaneo et al, 2018a). I use robust bias-corrected confidence intervals. This produces smaller confidence intervals than conventional inference (based on parametric least-squares assumptions) or standard bias correction, and has been found to be the only RD inference procedure that replicates experimental estimates (Hyytinen et al, 2018). By contrast, RD estimates using conventional confidence intervals or standard bias correction can lead to overestimation of effects (Hyytinen et al, 2018). However, since many RD papers use these approaches, I also report results using these alternative confidence intervals in the Appendix in Table A2, and they do indeed suggest even higher and statistically significant effects than in my preferred models.

5. Results

This section outlines the main results of my analysis. First, I estimate an RD treatment effect for the HPG using both parametric and non-parametric approaches. Second, I investigate whether the gains of the grant were larger for certain groups than others, namely by birthweight, maternal age and English index of income deprivation. Third, I conduct three robustness checks to validate the use of the RD methodology: placebo cut-off dates, a covariate balance test, and a McCrary manipulation test.

5.1 Impact on birthweight and low birthweight status (<2500g)

Table 2 and Table 3 summarise estimated treatment effects of the HPG on birthweight and low birthweight (<2500g) respectively using a selection of parametric and non-parametric models.

Table 2 RD estimates of the impact of the Health in Pregnancy Grant on birthweight (grams)

	(1)	(2)	(3)	(4)	Control mean
	10.72***	10.39***	10.17***	10.38***	3323.67
	(3.291)	(2.778)	(2.593)	(2.664)	
Observations	67,553	425,620	425,620	425,620	
MSE-optimal h	X				
Robust bias-corrected	X				
Linear	X	X			
Linear interaction			X		
Quadratic				X	

Notes: Standard errors, clustered by week of birth, in parentheses. Bandwidth refers to the size of the region either side of the cut-off (e.g., for a bandwidth of 16 weeks, the total region covered by the non-parametric analysis is 32 weeks). Model 1 is a non-parametric RD specific using the MSE-optimal bandwidth for each variable, with robust bias-corrected standard errors. Models 2-6 are parametric specifications all for a 16-week bandwidth around the cut-off: (2) is a linear polynomial; (3) is a linear interaction (ie spline), (4) is a quadratic polynomial. Preferred model is (1), the non-parametric model. Full results in Appendix Table A1 and Table A2. * p<0.1. ** p<0.05. *** p<0.01.

Table 3 RD estimates of the impact of the Health in Pregnancy Grant on low birthweight status (<2500g)

	(1)	(2)	(3)	Control mean
Low birthweight	-0.004	-0.002	-0.002	0.072
	(0.002)	(0.002)	(0.002)	
Observations	120,477	639,171	639,171	
MSE-optimal h	X			
Robust bias-corrected	X			
Probit		X		
Logit			X	

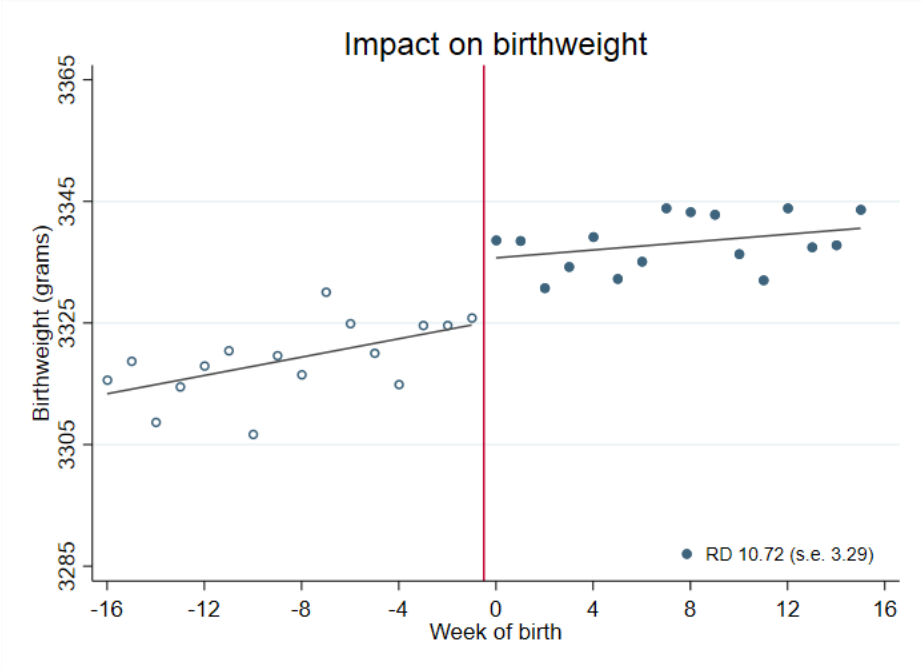
Notes: Standard errors, clustered by week of birth, in parentheses. Bandwidth refers to the size of the region either side of the cut-off (e.g., for a bandwidth of 24 weeks, the total region covered by the non-parametric analysis is 48 weeks). Model 1 is a non-parametric RD specific using the MSE-optimal bandwidth for each variable, with robust bias-corrected standard errors. (2) is a parametric probit model; (6) is a parametric logit model. Preferred model is (1), the non-parametric model. Full results in Appendix Table A2. * p<0.1. ** p<0.05. *** p<0.01.

There is a small positive impact on birthweight that is robust and statistically significant across almost all models, across parametric and non-parametric approaches, different bandwidths and polynomial specifications (see Appendix Table A1 and Table A2 for full results). My preferred model is (1), a non-parametric local linear model for an MSE-optimal bandwidth, with robust bias-corrected standard errors. This model suggests that the HPG led to an 11g (1.8 percent of a standard deviation, robust 95% confidence interval 4.27 to 17.17g) increase in birthweight on average that is statistically significant at the 1 percent level.

My results suggest that there was no significant impact on low birthweight at population level. While all parametric and non-parametric models have negative coefficients, they fail to reach significance for the most part. It is worth noting, however, that significant effects are reached when conventional confidence intervals are utilised in the non-parametric approach, an approach that many RD papers continue to use despite evidence that they can lead to overestimation of treatment effects (Hyytinen et al, 2018).

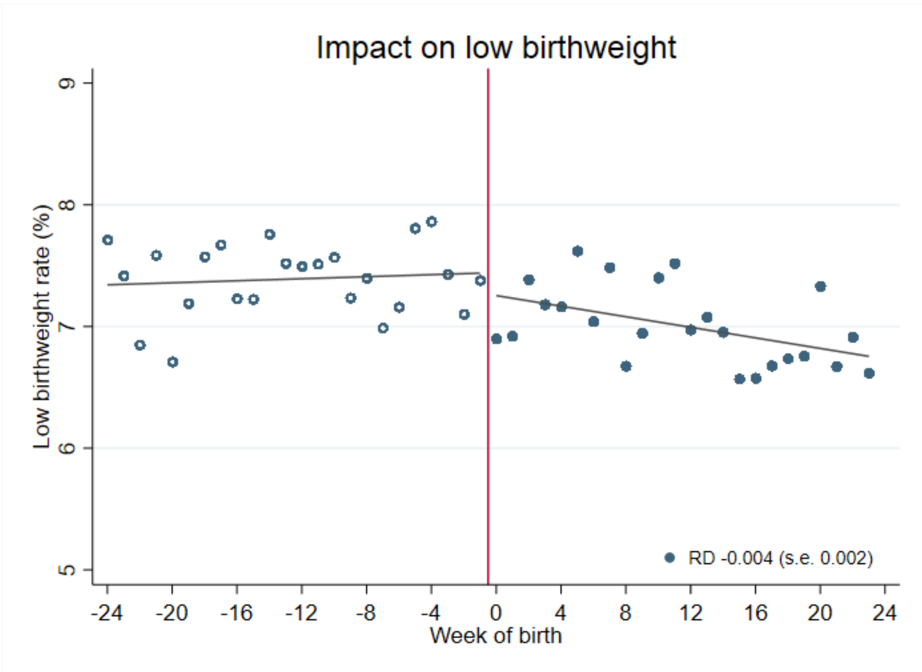
Figure 1 and Figure 2 illustrate these treatment effects graphically. In both cases, a linear specification appears appropriate. Figure 1 shows a clear but small discontinuous positive jump in mean birthweight at the cut-off. In line with my regression results, Figure 2 shows a much less clear drop in the low birthweight rate at the cut-off.

Figure 1 Impact of the Health in Pregnancy Grant on birthweight



Notes: 1. Scatterplot shows collapsed mean birthweight data by week of birth with a fitted linear spline around the cut-off. 2. A bandwidth of 16 weeks (preferred parametric bandwidth) is used for graphical purposes.

Figure 2 Impact of the Health in Pregnancy Grant on low birthweight status (<2500g)



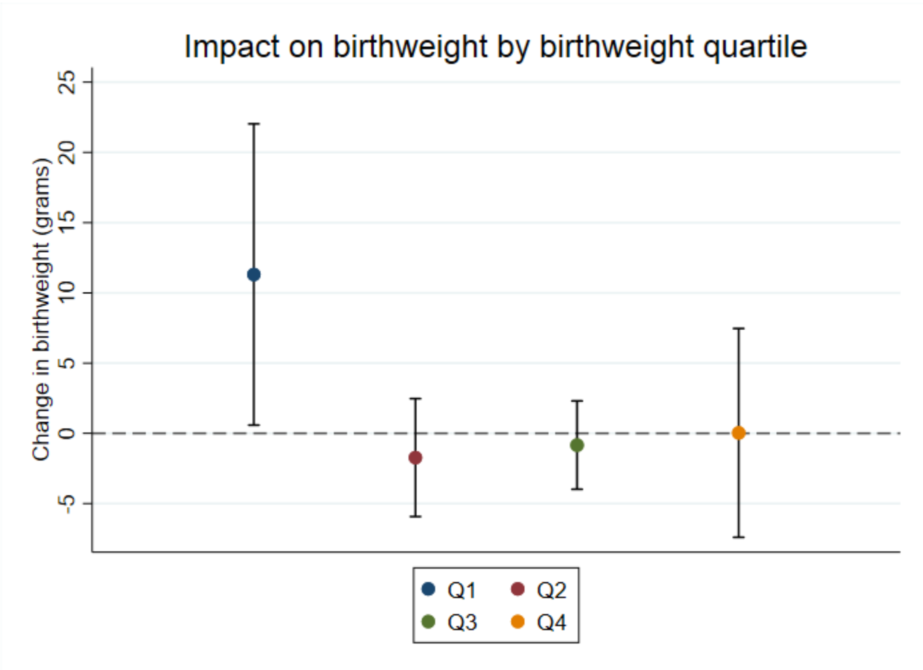
Notes: 1. Scatterplot shows collapsed low birthweight data by week of birth with a fitted linear spline around the cut-off. 2. A bandwidth of 24 weeks (preferred parametric bandwidth) is used for graphical purposes.

5.2 *Heterogenous effects by birthweight, maternal age and income deprivation*

Having established that the HPG led to a positive increase in birthweight on average across the population as a whole, I now investigate whether the gains of the grant were distributed differentially across groups.

Birthweight treatment effects are concentrated at the bottom of the birthweight distribution. When conducting separate RD analyses by birthweight decile, it is only the bottom decile of the birthweight distribution for whom there is a significant positive treatment effect, with this group seeing a 16g increase (3.2 percent of a standard deviation for this group, 95% confidence intervals: -0.70 to 31.87g) in birthweight at the cut-off. This bottom decile has a birthweight of 2279 grams on average, well below the low birthweight threshold of 2500 grams. When using birthweight quartiles, again it is only the bottom birthweight quartile that shows a significant effect of 11g (2.6 percent of a standard deviation for this group, 95% confidence intervals: 0.59 to 22.03g), and in this case it reaches significance at the 5 percent level due to the greater sample size of quartiles compared to deciles (see Figure 3 and Table A4). This clearly shows that the gains of the grant were concentrated among the smallest babies, particularly those at risk of low birthweight.

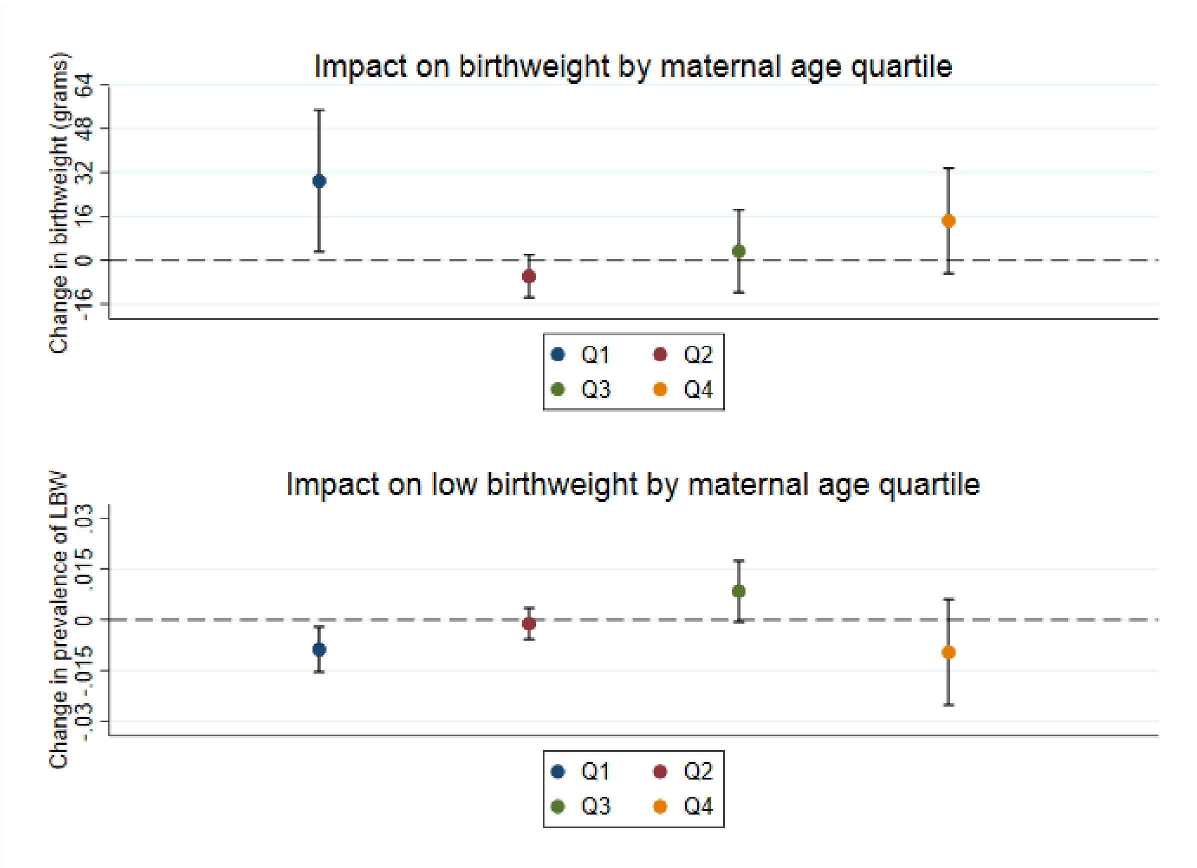
Figure 3 Heterogenous treatment effects of the HPG across the birthweight distribution



Notes: 1. Each point estimate represents estimates from a separate RD regression for each birthweight quartile, using an MSE-optimal bandwidth with robust bias-corrected standard errors (clustered by week of birth). 2. Birthweight deciles and quartiles are calculated among singleton births only (i.e., excluding multiple births). 3. Bars show 95% confidence intervals. 4. Full results in Appendix Table A3 and Table A4.

