Baby Bonuses: natural experiments in cash transfers, birth timing and child outcomes

Nathan Deutscher*
Crawford School of Public Policy
Australian National University

Robert Breunig
Crawford School of Public Policy
Australian National University

November 26, 2016

Abstract

In this paper we use the 1 July 2004 introduction of the Australian Baby Bonus to identify the effect of family income on child test scores at grade three. We use a difference-in-differences design. We find no evidence the Baby Bonus improved child outcomes in aggregate, but some evidence of a modest effect for children from disadvantaged backgrounds. We examine whether birth shifting associated with the Baby Bonus and two other Australian maternity payments has negative long-term effects on children. Despite concerns about this unintended treatment, we find no clear evidence of health or educational consequences.

Keywords: Natural experiment; child outcomes; family benefits; birth shifting

JEL: H31, I38, J13

*The authors would like to thank the Australian Curriculum, Assessment and Reporting Authority (ACARA), Social Research Centre (SRC), Department of Education and Training (DET), Australian Institute of Health and Welfare (AIHW) and Australian Taxation Office (ATO) for the data used in this paper. Discussions with Steve Croft (ACARA), Megan O'Connell (SRC), Phil Aungles (DET) and Georgina Chambers (AIHW) were particularly helpful in refining data requests. All findings and views in this paper are those of the authors and should not be attributed to others. Author contact details: `nathan.deutscher@anu.edu.au; robert.breunig@anu.edu.au`

1 Introduction

In 2011, all OECD nations provided some form of cash benefits to families, at an average cost of 1.3% of GDP.[1] Despite their ubiquity, there is a great deal of variety in how and why countries boost family incomes. Parental leave policies look to encourage parents, often mothers, to care for their children during infancy. Broader cash transfers to families can be seen as a form of redistribution, in recognition of the costs of raising children, or an investment in later child outcomes. This latter view, of children as an investment rather than a cost, is one that seems to have gained increasing traction.[2]

Despite significant public expenditure and interest in family benefit programs, the literature on the *causal* effect of family income on child outcomes remains relatively modest. This reflects the difficulty in finding genuinely exogenous sources of variation in family income. As noted in Currie and Almond (2011), there are few policy examples that increase incomes without potentially affecting child outcomes through other channels.

This paper's primary contribution is to exploit a natural experiment in cash transfers at birth to identify their effect on educational outcomes at grade three. We use the Australian introduction of a $3,000 Baby Bonus for parents of babies born on or after 1 July 2004. This allows a simple and transparent difference-in-differences evaluation comparing the outcomes of those born either side of 1 July 2004 with those born in the same windows in the years either side. In exploiting a date of birth cut off in this way, this paper is most closely related to the relatively recent and growing literature evaluating expansions in maternity leave provisions.

The paper's second contribution is to present the first tests for long-run consequences of birth shifting events. The introduction of the Baby Bonus was accompanied by a large shift in births, as documented in Gans and Leigh (2009). This complicates the identification of the treatment effect of the Baby Bonus, but also provides an opportunity to test whether the large shift in births and associated infant health consequences were accompanied by any long-term effect on child outcomes. We also examine two other large birth shifting events in Australia associated with the 1996 Maternity Allowance policy and the 2006 expansion of the Baby Bonus.

We begin by describing the related literature and the Australian policy environment,

[1] https://data.oecd.org/socialexp/family-benefits-public-spending.htm
[2] See Figure A.1 in the Appendix.

which has been relatively atypical among OECD countries. The methodology is outlined in Section 3. In Section 4, the data are described along with the construction of outcome variables and the analysis sample. The results are presented and alternative specifications tested in Section 5. Section 6 places the results in the context of existing literature and concludes.

2 Background

2.1 Related Literature

A number of previous papers have attempted to identify the causal effects of cash transfers using instrumental variable techniques exploiting variations in cash transfers over time, household type and location. In Dahl and Lochner (2012), variations in the amount of Earned Income Tax Credit that households are eligible for are exploited to identify relatively modest effects of household income on children's test scores. They find an additional $1,000 in permanent income raises combined math and reading test scores by 6 per cent of a standard deviation in the short run, with larger effects for children from disadvantaged families. Exploiting variations in Canadian child tax benefits, Milligan and Stabile (2011) find somewhat larger effects on test scores, and also identify positive effects on physical health for boys and mental health for girls. By design, these studies are best suited to examining the contemporaneous effects of family income on child outcomes. Yet there is a rich and well known literature on the general importance of early life experiences to later childhood and adult outcomes; see Currie and Almond (2011). In Morris et al. (2005), the authors draw on seven random assignment welfare and poverty evaluations to examine the age-specific pattern of effects of welfare policies on child achievement. Earnings supplement programs are found to be most effective between ages 2 and 5 (although the treatment effect for those aged 0 to 1 is not statistically different). The Australian Baby Bonus offers an alternate natural experiment to identify causal effects on child outcomes of family income at birth.

In using a date of birth cut off, this paper is related to recent studies of expansions in maternity leave provisions. Using registry-based data, maternity leave expansions in Germany (Dustmann and Schönberg (2012)), Denmark (Rasmussen (2010)) and Sweden (Liu and Skans (2010)) have been found to have had little or no effect on the long-term outcomes of children. More recently, Carneiro et al. (2015) found large positive effects for children affected by maternity leave expansions in Norway — which they suggest may partly be due to these reforms coming off a lower base level of entitlement and background support for child care than the expansions examined by other authors. Looking at earlier life outcomes, Rossin (2011) finds positive effects on infant health from maternity leave expansions in the United States while Baker and Milligan (2010) find little impact of increased maternal care on child development in Canada. The Australian Baby Bonus differs from these maternity leave expansions in that it was unconditional and not linked to prior employment or subsequent caring arrangements. This institutional background is described further in Section 2.3.