There is consistent evidence from my preferred and non-parametric models that the grant had a larger impact on teenage and young mothers (see Appendix Table A5 and Table A6).[10] This is noteworthy given that younger mothers are associated with significantly lower birthweights on average (see Appendix Figure A1), and a relatively high incidence of low birthweight (see Appendix Figure A2). As Figure 4 shows, the relationship between effect size in my preferred model and maternal age is U-shaped, with mothers in the bottom quartile of the age distribution being the only group with statistically significant effect sizes. Mothers in the bottom age quartile (25 years and under) see a 29g increase in birthweight (5 percent of a standard deviation for this group; 95% confidence intervals: 3.10 to 54.69g), a point estimate that is higher than that for the population as a whole.

Figure 4 Heterogenous treatment effects by maternal age



Notes: 1. Each point estimate represents estimates from a separate RD regression for each maternal age quartile, using an MSE-optimal bandwidth with robust bias-corrected standard errors (clustered by week of birth). 2. Bars show 95% confidence intervals. 3. Full results by maternal age quartile and decile in Appendix Table A5, Table A6 and Table A7.

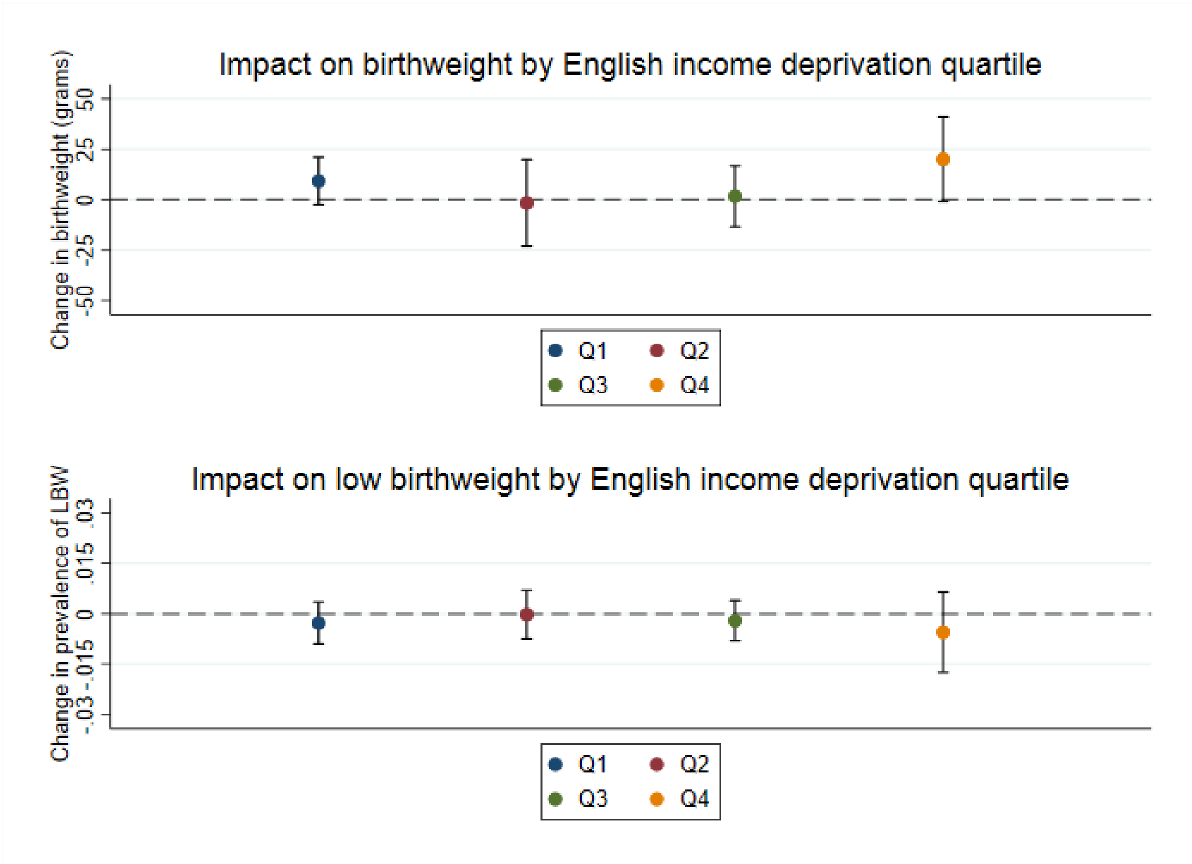
[10] Parametric models are less supportive of heterogenous effects across birthweight, maternal age and income deprivation. The parametric models should be treated with caution here, however, since the parametric assumption of normally distributed data is less likely to hold for these non-random subgroups of the population.

Furthermore, while the HPG did not reduce low birthweight for the overall population, it did significantly reduce the probability of low birthweight for younger mothers aged 25 or under, by 0.9 percentage points (95% confidence intervals: -0.015 to -0.002) or 12 percent relative to this group's mean low birthweight rate of 7.4 percent (see Appendix Table A7).

Matched data for births in England on the index of income deprivation indicate that the impact of the HPG was larger for mothers who live in more deprived areas.[11] Differences in effect size by deprivation quartile are more muted than for maternal age, likely due to the ecological fallacy and the limitations of the index as a predictor of individual-level socio-economic status or income (see Appendix Table A8). But as Figure 5 demonstrates, it is telling that coefficients only reach statistical significance for the most deprived quartile (Q4), for whom there is a positive treatment effect of 20 grams (3.4 percent of a standard deviation for this group; 95% confidence intervals: -0.98 to 41.04 grams), statistically significant at the 10 percent level. While none of the estimates for low birthweight by deprivation quartile are statistically significant, point estimates are larger for the most deprived quartile.

[11] Indices of income deprivation are not comparable for England and Wales, so analysis by deprivation was conducted for births in England only. NS-SEC occupation is recorded in the birth registrations data for 10 percent of observations only, so sample sizes were too small for meaningful inference.

Figure 5 Heterogeneous treatment effects by English index of income deprivation



Notes: 1. Results are for England only. 2. Each point estimate represents estimates from a separate RD regression for each English index of income deprivation quartile using an MSE-optimal bandwidth with robust bias-corrected standard errors (clustered by week of birth). 3. Bars show 95% confidence intervals. 4. Full results in Appendix Table A8 and Table A9.

5.3 Testing the validity of the RD methodology

Having presented the main findings of the RD analysis, I now conduct three robustness tests in order to test the validity of the RD methodology in this context. First, I conduct placebo RD tests for the treatment date (6 April) in four different years, both before and after the HPG was implemented. Second, I inspect baseline covariates to test whether there were any statistically significant discontinuities at the cut-off. Third, I investigate the possibility of due dates (proxied by date of birth) being manipulated.

- Placebo test

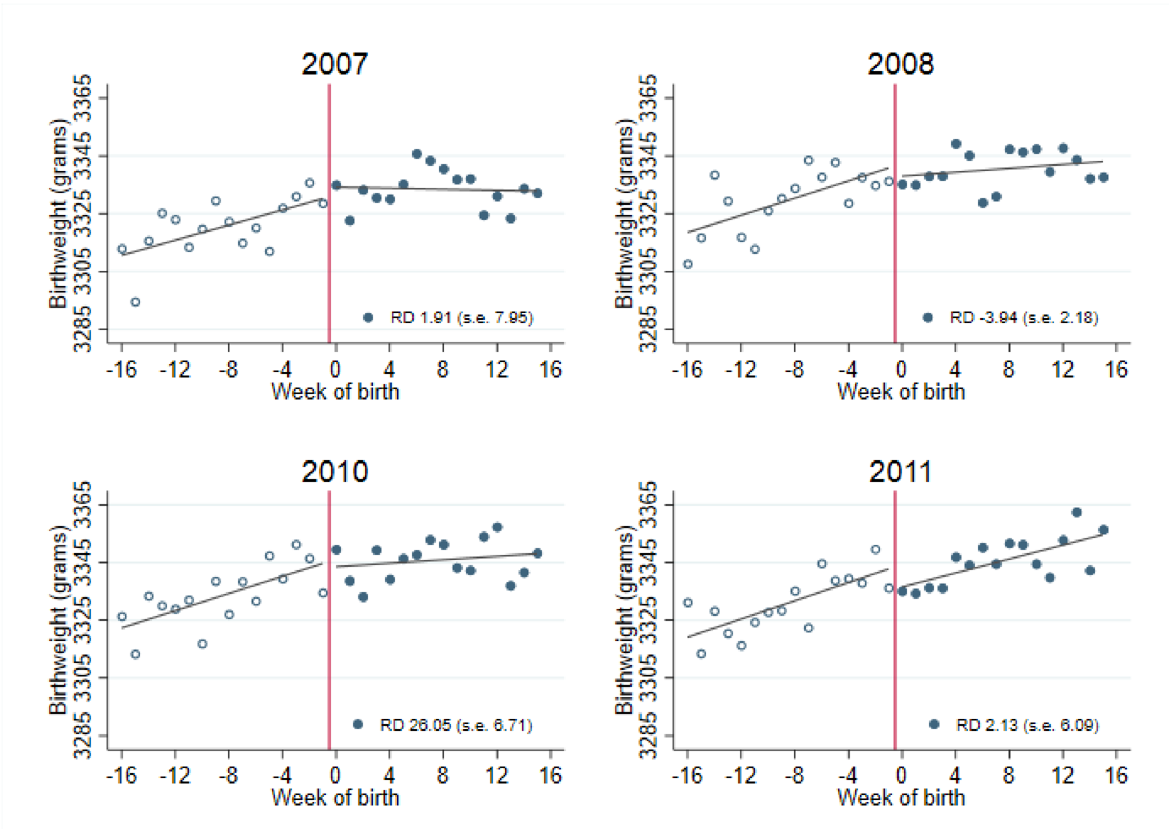
Given the tendency of birthweight to fluctuate throughout the year by season and its natural incline in April, it is important to test whether the observed increase in birthweight at the treatment cut-off is due to (discontinuous) seasonal variation. In order to do this, I create placebo cut-offs for two years before the grant was introduced (6 April 2007 and 6 April

2008), and for two years after the grant was introduced (6 April 2010 and 6 April 2011). Full results from these placebo tests are shown in Appendix Table A10 and Table A11.

There is no compelling evidence for a positive statistically significant jump in birthweight at any of these placebo cut-offs. All but one of the coefficients do not reach statistical significance, and several are negative. While 2010 demonstrates a comparable positive effect of 26g in the non-parametric analysis (as shown on Figure 6), the graphical evidence is weak and none of the parametric models corroborate that finding or achieve statistical significance at any level. The lack of robustness for these placebo cut-offs, therefore, does not compare favourably with the highly robust results at the real treatment cut-off, and supports the validity of my RD methodology.

Figure 6 illustrates these placebo cut-off checks graphically. They contrast sharply with the treatment graph in Figure 1, which by contrast shows a very clear positive discontinuity at the cut-off.

Figure 6 Robustness checks using placebo cut-off eligibility dates



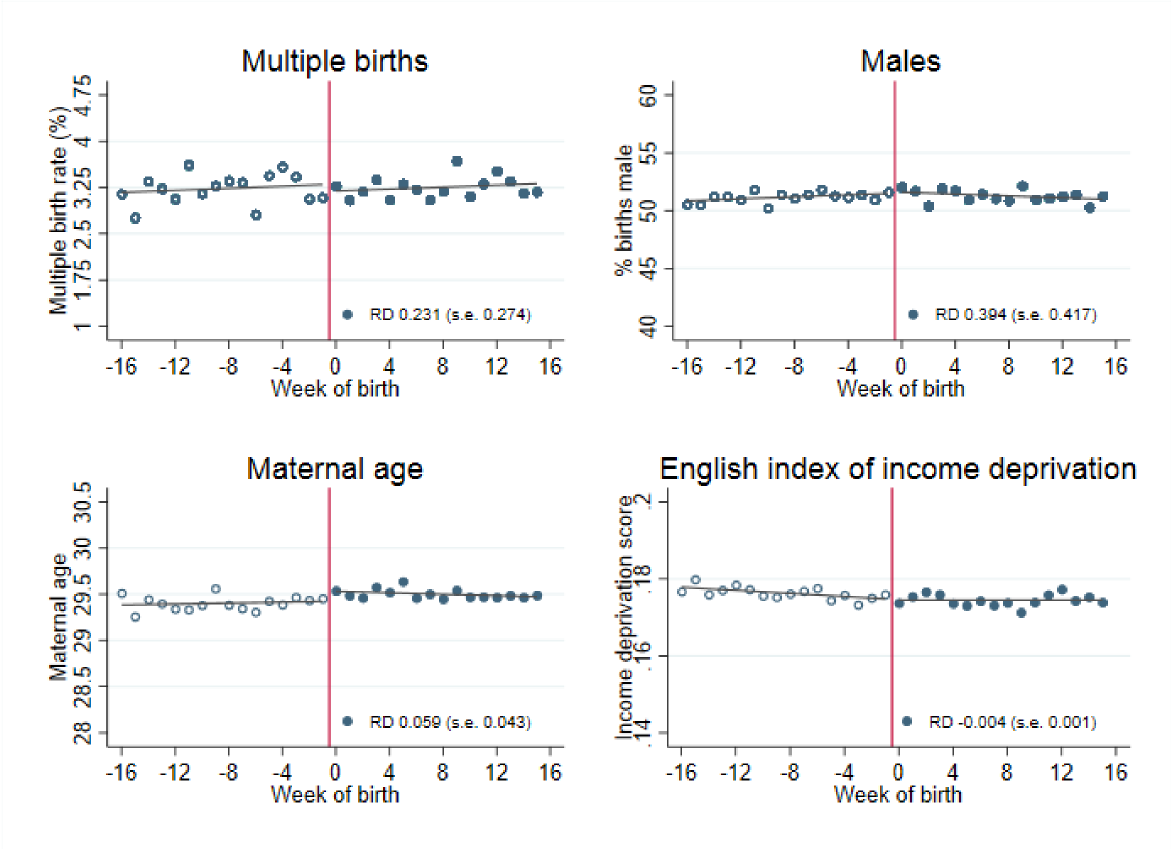
Notes: 1. Scatterplot shows collapsed mean birthweight data by week of birth. 2. A bandwidth of 16 weeks (preferred parametric bandwidth in the main RD analysis) is used for graphical purposes and for consistent comparison with the treatment graph in Figure 1.

- Covariate balance test

To check that the positive jump in birthweight is not attributable to a discontinuity in one of the predetermined covariates of birthweight, I also conduct RD tests for these variables. In other words, I re-estimate my main empirical specification, but instead of using birthweight as the outcome, I use maternal age, multiple birth status (where a pregnancy results in more than one child being born, such as twins or triplets), being male, and the English index of income deprivation. Full results are presented in Appendix Table A12; Figure 6 illustrates these graphically for my preferred models.

Multiple births and males are balanced around the cut-off. Parametric models suggest a small positive jump in maternal age, but they are not supported by my preferred model (non-parametric local linear polynomial with robust bias-corrected standard errors). To rule out any interference of maternal age, however, I also test whether maternal age exhibits a positive discontinuity on 6 April in the placebo years – 2007, 2008, 2010 and 2011 (see Table A13). Maternal age exhibits a mild positive discontinuity in some parametric models at the placebo cut-offs in 2010 and 2011, in a strikingly similar pattern to the treatment cut-off. This suggests that any discontinuity in maternal age at the cut-off is unlikely to be connected to the HPG. However, since treatment effects appear to be larger for younger mothers, it is possible that any such an increase in mean maternal age may have caused some minor downwards bias in my estimation of the impact of the HPG. English index of income deprivation scores meanwhile see a slight negative discontinuity of 0.4 percentage points in my preferred non-parametric model. However, this result is not consistent across other non-parametric models or parametric models. The index is also arguably less likely to prove a barrier to identification since it is measured at geographical rather than individual level.

Figure 7 Covariate balance tests as a robustness check



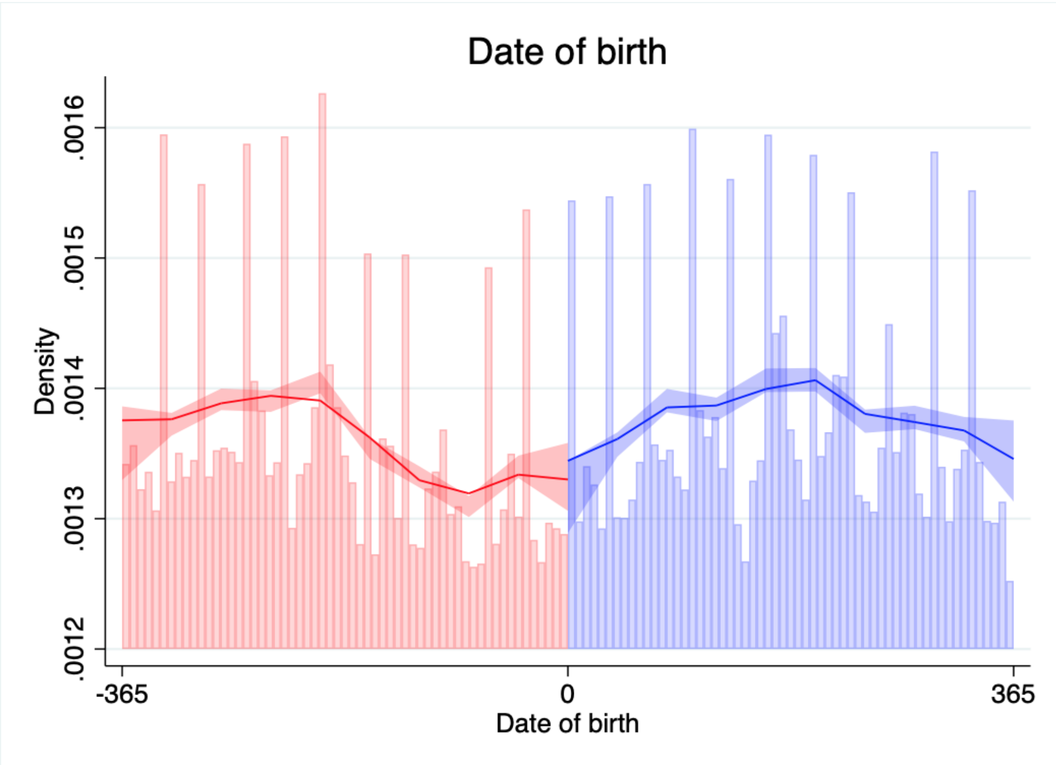
Notes: 1. Multiple births are those where more than one child is born as a result of a pregnancy (e.g., twins, triplets). 2. English index of income deprivation data is matched to the data using mother's postcode of residence. 3. Scatterplot shows collapsed data for each variable by week of birth. 2. A bandwidth of 16 weeks (preferred parametric bandwidth in the main RD analysis) is used for graphical purposes and for consistent comparison with the treatment graph in Figure 1.

- 'McCrary' manipulation test

Since the identifying assumption of RD designs is that individuals lack precise control over the running variable in question, the ability of pregnant mothers to manipulate dates of birth would be a clear impediment to the validity of the RD. Institutional background knowledge is essential to validate this assumption (Eggers et al, 2018). In the case of the HPG, the main source of manipulation of the running variable would be the postponement of conception in order to qualify for the grant. The first general discussion of eligibility requirements that mentions the approximate cut-off of April 2009, according to newspaper reports, appears to have been September 2007 (BBC News, 2007, September 8). However, no specific cut-off date was given, and it was not stated what point of gestation would have to been reached by April 2009 to be eligible. The possibility that prospective mothers had *precise* control over eligibility and the running variable is, therefore, remote.

To empirically verify this, I conduct a McCrary density test for the running variable. If manipulation had taken place, we can assume that there would be a 'bunching' of observations to the right of the cut-off, as women would be incentivised to postpone conception or birth until after the cut-off to be eligible for the grant (McCrary, 2008). I use date of birth as the most unsmoothed version of the running variable, and the one that in practice women could be expected to have control or manipulation over. The formal McCrary test fails to reject the null hypothesis of no discontinuity in the density at the cut-off: there is no significant evidence of manipulation or bunching (see Appendix Table A15 for full results for date of birth and week of birth). As Figure 8 shows, there is some regular variation in density of births, with spikes every Monday (including the cut-off date of Monday 6 April 2009). This is likely due to the managed scheduling of induced labour and elective caesarean sections during the week when services are at their highest capacity. The visible increase in the number of births at the cut-off is therefore likely to be attributable to a surge in births after the weekend, and the McCrary test confirms that this does not represent a discontinuity relative to the rest of the year.

Figure 8 Density plot of the running variable (date of birth)



Notes: 1. Histogram shows density of births by date of birth. 2. Overlapping confidence intervals at the cut-off indicate a failure to reject the null hypothesis of no discontinuity at the cut-off (i.e., no manipulation). 3. The daily variation in frequency of births and the notable spikes in the graph are likely to be due to the managed scheduling of induced labour and elective caesarean sections during weekdays, with the consequence that Mondays have the highest numbers of births and the weekends have the lowest.

6. Discussion

The results of my quasi-experimental RD analysis suggest that the Health in Pregnancy Grant was responsible for increases in birthweight of approximately 11g on average. A significant positive effect on birthweight is shown to be robust across parametric and non-parametric approaches, different orders of the polynomial and different bandwidths. This contrasts sharply with RD analyses of placebo cut-offs and baseline covariates – which demonstrate a lack of magnitude in the estimated effect, a lack of statistical significance and a lack of robustness. Both institutional context and a formal McCrary test support that manipulation of the running variable was unlikely.

A small positive jump in birthweight for the general population in the region of 11g (1.8 percent of a standard deviation) from a £190 cash transfer is small but impressive relative to other studies' estimates of the impact of income on birthweight. Quasi-experimental evidence from the US on the impact of a $1000 (2000 prices) increase in EITC income suggested an increase in mean birthweight of 10g for a high-impact sample of single low-educated mothers (Hoynes, Miller and Simon, 2015).

My analysis also shows that babies in the bottom decile of the birthweight distribution benefitted the most from the grant: this group, many of whom are at risk of low birthweight, saw a larger 16g (3.2 percent of a standard deviation) boost to birthweights on average. In fact, these are the only group within the birthweight distribution who saw a significant positive impact. This is consistent with other studies which have found that birthweight effects due to increases in income are strongest at the bottom of the birthweight distribution and very weak at the top (Hoynes, Miller, and Simon, 2015; Almond, Hoynes and Schanzenbach, 2011). This suggests that even though the grant was universal and unconditional, its effects were targeted on those babies most in need of increases in birthweight. For the bottom birthweight decile, an effect size of 16g represents a sizeable proportion of their total weight and a significant contribution towards improved health and weight at birth. Although this result was to be expected given the assumption of diminishing marginal returns to birthweight, it is hypothetically possible that the grant could have subsidised unhealthy eating and gestational diabetes, which can lead to foetal macrosomia (much larger than average birthweight) which is in turn associated with health risks to both mother and baby. The fact that the grant only led to increases in birthweight for small babies lends further support to the hypothesis that that the grant facilitated healthy eating and/or reduced stress in ways that benefitted those most at risk of low birthweight.

It is very interesting that the grant appears to have had a larger effect for babies with young mothers, of almost three times the magnitude of the overall birthweight treatment effect according to my preferred specification. Young mothers, particularly teenagers, are not only at heightened risk of low birthweight, but were a group of particular policy focus for the Labour government, with a ten-year strategy on teenage pregnancy having been set up in the decade prior to the HPG and a consistent emphasis on support for younger parents.

The higher effect for younger mothers (25 and under) could be driven by a number of mechanisms. First, younger mothers tend to have later engagement with antenatal health services, and the linking of the HPG to an antenatal check-up at 25 weeks may have improved this. A recent government publication found that 20.4 percent of women aged under 25 attended their antenatal appointment after 13 weeks of gestation, compared to 16.1 percent of women aged 25 or over (Public Health England, 2019). The linking of the HPG to seeking antenatal health advice at 25 weeks may have increased young mothers' access to information about staying healthy during pregnancy, thereby having a larger impact on birth outcomes than for other age groups. There is also some evidence that younger mothers can be more sensitive to 'learning effects' associated with cash transfers (see Sosa-Rubí et al, 2011, on Oportunidades/Progresa in Mexico). Other studies looking at the impact of income on birthweight have shown larger effect sizes for low-educated and low-skilled mothers (Hoynes, Miller and Simon, 2015; Moca et al, 2015) and for teenage mothers (Bitler and Currie, 2005). Younger mothers may therefore have been more responsive to the health advice and information surrounding the introduction of the HPG.

Second, younger mothers are likely to have lower incomes than their older counterparts, meaning that the flat-rate HPG is likely to represent a larger proportion of their budget constraint. This may have led to higher reductions in stress and/or improvements in nutrition. Third, older mothers may have a larger number of possibilities for how to spend the HPG other than health during pregnancy – they may have quickly absorbed the HPG into their overall budget constraint (to help with mortgage payments or to feed other children) and not have set it aside specifically for healthy food or lifestyle choices to improve their own health in pregnancy. Younger mothers may have been comparatively better placed to do so, particularly if it is their first birth and they do not have any other children. Fourth, since there is no official data on take-up, it is not possible to rule out the possibility that younger women were more likely to claim the grant and therefore to benefit from it. This possibility is unlikely to provide the explanation, as the grant only required one-off attendance at a GP or midwife check-up, and engagement with health services tends to be higher with older women (DCSF, 2009). Nevertheless, take-up of the (means-tested) US Special Supplemental Nutrition Program for Women, Infants and Children (WIC) has been shown to be higher among younger women

(Choniy, Currie and Sonchak, 2020), and the same could apply to the HPG, even though it is universal.

Despite the overall positive treatment effect on birthweight for small babies and the reduced probability of low birthweight for young mothers, this did not to translate into a significant reduction in low birthweight for the population as a whole. This threshold is admittedly an arbitrary one of heuristic value, and the most salient outcome for babies who are at risk of low birthweight is their final birthweight rather than the fact of whether they satisfy a threshold condition. Many studies also focus on 'high-impact' samples rather the general populations. But it is possible that the HPG's relatively late payment in pregnancy, and its relatively small financial value (the means-tested Sure Start Maternity Grant was and continues to be £500, by comparison), may have inhibited its impact on low birthweight at population level.

The main takeaway of this research, however, is that the HPG was remarkably effective at increasing mean birthweight and reducing the probability of low birthweight for young mothers, especially when taking into consideration its financial cost. While direct comparison with subgroups in other studies is not possible, the effect sizes are favourable compared to studies of much higher income increases. The largest effect sizes from the $1000 EITC boost in Hoynes, Miller and Simon (2015) were 18g (for Black mothers); meanwhile the largest effect sizes for the HPG in this paper are 29g (for young mothers). How did the HPG manage to achieve such comparatively large effect sizes despite its lower financial value?

There are three possible explanations of the HPG's larger effect sizes relative to other studies' estimates. First, the HPG was tied to seeking antenatal health advice at 25 weeks of pregnancy, unlike the EITC and other windfall cash interventions. It is not possible to determine from birth registrations data whether the HPG was associated with improved engagement with antenatal care, but it is possible given findings to this effect in Scotland (Leyland et al, 2017). Further research with Hospital Episode Statistics will investigate this within an English context. Second, unlike the other studied windfall income increases, the HPG was clearly labelled as a policy that was intended for mothers to use to improve health during pregnancy. Research has suggested that the impact of child benefit and child tax credit on parents' spending on healthy fruit and vegetables, books and toys may be in part due to the labelling of the benefits as designated for children; the same labelling effect may apply here with the HPG and healthy behaviours during pregnancy (Kaushal, Gao and Waldfogel, 2007). Third, unlike the EITC in the US, the HPG was universal and therefore less likely to attract stigma. This may have improved take-up rates among disadvantaged and younger mothers.