Examining whether the unintended birth shifting that accompanied the Baby Bonus had any effect on child outcomes is also of broad interest. There are now numerous examples of birth shifting in response to announcement effects, and the scheduling of births for a wider range of reasons is increasingly common. For example, birth delays and increases in birthweight have been identified in response to German parental leave reforms; see Tamm (2013). In the opposite direction, Schulkind and Shapiro (2014) find relatively small birth bring-forwards in response to the United States' tax system, which are nonetheless associated with lower birthweight and lower Agpar scores. More broadly, birth shifting has been identified by parents seeking to avoid inauspicious days and care providers wishing to avoid weekends, public holidays and conferences; see Gans and Leigh (2012); Gans et al. (2007).

2.2 Policy Background

Australia has a long history of relatively unconditional cash transfers for new parents. The first 'Baby Bonus' was a near universal £5 maternity allowance for births from midnight 10 October 1912; see Kewley (1973).[3] The allowance was intended to help pay for the attendance of a health professional at childbirth, with then Prime Minister Fisher suggesting it would:

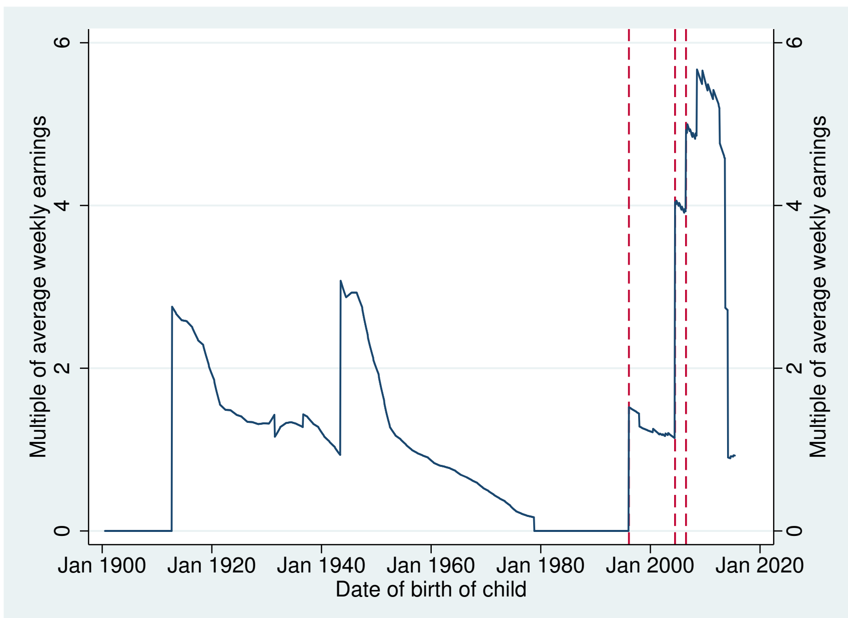
> *"be the means of helping poor parents to tide over an anxious period, and ensuring that their offspring's health, and perhaps life, shall not be jeopardised in the dawn*

[3]Near universal in that the allowance was not paid where the mother 'was an Asiatic or an Aboriginal of Australia, Papua or the Pacific Islands'; see Daniels (2009).

of existence ... I venture to say that it will not only protect many mothers, but to a great extent will preserve child life, and send it into the world equipped."

Maternity allowance was a popular initiative, but while subsequent reports found it led to an increase in the number of women attended by doctors there was no clear reduction in infant or maternal mortality; see Kewley (1973). This early version of maternity allowance was eventually allowed to erode in value before being abolished in 1978 (Figure 1).

Figure 1: Cash transfers on birth of a child.



Note: Dashed lines mark the policy changes examined in this paper: the 1 February 1996 reintroduction of Maternity Allowance; the 1 July 2004 introduction of the Baby Bonus; and the 1 July 2006 increase to the Baby Bonus. The above ignores income testing in place from 1 July 1931 to 1 July 1943 and from 1 January 2009 onwards and uses the minimum payment for a child where payments have varied according to the number of existing children in the family. Does not include the First Child Tax Refund, discussed in more detail below, or Paid Parental Leave, introduced from 1 January 2011. Historical payment rates are taken from Daniels (2009); *Family Assistance Guide* (2016) and average weekly earnings from Hutchinson (2016).

In recent times, Baby Bonuses have filled a hole left by the absence of mandatory paid maternity leave, introduced in Australia in 2011. In 1996, Australia and the United States were the only two OECD nations without mandatory paid maternity leave. Political pressure to lift support to new parents without disadvantaging stay-at-home mothers led to the reintroduction of, and increases to, cash transfers to new parents. It is this atypical experience that provides a series of natural experiments in the provision of cash transfers at birth.[4]

[4]In contrast to the original Baby Bonus, and perhaps more realistically, the stated aspirations for recent

2.3 The Baby Bonus

This paper focuses on the 1 July 2004 policy change — the introduction of the Baby Bonus, an unconditional cash grant of $3,000 on the birth of a child.[5] The Baby Bonus is one of the largest single changes in maternity payments in Australian history (see Figure 1). Interpreting the effect of the Baby Bonus is complicated by the fact it was not a uniform, unconditional increase in family incomes in *net* terms. The Baby Bonus replaced Maternity Allowance (a smaller unconditional cash transfer) and the variable First Child Tax Refund.[6]

Nevertheless, the Baby Bonus was an unambigious and large increase in entitlements for most families. Table 1 provides information about the distribution of the First Child Tax Rebate and the net benefits of the Baby Bonus. Around 75% of births in June 2004 (before the change) would have been better off under the new policy, while 44% would have benefited by the full $2,157 increase in cash transfers (as they were ineligible for the refund). Importantly, the change was more beneficial, on average, to parents from disadvantaged backgrounds: see Panels B and C of Table 1. For the mean (median) June 2004 birth the Baby Bonus would have been an increase of around $900 ($1,700) at birth — around one (two) week's average earnings.[7]

In addition, the extent of birth shifting suggests parents, on average, valued the Baby Bonus' introduction more than any other recent policy changes. In Table 2, estimates of the share of births delayed in the fortnight prior to the three largest policy changes since WWII are presented. The Baby Bonus estimates are taken from Gans and Leigh (2009), the Maternity Allowance estimate uses the same methodology and data as those authors to ensure comparability.[8] The introduction of the Baby Bonus induced by far the largest shift in

---

cash transfers to new parents have been more modest, and framed around assisting with the costs of parenthood. There may also have been political motivations, with the major policy changes — the introduction of Maternity Allowance in 1996 and the Baby Bonus in 2004 — coinciding with federal election years.