There are a number of limitations of my analysis. First, the data does not contain information on nutrition or stress, so it is difficult to draw substantive conclusions about the mechanisms behind the grant's impact other than those discussed above. Second, since the data does not include due dates, there is fuzziness brought about by using date of birth as a proxy for due date. While this is not a barrier to the regression discontinuity design as such, there are a number of spill overs between treatment and control group that detract from the quasi-random nature of the RD. Babies with a due date before 6 April 2009 (who were therefore ineligible for the HPG) but who were born late, after 6 April, are included in the treatment group even though they could not have received treatment. Likewise, babies with a due date after 6 April 2009 (who were therefore eligible for the HPG) but born early, before 6 April, are included in the control group, even though they had a high probability of receiving treatment. Such inconsistencies between due dates and dates of birth introduce additional noise into the estimation of a treatment effect and are likely to reduce statistical significance. In addition, since gestational age is not recorded at birth registration, I was unable to exclude premature babies born before twenty-five weeks from the treatment variable, who would have been ineligible for the HPG as payment occurred from twenty-five weeks. These factors are likely to introduce downwards bias in the estimation of the effect of the grant, and as such my estimates could be considered as underestimates. Further research using Hospital Episode Statistics will enable precise calculation of due dates to be conducted, and thus alleviate this problem.

Finally, the birth registrations data do not facilitate a holistic analysis of the impact of the HPG on birth outcomes beyond birthweight. The other central goal of the HPG – to reduce prematurity – cannot be investigated using this data, and instead requires Hospital Episode Statistics, which will be used in further research.

7. Conclusion and policy implications

The main aim of the Health in Pregnancy Grant (HPG) was to improve birth outcomes, specifically birthweight and prematurity. It was abolished in 2011 on the basis, in part, that it had failed to do so. This CASEpaper is the first to analyse the impact of the HPG in England and Wales, and the first to analyse the HPG more widely with an RD design.

First, my research suggests that the HPG did, contrary to the claims of its critics, have a remarkable positive impact on birthweight given its relatively small financial value and its late payment in pregnancy. The results are robust to parametric and non-parametric approaches, different specifications, and different bandwidths. They are also supported by three important robustness checks. It is beyond the scope of this CASEpaper to make substantive conclusions regarding the causal mechanisms driving the effect on birthweight, but it is possible, if we assume (as Leyland et al (2017) do) that the increase in birthweight must have been through nutrition, that the investment model is effective later in pregnancy than has been maintained by both critics and proponents of the HPG. Alternatively, if we maintain that the third trimester was indeed too late to improve birthweight through nutrition, this suggests that other factors such as reducing stress and reducing unhealthy behaviours such as smoking were effective instead, implying that the family stress and behavioural models can be effective later in pregnancy than the investment (nutrition) model. Of course, these mechanisms are not mutually exclusive, and it may be that some combination of the three causal mechanisms applied. Further research on these causal mechanisms would be helpful not only in an empirical sense, but also as a practical tool for refining future communications strategies on similar policies. While proponents of the HPG emphasised the role of nutrition in the media, in parliamentary discussions they conceded that the case for improved nutrition was a dubious one, making it vulnerable to critique and ultimately contributing to its abolition in 2011.

Second, the grant had significantly larger effects for younger mothers (25 and under). This could have been due to lower engagement with antenatal care and lower incomes among young mothers. It is unlikely to be explained by differential take-up, given that the grant was universal and there were relatively few barriers to entry. More plausible is the hypothesis that younger mothers saw greater increases in engagement with antenatal care, and perhaps that they were more sensitive to the labelling and 'learning effects' surrounding the HPG. If this is the case, then it implies that cash transfers do not have to have strict conditionality attached to them to influence behaviour.

Third, though the gains to birthweight were concentrated on smaller babies and younger mothers, they were not sufficient to significantly reduce low birthweight for the general population. This may have been because the grant was paid relatively late in pregnancy, or because the cash value was too low.

These results have striking policy implications for the targeting, conditionality and adequacy of cash transfers during pregnancy. First, as the only international example of a universal cash transfer to improve pregnancy and birth outcomes, the HPG represents an important case study in the debate about the targeting of cash transfers. My analysis suggests that universal programmes can have impacts that disproportionately benefit disadvantaged groups. The gains of the HPG were specific to smaller babies, and the gains were larger for younger mothers in particular. Younger mothers may be a difficult group to reach, as they have lower access to maternity and health services, information and support networks (DCSF, 2009). The fact that the HPG delivered specifically for young mothers therefore suggests that the universality of the grant may have been beneficial in offering non-stigmatised, accessible financial support for these women. Since the abolition of the HPG, the only remaining pregnancy-specific support in the UK – the Sure Start Maternity Grant and Healthy Start food vouchers – is means-tested, requiring a greater level of administrative input from the recipient. It is beyond the scope of this paper to compare the impact of the HPG with these means-tested programmes, but it is possible that the universality of the HPG may have been an important factor for reaching younger mothers, thereby explaining the scale of its impact for this group.

Second, my research suggests that cash transfers do not have to be conditional in order to elicit positive behavioural change. If the HPG was spent on plasma televisions, alcohol and cigarettes, as its critics had warned it would be, the treatment effect would be either zero or negative. The presence of a statistically significant increase in birthweight that was concentrated on smaller babies suggests, instead, that the HPG was spent on healthy nutrition, in ways that reduced antenatal stress or unhealthy behaviours, or some combination thereof. This supports the legitimacy of other labelled, universal and unconditional transfers to mothers that are founded on the principle of maternal choice, such as child benefit.

Third, the fact that the HPG boosted birthweights for small babies without leading to significant reductions in low birthweight suggests that despite these positive benefits, it is possible that the policy could have gone further by being paid earlier in pregnancy or by being larger in financial value. Policymakers interested in re-introducing the grant should consider whether it may be beneficial to bring the grant forwards to the first or second trimester and increase its cash value proportionately. The rationale for introducing the HPG was that birth was in a sense an arbitrary point to begin child benefit and state support for children: why not pay it during pregnancy and yield the health benefits? Arguably, the same rationale should have extended prior to the third trimester.

It is over a decade since the HPG was introduced, and nearly a decade since it was abolished. Since then, the social class gap in low birthweight has widened significantly (Stewart and Reader, 2021). COVID-19 and Marcus Rashford's food poverty campaign have brought health inequalities and food poverty to unprecedented public attention. The shock to living standards, the inevitable increase in child poverty, and the stress brought about to health systems and maternity services will exacerbate spiralling birthweight inequalities. Yet little policy action on early childhood, and none on pregnancy, has been forthcoming in the crisis. Much of the policy bandwidth has been taken up by debates on the important issues of school-age food poverty and in-kind provision such as free school meals, but children under five and cash transfers have attracted little attention.

The HPG may appear to be a small and insignificant policy of the past, but it is the only international example of a universal unconditional cash transfer during pregnancy to improve birthweight, and this research shows that it had a positive impact, especially on younger mothers. It could have gone further to achieve population-level improvements in low birthweight, and this paper has suggested some aspects of its design that could have increased its impact. It represents a unique policy experiment in the UK, and it is important to learn lessons from it. A re-vamped HPG with a higher cash value paid earlier in pregnancy could be a powerful intervention to reduce birthweight inequalities at a crucial moment.

References

Almond, D., Hoynes, H.W., and Schanzenbach., D.W. (2011). Inside The War On Poverty: The Impact Of Food Stamps On Birth Outcomes. Review of Economics and Statistics 93 (2): 387–403.

Averett, S., and Wang, Y. (2013). The effects of Earned Income Tax Credit payment expansion on maternal smoking. *Health Economics, 22*(11), 1344–1359.

BBC News (2007, September 8). *Eat well cash for mothers-to-be.* Retrieved from http://news.bbc.co.uk/1/hi/health/6984122.stm

BBC News (2017, June 5). *Pregnant women smoking rates 'concern' in England.* Retrieved from https://www.bbc.co.uk/news/health-40286702

Behrman, J. R., and Rosenzweig, M.R. (2004). Returns to Birthweight. *Review of Economics and Statistics, 86*(2), 586-601.

Benzeval, M., Bond, L., Campbell, M., Egan, M., Lorenc, T., Petticrew, M., and Popham, F. (2014). *How does money influence health? Technical Report.* Joseph Rowntree Foundation, York.

Bitler, M.P., and Currie, J. (2005). Does WIC Work? The Effects of WIC on Pregnancy and Birth Outcomes. Journal of Policy Analysis and Management 24(1),73-91.

Black, S., Devereux, P., and Salvanes, K. (2007). From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes. *The Quarterly Journal of Economics,122*(1), 409–439.

Blondel, B., Kogan, M., Alexander. G., Dattani, N., Kramer, M., Macfarlene, A., and Wen, S.W. (2002). The Impact of the Increasing Number of Multiple Births on the Rates of Preterm Birth and Low Birthweight: An International Study. *American Journal of Public Health, 92*(8), 1323–1330.

Bloom, T., Glass, N., Curry, M. A., Hernandez, R., and Houck, G. (2013). Maternal Stress Exposures, Reactions, and Priorities for Stress Reduction among Low-income, Urban Women. *Journal of Midwifery and Women's Health, 58*, 167–174. doi:10.1111/j.1542-2011.2012.00197.x

Boyd-Swan, C., Herbst, C. M., Ifcher, J., and Zarghamee, H. (2016). The Earned Income Tax Credit, Mental Health, and Happiness. *Journal of Economic Behaviour and Organization, 126,* 18–38.

De Brauw, A., and Peterman, A. (2020). Can Conditional Cash Transfers Improve Maternal Health Care? Evidence from El Salvador's Comunidades Solidarias Rurales Program. Health Economics, 29(6), 700-15.

Brooks-Gunn, J., and Duncan, G. (1997). The Effects of Poverty on Children. *The Future of Children, 7*(2), 55–71.

Card, D., and Shore-Sheppard, L. (2004). Using discontinuous eligibility rules to identify the effects of the federal Medicaid expansions on low income children. *Review of Economics and Statistics, 86*, 752–766.

Cattaneo, M.D., Idrobo, N., and Titiunik, R. (2018a) A Practical Introduction to Regression Discontinuity Designs: Volume I. Cambridge Elements: Quantitative and Computational Methods for Social Science, Cambridge University Press.

Cattaneo, M.D., Idrobo, N., and Titiunik, R. (2018b) A Practical Introduction to Regression Discontinuity Designs: Volume II. Cambridge Elements: Quantitative and Computational Methods for Social Science, Cambridge University Press.

Chersich, M.F., Luchters, S., Blaauw, D., Scorgie, F., Kern, E., van den Heever, A., Rees, H., Peach, E., Kharadi, S., and Fonn, S. (2016). Safeguarding maternal and child health in South Africa by starting the Child Support Grant before birth: Design lessons from pregnancy support programmes in 27 countries. *SAMJ Research, 106,* 1192–1210.

Chorniy, A., Currie, J., and Sonchak, L. (2020). Does prenatal WIC participation improve child outcomes? American Journal of Health Economics, 6(2). DOI: 10.1086/707832

Chung, W., Ha, H., and Kim, B. (2016). Money Transfer and Birth Weight: Evidence from the Alaska Permanent Fund Dividend. Economic Inquiry, 54(1), 576-590.

Conti, G., Hanson, M., Inskip, H., Crozier, S., Cooper, C., and Godfrey, K. (2018). *Beyond birth weight: the origins of human capital* (IFS Working Paper W18/30). London: Institute for Fiscal Studies.

Cooper, K., and Stewart, K. (2017). *Does Money Affect Children's Outcomes? An Update* (CASEpaper 203). Centre for Analysis of Social Exclusion. ISSN 1460-5023. Retrieved from http://sticerd.lse.ac.uk/dps/case/cp/casepaper203.pdf

Cowan, B., and Tefft, N. (2012). Education, Maternal Smoking, and the Earned Income Tax Credit. *B E Journal of Economic Analysis and Policy, 12*(1), 1–37.

Cramer, J. C. (1995). Racial and Ethnic Differences in Birthweight: The Role of Income and Financial Assistance. *Demography, 32*, 231–47.

Department for Children, Schools and Families (DCSF) (2009). Getting maternity services right for pregnant teenagers and young fathers. https://dera.ioe.ac.uk/10606/7/DCSF-00673-2009_Redacted.pdf

Dibben, C., Sigala, M., and Macfarlane, A. (2006). Area deprivation, individual factors and low birth weight in England: Is there evidence of an "area effect"? *Journal of Epidemiology and Community Health, 60*(12), 1053–1059.

Directgov (2010, May 19). *Information for mums-to-be: Health in Pregnancy Grant.* Retrieved from https://webarchive.nationalarchives.gov.uk/20100519165654/http://campaigns2.direct.gov.uk/money4mum2be/en/

Dunkel Schetter, C., and Lobel, M. (2012). Pregnancy and birth outcomes: A multilevel analysis of prenatal maternal stress and birth weight. In Baum, A., Revenson, T. A., and Singer, J. (Eds.), *Handbook of health psychology* (pp. 431–463). New York: Routledge.

Edmonds, T., and Kennedy, S. (2010). Savings Accounts and Health in Pregnancy Grant Bill (Research Paper 10/66). London, England: House of Commons Library.

Eggers, A., Freier, R., Grembi, V., and Nannicini, T. (2018). Regression Discontinuity Designs Based on Population Thresholds: Pitfalls and Solutions. *American Journal of Political Science, 62*(1), 210–229.

Fan, J., and Gijbels, I. (1996). *Local Polynomial Modelling and Its Applications.* London, England: Chapman and Hall.

Food Foundation. (2019). *The Broken Plate: Ten vital signs revealing the health of our food system, its impact on our lives and the remedies we must pursue.* Retrieved from https://foodfoundation.org.uk/wp-content/uploads/2019/02/The-Broken-Plate.pdf

Gelman, A., and Imbens, G. (2018). Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs. *Journal of Business and Economic Statistics,* 1–10.

Ghosh, R.E., Berild, J.D., Sterrantino, A.F., Toledano, M.B., and Hansell, A.L. (2017). Birth weight trends in England and Wales (1986–2012): babies are getting heavier. *Arch Dis Child Fetal Neonatal, 103,* F264–F270. doi:10.1136/archdischild-2016-311790

Glover, V., and O'Connor, T. (2002). Effects of antenatal stress and anxiety: Implications for development and psychiatry. *British Journal of Psychiatry, 180,* 389–391.

Goisis, A., Özcan, B., and Myrskylä, M. (2016). Decline in the negative association between low birth weight and cognitive ability. *Proceedings of the National Academy of Sciences, 114.* doi:10.1073/pnas.1605544114

Goisis, A., Schneider, D.C., and Myrskylä, M. (2018). Secular changes in the association between advanced maternal age and the risk of low birth weight: A cross-cohort comparison in the UK. *Population Studies, 72*(3), 381–397. doi:10.1080/00324728.2018.1442584

Gregg, P., Waldfogel, J., and Washbrook, E. (2006). Family Expenditures Post- Welfare Reform in the UK: Are Low-Income Families with Children Starting to Catch Up? *Labour Economics, 13*(6), 721–746.

Hahn, J., van der Klaauw, W., and Todd, P. (2001). Identification and estimation of treatment effects with a regression–discontinuity design. *Econometrica,69*(1), 201–209.

HC Deb, 22 November 2010, vol. 519, cc 102–136. Retrieved from https://hansard.parliament.uk/Commons/2010-11-22/debates/10112215000001/SavingsAccountsAndHealthInPregnancyGrantBill?highlight=health%20pregnancy#contribution-10112251000633

HMRC (2010). *HMRC Departmental Accounts 2009-10* (HC 299). London, England: The Stationery Office. Retrieved from https://webarchive.nationalarchives.gov.uk/20130127142016/http://www.hmrc.gov.uk/about/hmrc-accs-0910.pdf

Hollowell, J., Oakley, L., Kurinczuk, J.J., Brocklehurst, P., and Gray, R. (2011). The Effectiveness of Antenatal Care Programmes to Reduce Infant Mortality and Preterm Birth in Socially Disadvantaged and Vulnerable Women in High-income Countries: A Systematic Review. BMC Pregnancy and Childbirth 11.1: 13.

Howard, D. (2011, January 10). New parents 'hardest hit' by cuts and benefits changes. *BBC Newsbeat.* Retrieved from http://www.bbc.co.uk/newsbeat/article/12150582/new-parents-hardest-hit-by-cuts-and-benefits-changes

Hoynes, H., Miller, D., and Simon, D. (2015). Income, the earned income tax credit, and infant health. *American Economic Journal: Economic Policy, 7*(1), 172–211.

Hyytinen, A., Meriläinen, J., Saarimaa, T., Toivanen, O., and Tukiainen, J. (2018). When does regression discontinuity design work? Evidence from random election outcomes. *Quantitative Economics, 9*(2), 1019-1051.

Imbens, G., and Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics,142*(2), 615–635.

Jacob, R., and Zhu, Pei (2012). A Practical Guide to Regression Discontinuity. Retrieved from https://www.mdrc.org/sites/default/files/regression_discontinuity_full.pdf

Kapaya, H., Mercer, E., Boffey, F., Jones, G., Mitchell, C., and Anumba, D. (2015). Deprivation and poor psychosocial support are key determinants of late antenatal presentation and poor fetal outcomes – a combined retrospective and prospective study. *BMC Pregnancy and Childbirth, 15.* doi:10.1186/s12884-015-0753-3

Kaushal, N., Gao, Q., and Waldfogel, J. (2007). Welfare Reform and Family Expenditures: How Are Single Mothers Adapting to the New Welfare and Work Regime? *Social Service Review, 81*, 369–96.

Kelly, Y., Panico, Y., Bartley, M., Marmot, M., Nazroo, J., and Sacker, A. (2008). Why does birthweight vary among ethnic groups in the UK? Findings from the Millennium Cohort Study. *Journal of Public Health, 31*(1), 131–137. doi:10.1093/pubmed/fdn057

Koivu, A., Phan, Y.T.H, Abuhamed, J., Perry, B.L., Atkins, S., Perkiö, M., Koivusalo, M. (2020). The baby box: Enhancing the wellbeing of babies and mothers around the world. Kela.

Lee, D.S., and Card, D. (2008). Regression discontinuity inference with specification error. *Journal of Econometrics, 142*, 655–674.

Lee, D.S., and Lemieux, T. (2010). Regression Discontinuity Designs in Economics. *Journal of Economic Literature, 48,* 281–355.

Leyland, A.H., Ouédraogo, S., Nam, J., Bond, L., Briggs, A.H., Gray, R., Wood, R., and Dundas, R. (2017). Evaluation of Health in Pregnancy Grants in Scotland: a natural experiment using routine data. *Public Health Research, 5.* doi:10.3310/phr05060

Loopstra, R., and Tarasuk, V. (2013). Severity of household food insecurity is sensitive to change in household income and employment status among low-income families. *The Journal of Nutrition, 143*(8), 1316–1323.

Marmot, M. (2010). *Fair society, healthy lives: The Marmot Review: Strategic review of health inequalities in England post-2010*. Retrieved from http://www.instituteofhealthequity.org/resources-reports/fair-society-healthy-lives-the-marmot-review/fair-society-healthy-lives-full-report-pdf.pdf

Matsushima, M. (2018). Life course consequences of low birth weight: Evidence from Japan. *Journal of The Japanese and International Economies.* https://doi.org/10.1016/j.jjie.2018.07.001

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics, 142*(2), 698–714.

Milligan, K., and Stabile, M. (2011) Do Child Tax Benefits Affect the Well-Being of Children? Evidence from Canadian Child Benefit Expansions. *American Economic Journal: Economic Policy, 3,* 175–205.

Mocan, N., Raschke, C., and Unel, B. (2015). The Impact of Mothers' Earnings on Health Inputs and Infant Health. Economics and Human Biology, 19, 204-223. doi: http://www.sciencedirect.com/science/journal/1570677X

Mumsnet. (2009, April 17). *£190 grant for mums to be (news to me anyway).* Retrieved from https://www.mumsnet.com/Talk/childbirth/740256-pound-190-grant-for-mums-to-be-news-to-me

Nascimento, L.F.C., Machin, A.B., and dos Santos, D.A.A. (2017). Are there differences in birth weight according to sex and associations with maternal exposure to air pollutants? A cohort study. *Sao Paulo Med J., 135*(4), 347–354.

Newham, J.J., Wittkowski, A., Hurley, J., Aplin, J.D., and Westwood. M.

(2014). Effects of antenatal yoga on maternal anxiety and depression: A randomised controlled trial *Depression and Anxiety*, 31/8, 631–640. doi:10.1002/da.22268

Office for National Statistics. (2017). *Births Registrations in England and Wales, 1982-2016: Secure Access.* [data collection]. *3rd Edition.* UK Data Service. SN: 8043, http://doi.org/10.5255/UKDA-SN-8043-3

Public Bills Committee. (2008, January 10). Savings Accounts and Health in Pregnancy Grant Bill, 3rd sitting, cc 85–106. Retrieved from https://publications.parliament.uk/pa/cm200708/cmpublic/health/080110/am/80110s01.htm

Public Bills Committee. (2008, January 24). Savings Accounts and Health in Pregnancy Grant Bill, 11th sitting, cc 455–480. Retrieved from https://publications.parliament.uk/pa/cm200708/cmpublic/health/080124/am/80124s01.htm

Public Bills Committee. (2008, January 24). Savings Accounts and Health in Pregnancy Grant Bill, 12th sitting, cc 483–488. Retrieved from https://publications.parliament.uk/pa/cm200708/cmpublic/health/080124/pm/80124s01.htm

Public Bills Committee. (2010, November 2). Savings Accounts and Health in Pregnancy Grant Bill, 3rd sitting, cc 79–100. Retrieved from https://publications.parliament.uk/pa/cm201011/cmpublic/savingsacc/101104/am/101104s01.htm

Public Bills Committee. (2010, November 2). Savings Accounts and Health in Pregnancy Grant Bill, 4th sitting, cc 101–155. Retrieved from https://publications.parliament.uk/pa/cm201011/cmpublic/savingsacc/101104/pm/101104s01.htm

Public Health England (2019). Health of women before and during pregnancy: health behaviours, risk factors and inequalities. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/844210/Health_of_women_before_and_during_pregnancy_2019.pdf

Revill, J. (2007, September 9). Pregnant women to get cash for good diet. *The Observer.* Retrieved from https://www.theguardian.com/uk/2007/sep/09/health.healthandwellbeing

Ridge, T. (2013). 'We are All in This Together'? The Hidden Costs of

Poverty, Recession and Austerity Policies on Britain's Poorest Children. *Children and Society, 27,* 406–417. doi:10.1111/chso.12055

Roll, J. (2007). *Health and Social Care Bill: Bill 9 of 2007-08* (Research Paper 07/81). London, England: House of Commons Library.

Sharp-Jeffs, N. (2015). *Money matters: Research into the extent and nature of financial abuse within intimate relationships in the UK*. The Co-operative Bank and Refuge. Retrieved from http://cwasu.org/resource/research-into-the-extent-and-nature- of-financial-abuse-within-intimate-relationships/

Silverwood, R.J., Pierce, M., Hardy, R., Sattar, N., Whincup, P., Ferro, C., Savage, C., Kuh, D., and Nitsch, D. (2013). Low birth weight, later renal function, and the roles of adulthood blood pressure, diabetes, and obesity in a British birth cohort. *Kidney International, 84,* 1262–1270. doi:10.1038/ki.2013.223

Smith. T., Noble, M., Noble, S., Wright, G., McLennan, D., and Plunkett, E. (2015). *The English Indices of Deprivation 2015: Technical report.* Department for Communities and Local Government. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/464485/English_Indices_of_Deprivation_2015_-_Technical-Report.pdf

Sosa-Rubi, S.G., Walker, D., Servan, E., and Bautista-Arredondo, S. (2011). Learning Effect of a Conditional Cash Transfer Programme on Poor Rural Women's Selection of Delivery Care in Mexico. Health Policy and Planning, 26(6), 496-507.

Stewart, K., and Obolenskaya, P. (2016). 'Young children,' in R Lupton et al (eds), Social Policy in a Cold Climate: Policies and Their Consequences Since the Crisis. Bristol: The Policy Press.

Stewart, K., and Reader, M. (2021). The Conservative Governments' Record on Early Childhood from May 2015 to pre-COVID 2020: Policies, Spending and Outcomes. Social Policies and Distributional Outcomes Research Paper 8, Centre for Analysis of Social Exclusion.

Strully, K.W., Rehkopf, D.H., and Xuan, Z. (2010). Effects of Prenatal Poverty on Infant Health: State Earned Income Tax Credits and Birth Weight. *American Sociological Review, 75,* 534–562. doi: 10.1177/0003122410374086

Thistlethwaite, D. L., and Campbell., D.T. (1960). Regression Discontinuity Analysis: An Alternative to the Ex Post Facto Experiment. *Journal of Educational Psychology, 51*(6), 309–317.

UNICEF (2020). Universal child benefits: policy issues and options. UNICEF. https://www.unicef.org/media/72916/file/UCB-ODI-UNICEF-Report-2020.pdf

Victora, C.G., Adair, L., Fall, C., Hallal, P.C., Martorell, R., Richter, L., Singh Sachdev, H. (2008). Maternal and child undernutrition: consequences for adult health and human capital. *The Lancet, 371,* 340–357.

Williamson, C. (2006). Nutrition in pregnancy. *Nutrition Bulletin, 31,* 28–59.

WiredGov. (2009, January 20). *Pregnant mums urged – apply now for £190 cash boost.* Retrieved from https://www.wired-gov.net/wg/wg-news-1.nsf/0/D9E5D8C47BAF23DC80257544003B7260?OpenDocument

Wright, M. (2009, February 8). *New grant aimed at improving health of expectant mums*. The Observer. Retrieved from https://www.theguardian.com/money/2009/feb/08/pregnancy-grant

Appendix

Table A1 Parametric RD treatment effects of the Health in Pregnancy Grant on birthweight

	Bandwidth		24 weeks	20 weeks	16 weeks	12 weeks	8 weeks
Birthweight							
	Linear		15.84***	13.02***	10.39***	10.65***	10.85***
			(3.429)	(3.913)	(2.778)	(3.198)	(3.540)
		AIC	15.63	15.63	15.62	15.63	15.63
		BIC	1.444e+06	1.300e+06	1.133e+06	943696	716659
	Linear interaction		16.33***	13.38***	10.17***	10.36***	10.83***
			(3.350)	(3.953)	(2.593)	(2.802)	(3.485)
		AIC	15.63	15.63	15.62	15.63	15.63
		BIC	1.444e+06	1.300e+06	1.134e+06	943708	716671
	Quadratic		15.88***	13.03***	10.38***	10.65***	10.83**
			(3.099)	(3.636)	(2.664)	(2.908)	(3.691)
		AIC	15.63	15.63	15.62	15.63	15.63
		BIC	1.444e+06	1.300e+06	1.134e+06	943707	716671
	n		639,171	532,334	425,620	319,111	213,214
Low birthweight							
	Probit		-0.002	-0.002	-0.002	-0.003	-0.005**
			(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
		AIC	0.517	0.518	0.521	0.523	0.523
		BIC	-8.214e+06	-6.743e+06	-5.295e+06	-3.877e+06	-2.505e+06
	Logit		-0.002	-0.002	-0.002	-0.003	-0.005**
			(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
		AIC	0.517	0.518	0.521	0.523	0.523
		BIC	-8.214e+06	-6.743e+06	-5.295e+06	-3.877e+06	-2.505e+06
	n		639,171	532,334	425,620	319,111	213,214

Note: Standard errors, clustered by week of birth, in parentheses.

Table A2 Non-parametric RD effects of the HPG on birthweight

		(1)	(2)	(3)	(4)
	Bandwidth	MSE-optimal	16 weeks	12 weeks	8 weeks
Birthweight	Confidence intervals:				
	Conventional	12.21***	9.77***	9.91***	11.43***
		(0.798)	(2.386)	(2.570)	(2.370)
	Bias-corrected	10.72***	10.65***	11.67***	6.37***
		(0.798)	(2.386)	(2.570)	(2.370)
	Robust bias-corrected	10.72***	10.65***	11.67***	6.37*
		(3.291)	(2.911)	(3.019)	(3.566)
	MSE-optimal h	2.69			
	Effective n	67,553	412,078	306,131	200,012
Low birthweight	Confidence intervals:				
	Conventional	-0.004***	-0.003	-0.003*	-0.003**
		(0.001)	(0.002)	(0.002)	(0.002)
	Bias-corrected	-0.004***	-0.004**	-0.003*	-0.000
		(0.001)	(0.002)	(0.002)	(0.002)
	Robust bias-corrected	-0.004	-0.004**	-0.003	-0.000
		(0.002)	(0.002)	(0.002)	(0.003)
	MSE-optimal h	4.061			
	Effective n	120,477	412,078	306,131	200,012
	Triangular kernel	X	X	X	X
	MSE-optimal h	X			
	Linear	X	X	X	X

Note: Standard errors, clustered by week of birth, in parentheses.