[5] This was originally named Maternity Payment, but has always been popularly known as the Baby Bonus, and was eventually renamed as such.

[6] The First Child Tax Refund is described in detail in Appendix C.

[7] Any financial assistance provided by the First Child Tax Refund is likely to have been more heavily discounted than suggested in Table 1. First, the refund was paid out over the five years following the birth of the child, lowering its net present value. Second, the refund depended on the (uncertain) future taxable income of the primary carer, which would lower the certainty equivalent cash transfer at birth for the risk averse. Finally, the refund was far from salient — claims were still being submitted up to five years after they could have been paid out. These features could explain the lack of any sizeable bring-forward in births by those who would have been better off under the refund compared to the Baby Bonus. Our examination of Australian Taxation Office data, not presented here, shows no significant increase in number or size of claims around the cut-off.

[8] Maternity Allowance is not examined in Gans and Leigh (2009) but its introduction induced birth shifting that is apparent in the data underlying that paper.

births. Assuming birth shifting in line with the other two policy changes implies an average net value for the Baby Bonus change of $1,200-2,700, broadly consistent with the figures in Table 1.[9]

Table 1: Distribution of First Child Tax Refund claims and (hypothetical) net gain from Baby Bonus introduction (June 2004 births).

	Mean	Percentile				
		10^{th}	25^{th}	50^{th}	75^{th}	90^{th}
Panel A: All carers						
First Child Tax Refund	$1,280	0	0	$485	$2,266	$3,348
Net gain from Baby Bonus	$877	$2,157	$2,157	$1,672	-$109	-$1,191
Panel B: Carers in bottom quintile by base income						
First Child Tax Refund	$1,093	0	0	$957	$2,269	$2,282
Net gain from Baby Bonus	$1,064	$2,157	$2,157	$1,200	-$112	-$125
Panel C: Carers in top quintile by base income						
First Child Tax Refund	$2,179	0	0	$331	$3,907	$6,958
Net gain from Baby Bonus	-$22	$2,157	$2,157	$1,826	-$1,750	-$4,801

Note: All figures are expressed in September quarter 2004 dollars and calculated from Australian Taxation Office data. Base income is the primary carer's income in the year prior to the birth of the child. Net gain is simply the $2,157 increase in cash transfers (the $3,000 Baby Bonus minus the $843 Maternity Allowance) minus the refund claim.

Table 2: Births delayed in response to changes in cash transfers at birth.

Policy	Date of effect	Conception effect?	Value (2004 dollars)	Share of births moved
Maternity Allowance	1 Feb. 1996	No	A$1,030	4%
Baby Bonus	1 July 2004	No	Variable	10%
Baby Bonus	1 July 2006	Possible	A$780	6%

Note: Estimates of the share of births moved in response to the introduction and increase to the Baby Bonus are from Gans and Leigh (2009), Column 2 of Tables 1 and 8 respectively. They take the logarithm of the daily birth count for the fortnights either side of the relevant threshold and regress on week, day of year and public holiday fixed effects, with an indicator for the fortnight after the relevant policy change. They use data from 1975-2004 for the 1 July 2004 Baby Bonus introduction and from 1975-2006 (excluding 2004) for the 1 July 2006 Baby Bonus increase. We use the same data and method and all years from 1975-2006 to estimate birth shifting for the 1 February 1996 Maternity Allowance reintroduction.

[9] Comparing the re-introduction of Maternity Allowance on 1 February 1996 and the increase to the Baby Bonus on 1 July 2006 reveals a clear increase in the propensity of parents to shift births in response to policy incentives. This may in part reflect the greater prevalence of caesarean sections, which rose from around 20% to 30% of births over this period.

3 Methodology

3.1 Effect of intended treatment

We use a difference-in-differences design to estimate the treatment effect of the additional family income provided by the Baby Bonus. This design compares the outcomes of children born shortly before and after the introduction of the Baby Bonus (in 2004) with those born in the same windows the years either side (2003 and 2005). Similar approaches have been used to evaluate expansions in maternity leave coverage in numerous countries (as an example, see Dustmann and Schönberg (2012)). In particular, we estimate:

$$y_i = \beta_0 + \beta_1 \mathbb{I}_{2004,i} + \beta_2 \mathbb{I}_{afterJuly,i} + \beta_3 \mathbb{I}_{2004,i}.\mathbb{I}_{afterJuly,i} + \beta_4 X_i + \varepsilon_i$$

where, for child i, y_i is their test score, $\mathbb{I}_{2004}$ indicates whether they were born in 2004 and $\mathbb{I}_{afterJuly,i}$ indicates whether they were born on or after 1 July. The vector of covariates X_i includes: sex; location (dummy variables for each state or territory and for whether the child lives in a metropolitan area); language background other than English; state specific cubic terms in the child's age-at-test; and parental demographics.[10]

As previously noted the introduction of the Baby Bonus induced a large shift in births, with around 1,000 births delayed until after 1 July 2004 (Gans and Leigh (2009)). This poses a threat to the validity of the identification strategy by potentially altering the composition and outcomes of those born either side of 1 July in the treatment year. To address this we exclude those born in the fortnight either side of 1 July. This is likely to be sufficient as the birth shifting was highly localised — nearly 20% of the 'shifted' births occurred on the first day of the policy and nearly 80% within the first fortnight (Gans and Leigh (2009)).[11] In addition, we show (see Table 6 below) that the results also hold when excluding all those born a month either side of the policy introduction date.