Table A3 Heterogeneity in birthweight treatment effect by birthweight decile

Birthweight decile	(1)	(2)	(3)	(4)	(5)
D1 (mean 2279g)	15.58*	20.27***	16.50**	28.84**	4.86
	-8.307	(6.323)	(8.049)	(11.316)	(9.265)
MSE-optimal	13.37				
Effective n	35,639	40,857	30,399	19,847	42,200
D2 (mean 2851g)	0.55	0.80	1.94	-0.08	1.21
	-1.426	(1.456)	(1.637)	(1.196)	(1.345)
MSE-optimal	12.94				
Effective n	32,362	40,243	29,787	19,321	41,658
D3 (mean 3053g)	0.98	1.42	1.21	2.49*	-0.48
	-1.064	(0.957)	(1.017)	(1.288)	(0.808)
MSE-optimal	11.71				
Effective n	30,528	41,048	30,528	20,086	42,436
D4 (mean 3201g)	0.95	1.19	1.55	2.29**	-0.16
	-0.94	(0.958)	(1.027)	(1.115)	(0.729)
MSE-optimal	10.81				
Effective n	26,371	39,270	28,902	18,842	40,592
D5 (mean 3329g)	-3.00***	-2.26***	-2.75***	-3.06***	-0.55
	-0.381	(0.432)	(0.394)	(0.361)	(0.668)
MSE-optimal	4.861				
Effective n	11,591	39,769	29,329	19,107	41,092
D6 (mean 3458g)	0.17	-0.02	-0.47	-1.21	0.06
	-0.836	(0.879)	(0.859)	(0.877)	(0.741)
MSE-optimal	13.89				
Effective n	37,353	42,747	31,798	20,867	44,156
D7 (mean 3588g)	-1.73**	-1.71**	-1.59**	-1.28	-0.20
	-0.767	(0.789)	(0.784)	(0.828)	(0.558)
MSE-optimal	7.805				
Effective n	18,094	37,269	27,848	18,094	38,488
D8 (mean 3730g)	-0.05	-0.06	-0.65	-0.82	0.06
	-0.669	(0.629)	(0.632)	(0.602)	(0.581)
MSE-optimal	9.417				
Effective n	24,385	39,346	29,365	19,230	40,678
D9 (mean 3920g)	1.47	1.29	0.83	0.93	2.58***
	-0.999	(0.953)	(0.877)	(0.783)	(0.933)
MSE-optimal	10.07				
Effective n	27,250	40,118	29,879	19,589	41,319
D10 (mean 4304g)	2.52	1.67	-0.92	0.70	1.18
	-3.593	(3.421)	(3.821)	(3.544)	(3.859)
MSE-optimal	14.34				
Effective n	35,598	38,026	28,312	18,590	39,190
Non-parametric	X	X	X	X	
MSE-optimal h	X				
Robust bias-corrected	X	X	X	X	
Linear	X	X	X	X	X

Notes: Outcome is birthweight. Preferred model is Model 1. Model 1 is a non-parametric RD specification using the MSE-optimal bandwidth with robust bias-corrected standard errors. Models 2-4 are non-parametric RD specifications using 16-, 12- and 8-week bandwidths respectively with robust bias-corrected standard errors. Model 5 is a parametric linear specification for a 16-week bandwidth. Errors are clustered by week of birth. * p<0.1. ** p<0.05. *** p<0.01.

Table A4 Heterogeneity in birthweight treatment effect by birthweight quartile

Birthweight quartile	(1)	(2)	(3)	(4)	(5)
Q1 (mean 2654g)	11.31**	13.44**	8.95	4.89	3.00
	-5.469	(5.393)	(5.716)	(6.554)	(6.226)
MSE-optimal	8.875				
Effective n	55,556	101,429	75,297	49,092	104,889
Q2 (mean 3230g)	-1.72	-2.21	-2.29	0.37	0.37
	-2.141	(2.519)	(2.901)	(2.148)	(1.448)
MSE-optimal	11.89				
Effective n	73,648	99,758	73,648	48,111	103,089
Q3 (mean 3555g)	-0.84	-0.96	-0.71	-2.83***	0.30
	-1.604	(1.586)	(1.598)	(1.081)	(1.151)
MSE-optimal	8.33				
Effective n	55,603	100,666	75,130	49,117	104,025
Q4 (mean 4044g)	0.03	2.65	0.69	-0.46	4.40
	-3.793	(3.787)	(4.279)	(4.401)	(3.228)
MSE-optimal	6.121				
Effective n	40,863	96,840	72,072	47,253	99,806
Non-parametric	X	X	X	X	
MSE-optimal h	X				
Robust bias-corrected	X	X	X	X	
Linear	X	X	X	X	X

Notes: Outcome is birthweight. Preferred model is Model 1. Model 1 is a non-parametric RD specification using the MSE-optimal bandwidth with robust bias-corrected standard errors. Models 2-4 are non-parametric RD specifications using 16-, 12- and 8-week bandwidths respectively with robust bias-corrected standard errors. Model 5 is a parametric linear specification for a 16-week bandwidth. Errors are clustered by week of birth. * p<0.1. ** p<0.05. *** p<0.01.

Table A5 Heterogeneity in impact on birthweight by maternal age decile

Maternal age decile	(1)	(2)	(3)	(4)	(5)
D1 (21 and under)	32.52*	25.74	35.32**	40.86***	18.62
	(17.508)	(18.642)	(16.774)	(12.256)	(13.718)
MSE-optimal	8.034				
Effective n	23,133	42,251	31,263	20,374	43,603
D2 (21-24 yrs)	26.90*	22.61*	26.81*	15.36	-1.14
	(14.338)	(13.293)	(14.537)	(17.795)	(12.089)
MSE-optimal	8.171				
Effective n	22,528	41,136	30,485	19,875	42,499
D3 (24-26 yrs)	12.11	8.09	12.22	7.19	15.38
	(8.944)	(9.219)	(11.218)	(15.554)	(11.839)
MSE-optimal	8.476				
Effective n	22,634	41,154	30,566	19,967	42,493
D4 (26-28 yrs)	5.56**	12.49**	7.53	5.60	14.59***
	(2.827)	(5.190)	(4.613)	(4.515)	(5.218)
MSE-optimal	4.551				
Effective n	12,366	41,926	31,224	20,442	43,297
D5 (28-30 yrs)	-26.69**	-26.06**	-24.20*	-15.97	-13.51
	(12.197)	(12.462)	(12.466)	(17.055)	(11.536)
MSE-optimal	9.887				
Effective n	24,964	40,895	30,293	19,656	42,236
D6 (30-31 yrs)	9.85	8.08	-11.37	-42.35**	21.69
	(21.422)	(21.067)	(18.681)	(16.459)	(14.587)
MSE-optimal	9.881				
Effective n	25,016	40,648	30,201	19,831	41,953
D7 (31-33 yrs)	25.18*	-2.30	18.50	40.81***	18.63*
	(13.510)	(14.373)	(12.037)	(9.412)	(9.319)
MSE-optimal	5.566				
Effective n	14,942	41,788	31,063	20,307	43,191
D8 (33-35 yrs)	27.68**	28.43**	20.36*	0.69	18.42*
	(11.649)	(11.148)	(11.043)	(12.310)	(9.696)
MSE-optimal	9.363				
Effective n	25,438	41,186	30,650	20,149	42,526
D9 (35-37 yrs)	24.95*	16.88	33.88***	21.84	-6.35
	(13.183)	(14.374)	(12.161)	(16.592)	(11.089)
MSE-optimal	5.877				
Effective n	14,231	39,795	29,666	19,358	41,134
D10 (37 and over)	6.63	10.41	-1.31	-10.77	10.77
	(15.621)	(14.365)	(14.674)	(15.910)	(11.665)
MSE-optimal	8.788				
Effective n	22,678	41,299	30,720	20,053	42,688
Non-parametric	X	X	X	X	
MSE-optimal h	X				
Robust bias-corrected	X	X	X	X	
Linear	X	X	X	X	X

Notes: Outcome is birthweight. Preferred model is Model 1. Model 1 is a non-parametric RD specification using the MSE-optimal bandwidth with robust bias-corrected standard errors. Models 2-4 are non-parametric RD specifications using 16-, 12- and 8-week bandwidths respectively with robust bias-corrected standard errors. Model 5 is a parametric linear specification for a 16-week bandwidth. Errors are clustered by week of birth. * p<0.1. ** p<0.05. *** p<0.01.

Table A6 Heterogeneity in impact on birthweight by maternal age quartile

Maternal age quartile	(1)	(2)	(3)	(4)	(5)
Q1 (25 and under)	28.89**	24.02*	30.10**	26.48***	10.53
	-13.161	(13.222)	(11.961)	(8.635)	(8.888)
MSE-optimal	7.927				
Effective n	49,982	103,232	76,566	49,982	106,611
Q2 (25-30 yrs)	-5.83	-6.42	-6.41**	-3.96	3.68
	-3.963	(4.023)	(3.226)	(4.037)	(4.816)
MSE-optimal	6.066				
Effective n	43,703	104,130	77,265	50,332	107,517
Q3 (30-34 yrs)	3.29	3.21	1.95	-4.48	16.33**
	-7.661	(7.494)	(7.883)	(9.699)	(7.113)
MSE-optimal	8.329				
Effective n	56,063	101,748	75,652	49,616	105,084
Q4 (34 and over)	14.4	21.53**	22.08**	8.55	9.47
	-9.762	(9.899)	(9.178)	(10.178)	(7.677)
MSE-optimal	4.224				
Effective n	30,202	102,968	76,648	50,082	106,408
Non-parametric	X	X	X	X	
MSE-optimal h	X				
Robust bias-corrected	X	X	X	X	
Linear	X	X	X	X	X

Notes: Outcome is birthweight. Models 1 and 3 are non-parametric RD specifications using the MSE-optimal bandwidth for each year with robust bias-corrected standard errors. Models 2-4 are non-parametric RD specifications using 16-, 12- and 8-week bandwidths respectively with robust bias-corrected standard errors. Model 5 is a parametric linear specification for a 16-week bandwidth. Errors are clustered by week of birth. * p<0.1. ** p<0.05. *** p<0.01.

Table A7 Heterogeneity in impact on low birthweight by maternal age quartile

Maternal age quartile	(1)	(2)	(3)	(4)	(5)
Q1 (25 and under)	-0.009**	-0.011***	-0.011***	-0.008**	-0.019
	(0.003)	(0.004)	(0.004)	(0.003)	(0.018)
MSE-optimal	11.59				
Effective n	76,566	103,232	76,566	49,982	106,611

Q2 (25-30 yrs)	-0.001	-0.002	-0.001	-0.001	0.006
	(0.002)	(0.002)	(0.002)	(0.003)	(0.018)
MSE-optimal	6.205				
Effective n	43,703	104,130	77,265	50,332	107,517
Q3 (30-34 yrs)	0.008*	0.007*	0.010***	0.013***	-0.009
	(0.005)	(0.004)	(0.003)	(0.003)	(0.028)
MSE-optimal	7.038				
Effective n	49,616	101,748	75,652	49,616	105,084
Q4 (34 and over)	-0.010	-0.010	-0.010	-0.005	-0.027
	(0.008)	(0.008)	(0.009)	(0.011)	(0.035)
MSE-optimal	9.917				
Effective n	63,544	102,968	76,648	50,082	106,408
Non-parametric	X	X	X	X	
MSE-optimal h	X				
Robust bias-corrected	X	X	X	X	
Linear	X	X	X	X	X

Notes: Outcome is a low birthweight dummy. Preferred model is Model 1. Model 1 is a non-parametric RD specification using the MSE-optimal bandwidth with robust bias-corrected standard errors. Models 2-4 are non-parametric RD specifications using 16-, 12- and 8-week bandwidths respectively with robust bias-corrected standard errors. Model 5 is a parametric linear specification for a 16-week bandwidth. Errors are clustered by week of birth. * p<0.1. ** p<0.05. *** p<0.01.

Table A8 Heterogeneity in impact on birthweight by English index of income deprivation

English index of income deprivation quartile	(1)	(2)	(3)	(4)	(5)
Q1 (mean deprivation rate 5%)	9.27	12.57**	11.18*	0.24	19.94***
	(5.995)	(5.987)	(5.794)	(6.731)	(5.266)
MSE-optimal	6.343				
Effective n	49,193	116,103	86,638	56,615	119,874
Q2 (mean deprivation rate 11%)	-1.64	4.69	6.07	-10.76	-0.35
	(10.986)	(9.104)	(9.472)	(8.518)	(6.654)
MSE-optimal	5.634				
Effective n	29,379	81,992	60,938	39,944	84,666
Q3 (mean deprivation rate 20%)	1.71	2.19	2.31	8.44	7.56
	(7.712)	(6.773)	(7.951)	(7.688)	(4.649)
MSE-optimal	8.046				
Effective n	53,122	97,055	71,993	46,821	100,241
Q4 (mean deprivation rate 34%)	20.03*	18.44*	20.41*	13.30	7.57
	(10.719)	(10.582)	(10.731)	(9.592)	(7.679)
MSE-optimal	8.471				
Effective n	52,863	96,451	71,315	46,730	99,665
Non-parametric	X	X	X	X	
MSE-optimal h	X				
Robust bias-corrected	X	X	X	X	
Linear	X	X	X	X	X

Notes: Outcome is birthweight. Models 1 and 3 are non-parametric RD specifications using the MSE-optimal bandwidth for each year with robust bias-corrected standard errors. Models 2-4 are non-parametric RD specifications using 16-, 12- and 8-week bandwidths respectively with robust bias-corrected standard errors. Model 5 is a parametric linear specification for a 16-week bandwidth. Errors are clustered by week of birth. * p<0.1. ** p<0.05. *** p<0.01.