Although birth shifting poses a challenge when identifying the effect of the Baby Bonus'

[10] A student is defined as being of a language background other than English if either the student, the student's mother or the student's father speaks a language other than English at home. Parental demographics are captured by categorical variables covering the mother's and father's education — unknown, Year 9, Year 10, Year 11, Year 12, certificate, diploma, and bachelor degree or higher — and occupation — unknown, not in paid work, blue collar, white collar, management, and senior management.

[11] Note birth shifting outside this window need not imply that any births were moved more by than a fortnight to cross the 1 July threshold. Rather it may reflect flow-on birth reschedulings by providers to smooth out their workflows — for example moving births from early to late June or early to late July.

intended treatment — additional family income on the birth of a child — it also provides an opportunity. As Gans and Leigh note "with more data, this event provides an opportunity for health researchers and economists to study the impact of a large disruption in a well-developed, modern medical system". This is the focus of the second part of this paper.

3.2 Effect of unintended treatment

A regression discontinuity design is used to estimate the combined treatment effect of the additional family income and birth shifting that accompanied the Baby Bonus. Birth shifting may result in three key differences between those born immediately after the introduction of the Baby Bonus, relative to those born beforehand.

- **Hospital crowding:** the primary concern raised by Gans and Leigh (2009) was the additional pressure placed on the hospital system. This was also the focus of media reporting at the time. In all, 1,005 children were born on 1 July 2004, over 200 more than what might have otherwise been expected Gans and Leigh (2009).

- **Increased gestation length and birthweight:** a further concern was that birth shifting may lead to suboptimal infant health. Both low and high birthweight can be associated with poorer later life outcomes; see Thorngren-Jerneck and Herbst (2001). Thus, if planned births are postponed for longer than optimal, it is possible that child health and outcomes will be compromised.

- **Differences in family backgrounds:** finally, it may be that some families are better able to shift births around the introduction date than others. These same families may also differ in ways (observable or not) that matter for their children's later life success. This could raise equity concerns around the introduction of the policy.

The existing work by Gans and Leigh (2009) provides a good sense of the magnitude of birth shifting that accompanied the introduction of the Baby Bonus. The birth shifting and potential hospital crowding is thus relatively well understood. A regression discontinuity approach is used to estimate other discontinuities around 1 July 2004. As recommended in Lee and Lemieux (2010) we use local linear regressions either side of the threshold, allowing both intercept and slope to vary. Let h_i be the health or family background variable of interest. We now estimate the regression below in a window around the introduction of the Baby Bonus, where t is number of days after 1 July 2004 and $\mathbb{I}_{afterJuly,i}$ remains an indicator for whether the individual was born on or after 1 July 2004.

$$h_i = \alpha_0 + \alpha_1 \mathbb{I}_{afterJuly,i} + \alpha_2 t + \alpha_3 \mathbb{I}_{afterJuly,i} t + \alpha_4 X + \varepsilon_i \tag{1}$$

The estimated discontinuity in the expected value of h at 1 July 2004 is now α_1, and standard hypothesis testing can be used. The same regression can be used to test for a discontinuity in test scores — whether the policy-induced differences in family income, background (due to birth shifting) and infant health had any lasting consequences. A vector of covariates X is included in the regressions to control for day of week effects in health and family background variables. In testing for a discontinuity in test scores we use the same vector of covariates as in the earlier difference-in-differences design.[12] To complement this analysis, we run similar regressions testing for lasting consequences from the two other policy changes in Table 2.

A key question in the regression discontinuity literature is the choice of bandwidth, balancing the reduced variance that comes with a larger bandwidth against any bias arising from the inaccuracy of the linear specification over too large a range. Two broad approaches are used: estimating an optimal bandwidth given the joint distribution of all variables, beginning with estimating a rule-of-thumb bandwidth, and cross-validation procedures (Lee and Lemieux (2010)). In our case, for local linear regression with a rectangular kernel, the rule of thumb bandwidth is given by:

$$bw_{ROT} = 2.702 \left[\frac{\tilde{\sigma}^2 R}{\sum_{i=1}^{N} \{\tilde{m}''(x_i)\}^2} \right]^{\frac{1}{5}}$$

where $\tilde{m}''(.)$ is the second derivative (curvature) of a regression of h on the independent variables in Equation 1, $\tilde{\sigma}$ is the estimated standard error of the regression and R is the range of the assignment variable, here date of birth, over which the regression is estimated. For simplicity and given the nature of the results, we simply present the regression discontinuity estimates over a range of bandwidths, including for the rule of thumb bandwidth

[12] We thus only test for a discontinuity in expected test scores adjusted for observable family background traits.

when estimating any discontinuity in outcomes. Following Lee and Lemieux (2010) we use a quartic in the assignment variable when calculating the rule of thumb bandwidth. Note these regression discontinuity estimates capture the joint treatment effect of the additional family income, hospital crowding, infant health and (unobservable) family background differences.[13] It would be difficult to convincingly disentangle these effects even with unit record data linking family background, infant health and later life outcomes, which is not available here. Given this, our estimates should be interpreted as a combined treatment effect.