Table A9 Heterogeneity in impact on low birthweight by English income deprivation quartile

English index of income deprivation quartile		(1)	(2)	(3)	(4)	(5)
Q1 (mean deprivation rate 5%)		-0.003	-0.003	-0.002	-0.000	-0.039*
		(0.003)	(0.003)	(0.004)	(0.006)	(0.021)
	MSE-optimal	9.252				
	Effective n	71,710	116,103	86,638	56,615	119,874
Q2 (mean deprivation rate 11%)		-0.000	-0.000	0.000	0.007	0.007
		(0.004)	(0.003)	(0.004)	(0.008)	(0.033)
	MSE-optimal	11.4				
	Effective n	60,938	81,992	60,938	39,944	84,666
Q3 (mean deprivation rate 20%)		-0.002	-0.002	-0.003	-0.006**	-0.006
		(0.003)	(0.003)	(0.003)	(0.003)	(0.014)
	MSE-optimal	8.903				
	Effective n	53,122	97,055	71,993	46,821	100,241
Q4 (mean deprivation rate 34%)		-0.006	-0.007	-0.005	0.005	0.008
		(0.006)	(0.006)	(0.006)	(0.005)	(0.023)
	MSE-optimal	7.345				
	Effective n	46,730	96,451	71,315	46,730	99,665
Non-parametric		X	X	X	X	
MSE-optimal h		X				
Robust bias-corrected		X	X	X	X	
Linear		X	X	X	X	X

Notes: Outcome is a low birthweight dummy. Preferred model is Model 1. Model 1 is a non-parametric RD specification using the MSE-optimal bandwidth with robust bias-corrected standard errors. Models 2-4 are non-parametric RD specifications using 16-, 12- and 8-week bandwidths respectively with robust bias-corrected standard errors. Model 5 is a parametric linear specification for a 16-week bandwidth. Errors are clustered by week of birth. * p<0.1. ** p<0.05. *** p<0.01.

Table A10 Full parametric results for birthweight from placebo cut-off tests

Placebo cut-off date			24 weeks	20 weeks	16 weeks	12 weeks	8 weeks
2007	Linear		8.07**	3.63	3.73	2.22	-7.32
			-3.966	-4.244	-4.979	-5.035	-5.663
		AIC	15.63	15.63	15.63	15.63	15.63
		BIC	1.425e+06	1.277e+06	1.111e+06	920332	695503
	Linear interaction		8.05**	3.34	3.05	2.00	-7.36
			-3.991	-4.093	-4.465	-4.784	-5.549

		AIC	15.63	15.63	15.63	15.63	15.63
		BIC	1.425e+06	1.277e+06	1.111e+06	920344	695515
	Quadratic		8.08**	3.59	3.63	2.25	-7.34
			-4.01	-4.125	-4.417	-4.861	-5.737
		AIC	15.63	15.63	15.63	15.63	15.63
		BIC	1.425e+06	1.277e+06	1.111e+06	920344	695515
2008	Linear		4.82	-0.94	-3.6	-6.26	3.88
			-3.659	-3.75	-4.196	-3.783	-3.402
		AIC	15.63	15.63	15.63	15.64	15.64
		BIC	1.454e+06	1.310e+06	1.146e+06	952787	723358
	Linear interaction		4.82	-1.26	-4.16	-6.74	3.94
			-3.702	-3.764	-4.108	-4.072	-3.278
		AIC	15.63	15.63	15.63	15.64	15.64
		BIC	1.454e+06	1.310e+06	1.146e+06	952797	723371
	Quadratic		4.84	-0.97	-3.64	-6.24	3.89
			-3.698	-3.675	-3.748	-3.796	-3.503
		AIC	15.63	15.63	15.63	15.64	15.64
		BIC	1.454e+06	1.310e+06	1.146e+06	952795	723371
2010	Linear		5.33	1.43	-1.75	-4.66	-5.84
			-4.23	-4.741	-5.363	-6.13	-8.149
		AIC	15.62	15.62	15.62	15.62	15.62
		BIC	1.457e+06	1.310e+06	1.144e+06	949938	718530
	Linear interaction		5.38	1.24	-2.24	-5.02	-6.00
			-4.305	-4.681	-5.097	-6.195	-8.312
		AIC	15.62	15.62	15.62	15.62	15.62
		BIC	1.457e+06	1.310e+06	1.144e+06	949948	718542
	Quadratic		5.31	1.45	-1.71	-4.58	-5.79
			-4.293	-4.652	-4.975	-6.087	-8.161
		AIC	15.62	15.62	15.62	15.62	15.62
		BIC	1.457e+06	1.310e+06	1.144e+06	949948	718542
2011	Linear		7.32	-0.36	-7.23*	-9.57**	-10.95*
			-4.575	-3.938	-3.927	-4.394	-5.374
		AIC	15.62	15.62	15.62	15.63	15.63
		BIC	1.468e+06	1.322e+06	1.155e+06	959916	727209
	Linear interaction		7.60*	-0.28	-7.38*	-9.96**	-10.66*
			-4.484	-3.974	-4.091	-4.657	-5.373
		AIC	15.62	15.62	15.62	15.63	15.63
		BIC	1.468e+06	1.322e+06	1.155e+06	959926	727221
	Quadratic		7.24	-0.38	-7.22*	-9.51**	-10.99**
			-4.371	-3.9	-3.968	-4.428	-5.145
		AIC	15.62	15.62	15.62	15.63	15.63
		BIC	1.468e+06	1.322e+06	1.155e+06	959924	727221

Notes: Outcome is birthweight. Errors are clustered by week of birth. * $p<0.1$. ** $p<0.05$. *** $p<0.01$.

Table A11 Full non-parametric results for birthweight from placebo cut-off tests

Placebo cut-off date		(1)	(2)	(3)	(4)
	Bandwidth	MSE-optimal	16 weeks	12 weeks	8 weeks
2007	Confidence intervals:				
	Conventional	1.57	0.99	-3.28	-7.47
		(4.980)	(4.185)	(4.661)	(6.561)
	Bias-corrected	1.91	-7.33*	-8.90*	2.64
		(4.980)	(4.185)	(4.661)	(6.561)
	Robust bias-corrected	1.91	-7.33	-8.90	2.64
		(7.946)	(6.651)	(8.127)	(5.969)
	MSE-optimal h	4.529			
	Effective n	114,143	397,500	294,489	191,495
2008	Confidence intervals:				
	Conventional	-2.75***	-2.96	-1.33	1.75
		(0.329)	(2.924)	(2.246)	(1.997)
	Bias-corrected	-3.94***	0.61	4.37*	-5.97***
		(0.329)	(2.924)	(2.246)	(1.997)
	Robust bias-corrected	-3.94*	0.61	4.37	-5.97
		(2.181)	(2.528)	(3.689)	(4.082)
	MSE-optimal h	2.559			
	Effective n	67,954	417,371	309,248	202,312
2010	Confidence intervals:				
	Conventional	22.08***	-3.27	-3.38	1.93
		(2.924)	(6.159)	(7.471)	(8.042)
	Bias-corrected	26.05***	-0.83	7.07	20.01**
		(2.924)	(6.159)	(7.471)	(8.042)
	Robust bias-corrected	26.05***	-0.83	7.07	20.01***
		(6.708)	(8.944)	(8.083)	(6.319)
	MSE-optimal h	3.376			
	Effective n	94,231	419,552	310,219	201,723
2011	Confidence intervals:				
	Conventional	0.15	-9.82**	-9.73**	-8.45
		(5.292)	(4.303)	(4.953)	(6.134)
	Bias-corrected	2.13	-9.05**	-6.75	0.43
		(5.292)	(4.303)	(4.953)	(6.134)
	Robust bias-corrected	2.13	-9.05	-6.75	0.43
		(6.094)	(6.383)	(6.608)	(5.128)
	MSE-optimal h	3.931			
	Effective n	94,911	424,999	314,320	203,954
	Triangular kernel	X	X	X	X
	MSE-optimal h	X			
	Linear	X	X	X	X

Notes: Outcome is birthweight. Model 1 is a non-parametric RD specific using the MSE-optimal bandwidth for each year. Errors are clustered by week of birth.
* p<0.1. ** p<0.05. *** p<0.01.

Table A12 Full results from covariate balance tests

Covariate		(1)	(2)	(3)	(4)	(5)	Control mean
% multiple births		0.002	0.001	0.002	0.005***	-0.111	3.30
		(0.003)	(0.001)	(0.002)	(0.002)	(0.133)	
	MSE-optimal h	2.629					
	Effective n	67,553	412,078	306,131	200,012	425,620	
% males		0.004	0.003	0.005	0.001	-0.110	51.18
		(0.004)	(0.004)	(0.004)	(0.003)	0.366	
	MSE-optimal h	9.182					
	Effective n	253,087	412,078	306,131	200,012	425,620	
Maternal age		0.059	0.045	0.011	0.064*	0.111***	29.40
		(0.043)	(0.035)	(0.038)	(0.034)	(0.040)	
	MSE-optimal h	4.042					
	Effective n	120,477	412,078	306,131	200,012	425,620	
English index of income deprivation		-0.004***	0.000	-0.000	-0.003***	-0.000	0.175
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	
	MSE-optimal h	2.635					
	Effective n	64,268	391,601	290,884	190,110	404,446	
Non-parametric		X	X	X	X		
MSE-optimal		X					
Robust bias-corrected		X					
Linear		X	X	X	X	X	

Note: Model 1 is a non-parametric RD specific using the MSE-optimal bandwidth for each year. Models 2-4 are non-parametric RD specifications using 16-, 12- and 8-week bandwidths respectively with robust bias-corrected standard errors. Model 5 is a parametric linear specification for a 16-week bandwidth. Errors are clustered by week of birth. * p<0.1. ** p<0.05. *** p<0.01.

Table A13 Investigating discontinuities in maternal age at placebo cut-offs

Placebo cut-off date		(1)	(2)	(3)	(4)	(5)	Control mean
2007		-0.18*	-0.14	-0.19*	-0.23**	0.09	29.43
		-0.107	(0.099)	(0.103)	(0.109)	(0.061)	
	MSE-optimal h	7.02					
	Effective n	191,495	397,500	294,489	191,495	621,106	
2008		-0.16***	-0.13***	-0.15***	-0.14***	0.05	29.41
		-0.037	(0.047)	(0.044)	(0.023)	(0.051)	
	MSE-optimal h	4.291					
	Effective n	121,677	417,371	309,248	202,312	645,076	
2010		0.04	0.06	0.04	0.16*	0.10**	29.46
		-0.081	(0.082)	(0.097)	(0.092)	(0.047)	
	MSE-optimal h	11.21					
	Effective n	310,219	419,552	310,219	201,723	652,854	

2011	-0.03	-0.02	-0.03	-0.04	0.08*	29.59
	-0.026	(0.036)	(0.033)	(0.042)	(0.042)	
MSE-optimal h	5.423					
Effective n	149,405	424,999	314,320	203,954	661,416	
Non-parametric	X	X	X	X		
MSE-optimal h	X					
Robust bias-corrected	X					
Linear	X	X	X	X	X	

Note: Outcome is maternal age. Model 1 is a non-parametric RD specific using the MSE-optimal bandwidth for each year. Models 2-4 are non-parametric RD specifications using 16-, 12- and 8-week bandwidths respectively with robust bias-corrected standard errors. Model 5 is a parametric linear specification for a 16-week bandwidth. Errors are clustered by week of birth. * p<0.1. ** p<0.05. *** p<0.01.

Table A14 Sensitivity analysis with date of birth as the running variable

	(1)	(2)	(3)	(4)	(5)	(6)	Control mean
Birthweight	12.28*	17.45	18.15	17.58			3322.79
	(6.931)	(10.419)	(11.267)	(10.438)			
AIC		8.370	8.369	8.366			
BIC		-157094	-157169	-157339			
MSE-optimal h	69.35						
Effective n	262,496	59,780	59,780	59,780			
Low birthweight	-0.005*				-0.002	-0.002	
	(0.003)				(0.002)	(0.002)	
MSE-optimal h	63.48						
Effective n	239,826				425,620	425,620	
Non-parametric	X						
MSE-optimal h	X						
Robust bias-corrected	X						
Linear	X	X					
Linear interaction			X				
Quadratic				X			
Probit					X		
Logit						X	

Notes: Standard errors, clustered by week of birth, in parentheses. Model 1 is a non-parametric RD specific using the MSE-optimal bandwidth for each variable with robust bias-corrected standard errors. Models 2-6 are parametric specifications all for a 16-week bandwidth around the cut-off: (2) is a linear polynomial; (3) is a linear interaction (ie spline), (4) is a quadratic polynomial; (5) is a probit model; (6) is a logit model. Errors are clustered by week of birth. * p<0.1. ** p<0.05. *** p<0.01.

Table A15 Full results from McCrary density test for the running variable

	Date of birth	Week of birth
t-statistic	-1.128	-5.632
p-value	0.259	0.000
Observations	1,401,432	1,389,797
Robust bias-corrected	X	X

Figure A1 Relationship between birthweight and maternal age decile

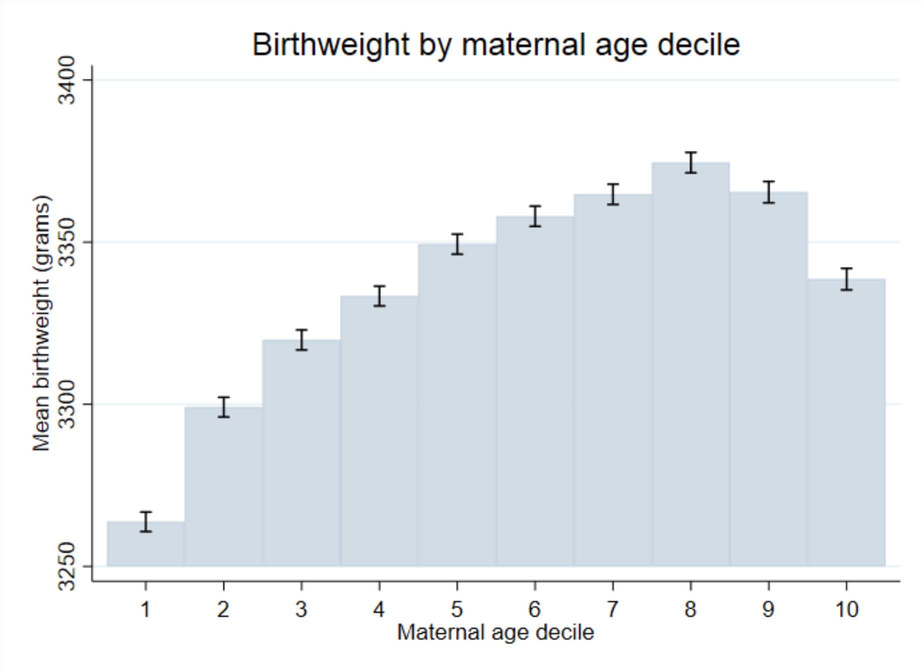


Figure A2 Relationship between birthweight and maternal age decile

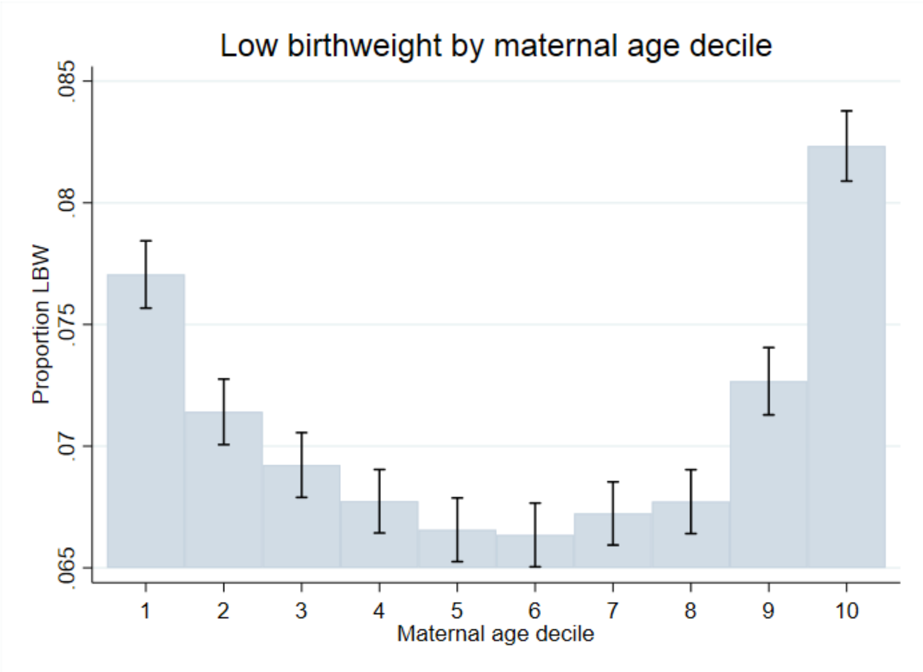


Figure A3 Relationship between birthweight and maternal age quartile

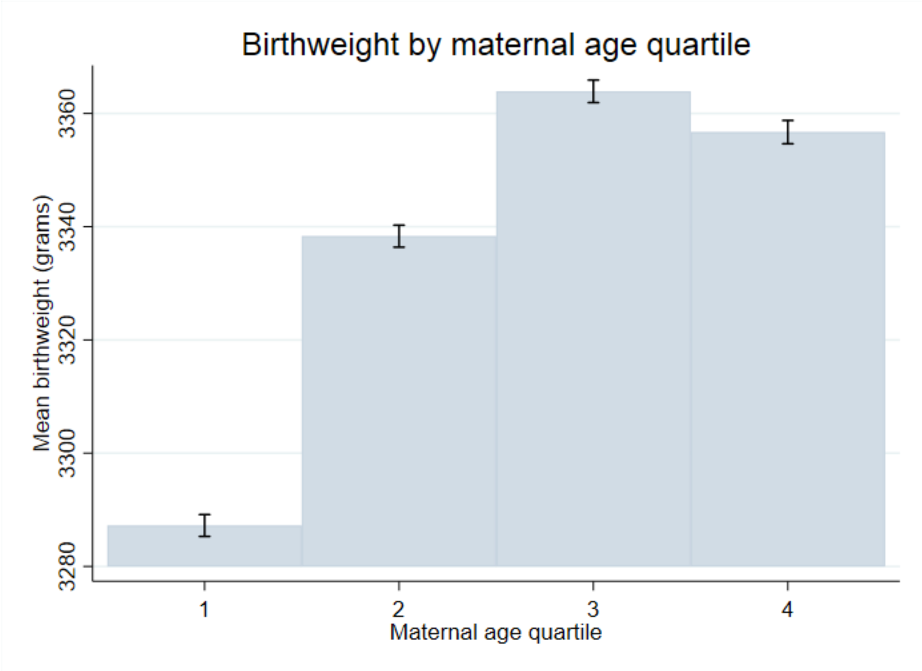


Figure A4 Relationship between low birthweight and maternal age quartile

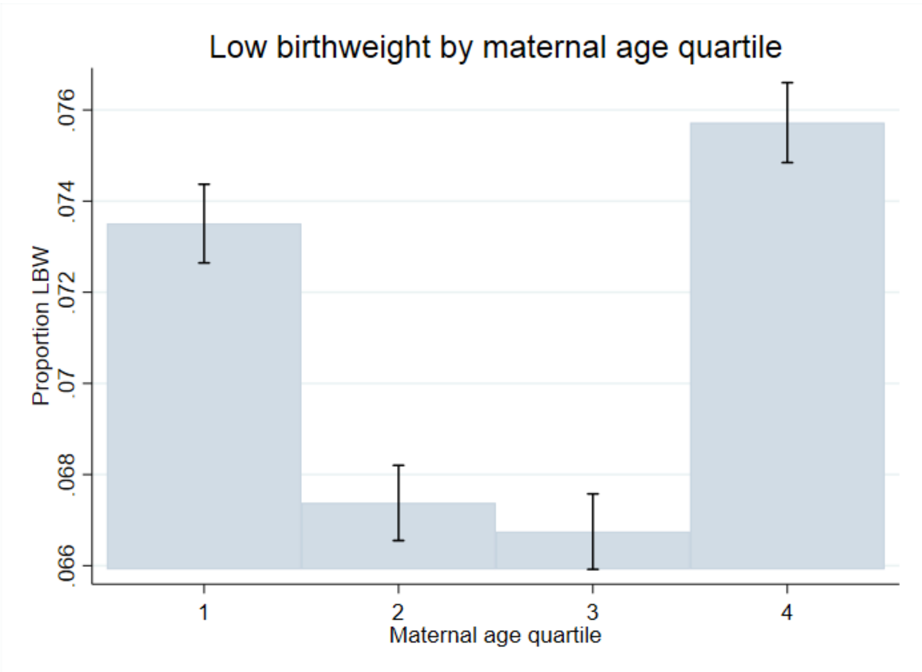


Figure A5 Relationship between birthweight and English index of income deprivation

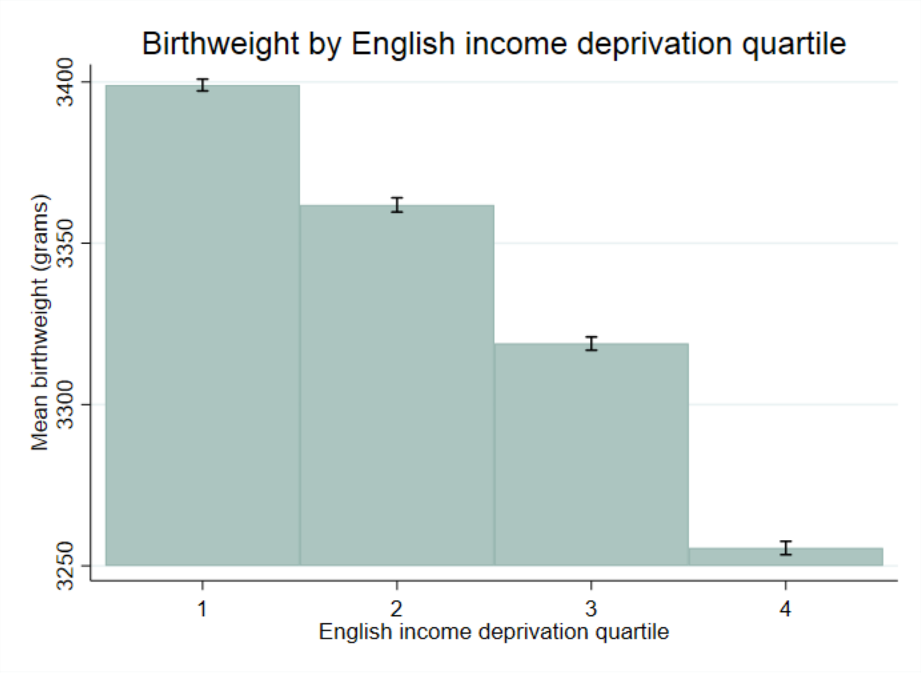


Figure A6 Relationship between low birthweight and English index of income deprivation

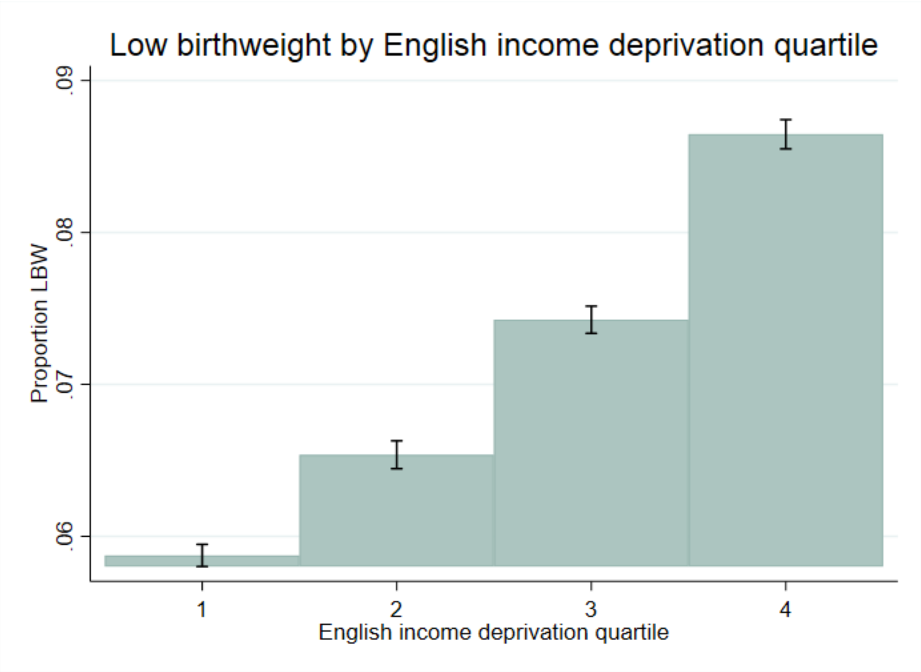
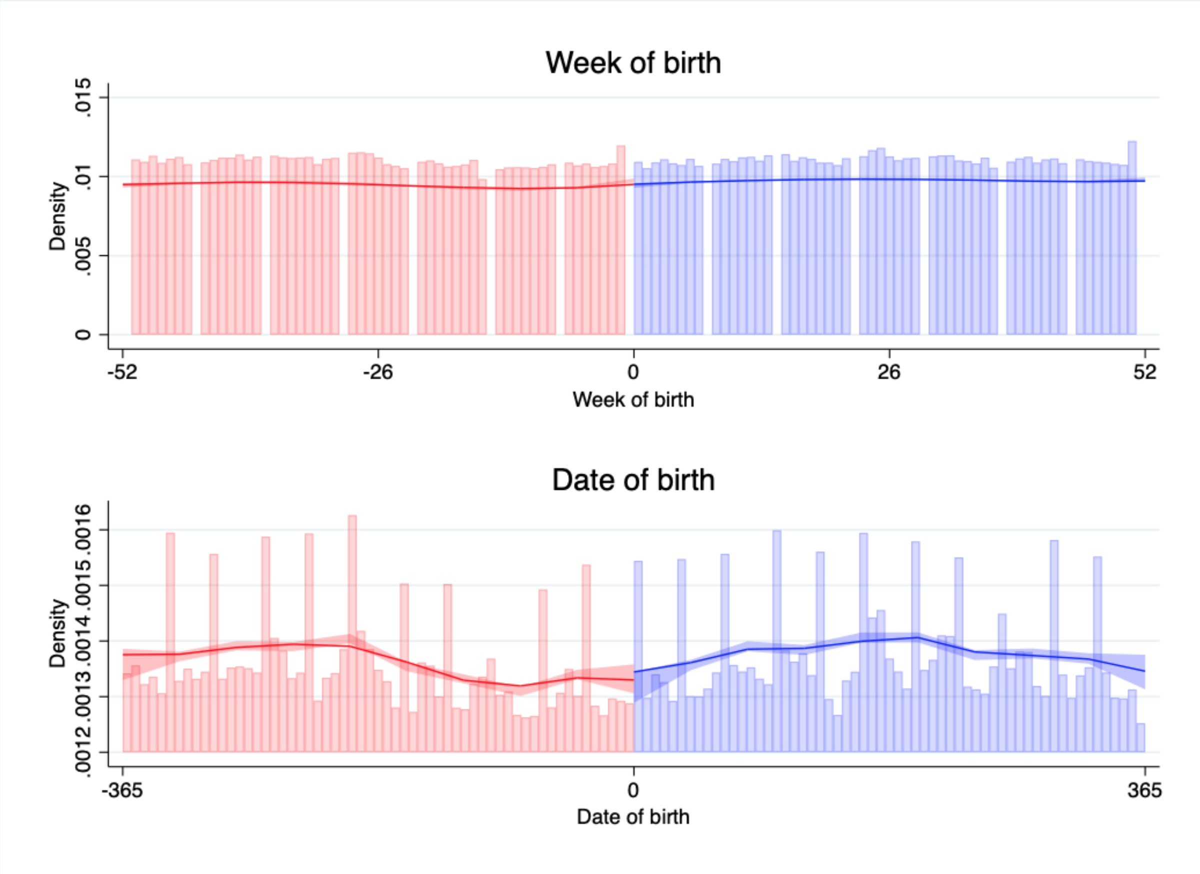


Figure A7 Density of births around the cut-off, by the running variable (week/date of birth)



Notes: 1. Histogram shows density of births by date and week of birth. 2. Overlapping confidence intervals at the cut-off indicate a failure to reject the null hypothesis of no discontinuity at the cut-off (i.e., no manipulation). 3. The daily variation in frequency of births by date of birth and the notable spikes in the graph are likely to be due to the managed scheduling of induced labour and elective caesarean sections during weekdays, with the consequence that Mondays have the highest numbers of births and the weekends have the lowest.