4 Data and sample selection

4.1 Child outcomes data

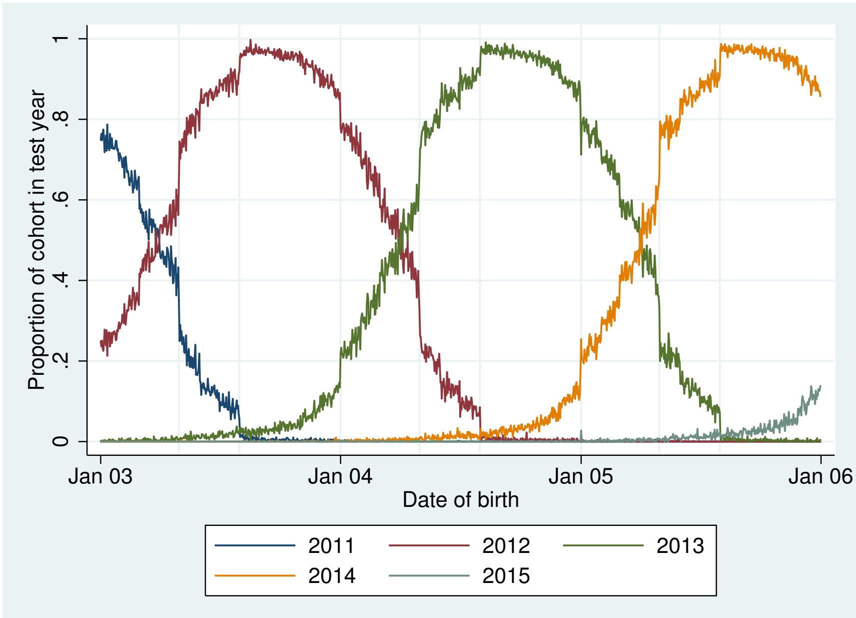
Early child outcomes are drawn from Australia's national system of school testing: the National Assessment Program — Literacy and Numeracy (NAPLAN).[14] Since 2009, students in Years 3, 5, 7 and 9 have sat the NAPLAN tests in mid May of each year. Students are tested across five domains — reading, numeracy, spelling, grammar and written work. Raw marks in each of these domains are mapped onto an achievement scale spanning all four year levels and ranging from approximately 0 to 1000. These measures of student achievement form our initial outcome variables. Details of the process of mapping raw marks onto these scales are in Australian Curriculum, Assessment and Reporting Authority (2015); many of the procedures are similar if not identical to those used in the Programme for International Student Assessment (PISA).

In Australia, primary and secondary education policy is primarily the responsibility of the six state and two territory authorities. An important difference between states and territories is in their school commencement policies: children can only start school if they will be six years old by the relevant cut off date. The degree to which children eligible to start school can be held back ('redshirted') also differs across states. In the years examined, three jurisdictions

[13] The introduction of the Baby Bonus is a classic case where a regression discontinuity design would not be a valid approach to estimating the treatment effect of the policy alone. While individual control over the assignment variable is imprecise in some senses — a delayed caesarean section may need to be brought forward for health reasons — it is nonetheless precise enough for some individuals to generate a discontinuity in the distribution of births over the assignment variable. For this not to invalidate the design would require no differences relevant to child outcomes between families with an a priori discontinuous conditional expectation of date of birth relative to those with an a priori continuous conditional expectation of date of birth.

[14] The NAPLAN data used in this publication are sourced from the Australian Curriculum, Assessment and Reporting Authority (ACARA) and are available from ACARA in accordance with its Data Access Protocols.

Figure 2: Distribution of individuals across test years by date of birth



Note: Excludes jurisdictions with 1 July cut off dates. Vertical lines indicate remaining cut off dates for school entry (30 April, 31 July and 1 January) and correspond to discontinuous jumps in the probability of sitting the test in a later year.

— Queensland, Western Australia and the Northern Territory — have 1 July cut off dates. They also have virtually no redshirting. As a result those born either side of 1 July 2004 in these jurisdictions are in different test years and differ by nearly a year in the age at which they sit the test. This difference in age at test induces a large discontinuity in test scores (Appendix A, Figure A.2) which would render the regression discontinuity design invalid. These jurisdictions are dropped from the sample. The remaining jurisdictions account for around two thirds of the population of interest.

As noted earlier, we focus on those born from 2003 to 2005. These children can end up in different test years based on state and territory policies and whether they were held back (see Figure 2). We restrict attention to those sitting the test in the year in which they turn 8, 9 or 10. This is necessary as we only have NAPLAN data for Year 3 students sitting the tests from 2011 to 2015.[15] It is also a fairly benign restriction, as it captures the vast majority of each cohort: only 0.05% of those born in 2004 in our data took the test outside this window.

[15]Specifically, we have NAPLAN data from 2013, 2014 and 2015. This includes students' past results, which allows us to infer the Year 3 test results for students who sat the tests in 2011 and 2012 and were eligible to sit the Year 5 tests in 2013 and 2014. This gives us Year 3 test results from 2011-2015 inclusive.

For the difference-in-differences design the key identifying assumption is that of common trends across treatment and control. Here this requires that the relationship between date of birth and test score remains the same across birth years — the relative penalties or bonuses from being born in a certain part of the year do not change. This could fail in a number of ways. First, NAPLAN tests change between years, and different tests may, unintentionally, have different age-at-test effects. A large amount of effort goes into the process of mapping raw marks to the NAPLAN achievement scale to provide a measure that is comparable across the years (Australian Curriculum, Assessment and Reporting Authority (2015)), but this may nonetheless be insufficient for our purposes. Trends in redshirting may also influence the relationship between when a child is born during the year and their eventual test scores. Some parts of the year are more prone to the influence of redshirting than others: Figure 2 shows the distribution of children across test years changes most rapidly prior to 1 May. Importantly, test year is not something that can sensibly be controlled for in our regressions as it is in part endogenous to ability — students are probably more likely to be held back, all else equal, if it is felt they are not yet ready for school.

To inform the choice of tests and windows we estimate the treatment effects for a placebo policy introduced on 1 July 2005, using 2005 as the treatment year and 2003 as the control. This tests the common trends assumption for the two control years. Of course, it is possible there are variations in the age-at-test effect in 2004 for the chosen tests and windows that are not apparent across 2003 and 2005, so we later test the sensitivity of our results in other ways. The outcome variables are (normalised) test scores for reading, numeracy, spelling, grammar and written work. Windows are progressively widened to look one, two or six months either side of 1 July (less the fortnight either side of July 1). A hybrid window from May to December is also included to make the most of the relative stability in test year distribution in the second half of the year. Results are in Table 3. Highly statistically significant differences are found for the spelling and written tests across most windows and for the full year window across all but one test. These deviations from the common trends assumption are also apparent from a visual inspection of test scores by date of birth, as shown in Appendix A, Figure A.3. Based on these results we restrict attention to reading and numeracy test scores and the first three windows.[16] Reading and numeracy test scores

[16]Grammar scores are excluded as they arise from the same test as spelling scores. While spelling and

Table 3: Estimated treatment effects of placebo policy on component tests.

Window	Jun-Jul (1)	May-Aug (2)	May-Dec (3)	Jan-Dec (4)
Reading	-0.007	-0.018	-0.008	-0.008
	(0.021)	(0.012)	(0.010)	(0.007)
Numeracy	-0.022	-0.006	-0.000	0.036***
	(0.021)	(0.013)	(0.010)	(0.007)
Spelling	-0.048**	-0.027**	-0.021**	-0.032***
	(0.021)	(0.013)	(0.010)	(0.007)
Grammar	-0.023	-0.019	-0.011	-0.017**
	(0.021)	(0.012)	(0.010)	(0.007)
Written	-0.025	-0.025**	-0.028***	-0.080***
	(0.020)	(0.012)	(0.010)	(0.007)

Note: All windows exclude a fortnight either side of 1 July.
Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

are summed to generated a combined test score, which is normalised to have mean zero and standard deviation one for ease of interpretation.[17]

Reading and numeracy test scores are available for 90% of students. Missing test scores are potentially problematic for two reasons. First, missing test scores are not random. Students with missing scores are more likely to have parents with no further education beyond high school and who are unemployed; children from such backgrounds who are not missing test scores have lower test scores on average. Second, the pattern of missing test scores across the calendar years does not follow the common trends assumption, as is clear from Appendix A, Figure A.4. However, if we assume those with missing test scores tend to perform worse, then the trends observed will, if anything, bias our results downwards.[18]

grammar have different questions on this test, this could nonetheless generate an interdependency.

[17]This data driven rejection of spelling, grammar and written work test scores as outcome variables also has a sound intuitive basis — there are reasons why these scores may be less discriminating and thus harder to equate across years. The reading and numeracy scores are based on separate tests which, for Year 3 students in 2014, had 38 and 35 questions respectively (marked either as right or wrong). In contrast, spelling and grammar scores were based on only 25 and 26 questions respectively, and were bundled into the same test. While other tests are year-level specific, the written work score is based on a persuasive response to a writing prompt that is identical across year-levels. This is marked on ten criteria, and while the total mark is out of 48 there was evidence of multi-modal distribution in 2014 that made equating scores more difficult, see Australian Curriculum, Assessment and Reporting Authority (2015).

[18]There is a lower prevalence of missing scores in the second half of 2004 relative to what one would expect

4.2 Births and birth outcomes data

As a prelude to looking for lasting consequences of birth shifting we extend the work on intermediate infant health outcomes by Gans and Leigh (2009). To do this we use Australian Institute of Health and Welfare (AIHW) data on daily birth outcomes in 2004. These are drawn from the National Perinatal Data Collection, which includes information from midwives and other health professionals attending births. In particular, we have information on the distribution of gestational age at birth and birthweight by date of birth cohort. In Gans and Leigh (2009), the authors had birthweight data for only three jurisdictions (covering 36% of births) and did not have gestational age data. Further details of the AIHW data can be found in Australian Institute of Health and Welfare (2014).[19]

4.3 Additional child outcomes data

We use the NAPLAN test score data to look at the combined treatment effect of the 1 July 2004 Baby Bonus introduction. However, the time periods spanned by the NAPLAN data prevent us from looking at the re-introduction of Maternity Allowance on 1 February 1996 or the increase in the Baby Bonus on 1 July 2006. To look at the effects of these increases in cash transfers at birth we use two alternative outcome measures.

For Maternity Allowance, we use aggregate data on university offers by precise date of birth. This covers all offers made to Australian residents by Australian universities in either 2013 or 2014 — the years in which the majority of those born around 1 February 1996 will have been completing high school and applying for university. More information on the Higher Education Statistics is available at Department of Education and Training (2016). For the Baby Bonus increase, we use unit record data including precise date of birth from the Australian Early Development Census (AEDC).[20] The AEDC is a triennial census of children in their first year of full-time school. Data are collected from schools

based on trends in 2003 and 2005 and the prevalence of missing scores in the first half of 2004. If those missing scores do worse, this will lead scores in the second half of 2004 to be lower than expected.

[19] Data on daily birth counts, used in Table 2 and Figure 5 is from the Australian Bureau of Statistics (ABS). These data are precisely as used and described in Gans and Leigh (2009) and have been provided by those authors. They are drawn from state and territory birth registers and includes all registered births in Australia. Further details can be found in Australian Bureau of Statistics (2014).

[20] The AEDC is funded by the Australian Government Department of Education and Training. The findings and views reported are those of the authors and should not be attributed to the Department or the Australian Government.

and teachers using the Australian version of the Early Development Instrument (AvEDI), adapted from the Canadian Early Development Instrument. The data include information on child development across five domains: physical health and wellbeing; social competence; emotional maturity; language and cognitive skills; and communication skills and general knowledge. Children have scores from 0-10 in each domain — a higher score indicates a higher level of development. We sum these scores to give a single outcome metric.[21] More information on the AEDC is available in Australian Government (2013).

5 Results

We begin with the difference-in-differences estimates of the treatment effect of the additional family income provided by the Baby Bonus, including estimates for population subgroups and alternative specifications. We then extend the work of Gans and Leigh (2009) looking at birth shifting in response to the introduction of the Baby Bonus. This provides a better understanding of the potential treatment applying immediately after 1 July 2004 — additional family income, hospital crowding, and differences in infant health and family background. We use a regression discontinuity approach to assess whether this combined treatment had any effect on child outcomes. We also test for discontinuities in outcomes associated with the 1 February 1996 Maternity Allowance reintroduction and 1 July 2006 Baby Bonus increase.

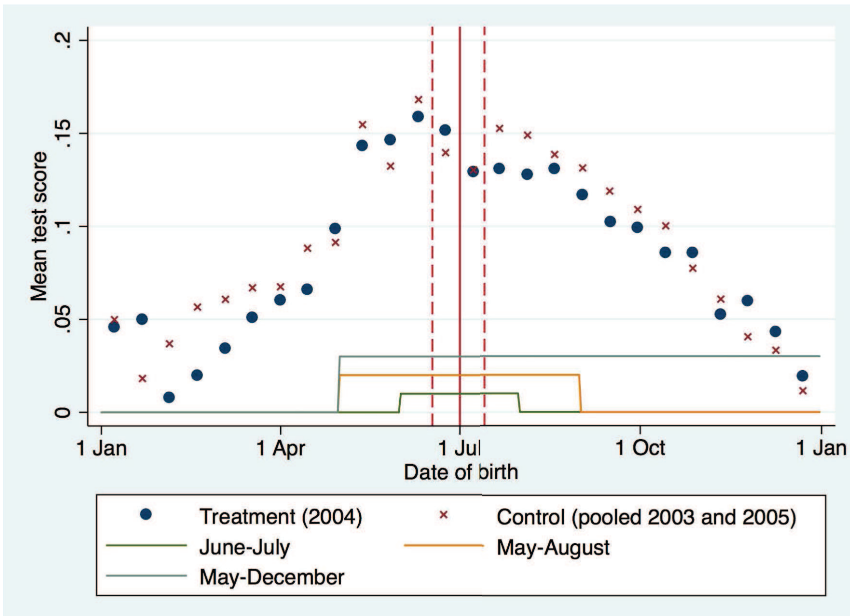
5.1 Child outcomes: difference-in-differences

In Figure 3 we plot average test scores by date of birth for the treatment and control groups in fortnightly bins either side of 1 July. This provides an immediate feel for the results that follow and allows a visual inspection of the assumptions underlying the difference-in-differences design. Note the inverted U-shape apparent for both treatment and control groups simply reflects the well known age-at-test effect — older children tend to get better results. In our sample, the interaction of school entry policies and redshirting behaviours result in average age-at-test following an inverted U-shape, see Figure A.5 in Appendix A. Similarly,

[21] This combined score is coded as missing if any one of a child's five domain scores is missing or flagged as invalid. Domain scores are flagged as invalid for children who have been in the class for less than a month, are less than four years old, have special needs or where teachers complete less than 75 per cent of the items in any given domain. This affects less than 7 per cent of children. We also conducted the analysis keeping 'invalid' scores — this had no meaningful impact on the results.

the apparent discontinuity between April and May reflects a discontinuity in age induced by the school starting age cut off in the state of Victoria and in the Australian Capital Territory.

Figure 3: Mean test score by date of birth



Note: means are calculated for fortnightly bins either side of 1 July. Dashed lines indicate the exclusion window — a fortnight either side of 1 July.

Figure 3 provides some comfort with respect to the common trends assumption required for the difference-in-differences design. Here this requires the test score by date of birth profiles to be parallel, aside from any difference emerging with the introduction of the policy. This assumption appears reasonable in the May to December window — the test scores profiles have the same basic shape. Perhaps the clearest violations are in the beginning of the year (and we have already rejected using a full year window of analysis) and around the introduction of the Baby Bonus.

A visual inspection of Figure 3 leaves mixed impressions as to any impact of the Baby Bonus. There is a hint of a *negative* discontinuity around the 1 July threshold, which is investigated further in Section 5.2. Yet when we consider wider windows, scores in the treatment year (on average a little lower than those in the control) do not appear to have been shifted by the introduction of the Baby Bonus.

Table 4 provides difference-in-differences estimates of the effect of the additional family income provided by the Baby Bonus over progressively wider windows of analysis. All re-

Table 4: Estimated treatment effects of the Baby Bonus

Window	Jun-Jul		May-Aug		May-Dec	
	(1)	(2)	(3)	(4)	(5)	(6)
Covariates	No	Yes	No	Yes	No	Yes
β_3	-0.009	-0.014	-0.017	-0.014	-0.005	-0.002
	(0.020)	(0.018)	(0.012)	(0.010)	(0.009)	(0.008)
N	46,349	45,971	132,531	131,477	306,178	303,630
R-squared	0.000	0.201	0.000	0.199	0.001	0.193

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

gressions exclude those born a fortnight either side of introduction and report estimates with and without covariates. Heteroskedacity-robust standard errors, clustered within schools, are reported alongside the results. The results confirm the visual impression left by Figure 3. All point estimates are small and none are statistically different from zero. In an encouraging sign for the validity of the design, none of the estimates are statistically different from each other — violations of the common trends assumption could have seen the window of analysis and/or the inclusions of covariates leading to significantly different estimates. Using the treatment effect presented in column (6) as our benchmark, a one-sided t-test would lead us to reject at the 5% level the hypothesis that the Baby Bonus lifted average test scores by at least 1.2% of a standard deviation.

5.1.1 Subpopulation estimates

It may not be surprising that the small, temporary boost to family income provided by the Baby Bonus had no discernible effect across the population as a whole. Improved outcomes are more plausible in disadvantaged households with lower household incomes and more significant credit constraints. To test for effects here, we repeat regressions (2), (4) and (6) from Table 4 for different population subgroups: those where the highest level of parental education is high school dropout, those where the highest level of parental education is high school (including dropouts), those with a university degree, and those where the highest status parental occupation is not employed, and those where the highest status parental occupation is senior management. The results are in Table 5. Covariates are included in all regressions to increase precision.

The results are mixed, though they are consistent with the introduction of the Baby Bonus improving outcomes for children from disadvantaged backgrounds. While only one of the point estimates is statistically significant, the treatment effects for those from disadvantaged backgrounds are typically positive and larger than those from more advantaged backgrounds. For the children of those with no more than a high school education, the estimated treatment effect on test scores is 4% of a standard deviation, and statistically significant at the 5% level. These children make up 19% of the sample and have an average percentile rank of 36.8. The effect is still relatively modest; for children between the 20^{th} and 80^{th} percentiles of the test score distribution an increase of 4% of a standard deviation in their score will increase their rank by about 1.5 percentiles.[22] A caveat on the statistical significance of these results is that the uncertainty in transforming raw grades to NAPLAN scores is not accounted for: true standard errors will be slightly larger than presented here.

5.1.2 Sensitivity analysis

As shown earlier, the identifying assumptions underlying the difference-in-differences design fail for some NAPLAN test domains and for the full calendar year window. Testing the sensitivity of the above results is thus particularly important. In Table 6, we repeat regression (3) from Table 5 under alternative specifications. In (1), all those born in June and July are dropped to allow for birth shifting over a wider window. In (2) and (3), the control year of birth is restricted to 2003 or 2005 respectively. Finally, in (4) and (5), results are estimated separately for New South Wales and Victoria — the two most populous states in Australia, accounting for 85% of our sample. This latter test is particularly useful as the two states have different education policies (including school entry cut off dates).

The results are mixed. First, the estimates change little when both June and July are excluded. For the remaining specifications the main finding of no statistically significant effect (and an ability to reject even modest effects) is robust. The suggestive evidence of a positive effect for those from disadvantaged backgrounds is more tenuous: the case for such

[22] While the relatively more disadvantaged group of children with high school dropout parents has a smaller and less significant treatment effect, it is not statistically significantly different from that for the larger group that includes high school graduates. In addition, children from severely disadvantaged backgrounds may face challenges that are not as easily ameliorated by additional family income as those a little less worse off.

Table 5: Estimated treatment effects of Baby Bonus policy: subpopulations

Window	Jun-Jul	May-Aug	May-Dec
	(1)	(2)	(3)
Panel A: Full sample			
β_3	-0.014	-0.014	-0.002
	(0.018)	(0.010)	(0.008)
N	45,971	131,477	303,630
R-squared	0.201	0.199	0.193
Panel B: Highest parental education = high school dropout			
β_3	0.031	0.029	0.029
	(0.053)	(0.031)	(0.025)
N	4,617	13,246	29,920
R-squared	0.081	0.074	0.072
Panel C: Highest parental education = high school or less			
β_3	0.004	0.022	0.040**
	(0.039)	(0.023)	(0.018)
N	8,383	24,002	54,566
R-squared	0.119	0.114	0.112
Panel D: Highest parental education = university degree			
β_3	-0.009	-0.019	-0.002
	(0.030)	(0.018)	(0.014)
N	17,223	49,221	115,725
R-squared	0.093	0.086	0.079
Panel E: Highest parental occupation = not employed			
β_3	-0.041	0.040	0.033
	(0.059)	(0.035)	(0.028)
N	3,891	11,236	25,324
R-squared	0.143	0.136	0.128
Panel F: Highest parental occupation = senior manager			
β_3	0.018	0.002	0.009
	(0.036)	(0.021)	(0.017)
N	11,812	33,455	78,481
R-squared	0.119	0.118	0.108

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 6: Key estimated treatment effects under alternative specifications.

	Robustness test				
	(1)	(2)	(3)	(4)	(5)
Exclude	± month	± fortnight	± fortnight	± fortnight	± fortnight
Control	pooled	2003	2005	pooled	pooled
States	all	all	all	NSW	VIC
Panel A: Full sample					
β_3	-0.001	-0.004	0.000	-0.009	0.002
	(0.010)	(0.010)	(0.010)	(0.012)	(0.014)
N	257,659	196,975	205,968	147,821	109,495
R-squared	0.192	0.192	0.198	0.217	0.163
Panel B: Highest parental education = high school dropout					
β_3	0.014	0.018	0.043	0.025	0.002
	(0.031)	(0.030)	(0.029)	(0.036)	(0.042)
N	25,303	19,934	19,690	14,814	10,044
R-squared	0.072	0.077	0.070	0.069	0.058
Panel C: Highest parental education = high school or less					
β_3	0.039*	0.035	0.046**	0.051*	0.007
	(0.022)	(0.022)	(0.021)	(0.027)	(0.030)
N	46,183	36,100	36,315	25,692	19,543
R-squared	0.112	0.111	0.119	0.116	0.091
Panel D: Highest parental education = university degree					
β_3	-0.005	-0.001	-0.002	-0.014	0.008
	(0.017)	(0.017)	(0.016)	(0.021)	(0.022)
N	98,502	73,512	80,331	55,886	44,373
R-squared	0.077	0.081	0.080	0.089	0.069
Panel E: Highest parental occupation = not employed					
β_3	0.048	0.017	0.048	0.020	-0.006
	(0.034)	(0.034)	(0.032)	(0.044)	(0.041)
N	21,433	16,219	17,323	11,120	11,240
R-squared	0.127	0.131	0.137	0.123	0.124
Panel F: Highest parental occupation = senior manager					
β_3	0.009	0.020	-0.002	0.006	0.003
	(0.020)	(0.019)	(0.019)	(0.024)	(0.027)
N	66,669	50,240	53,871	39,488	27,463
R-squared	0.107	0.112	0.110	0.130	0.083

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

an effect is weaker when the control year of birth is 2003 and is absent in the Victorian subsample. The earlier effects could thus be driven by unobserved changes in education practices in New South Wales, for example, rather than the introduction of Baby Bonus.

As a final test, we extend the earlier exercise estimating treatment effects for a placebo policy introduced on 1 July 2005 in Appendix B, Table B.1. This time we report results for reading, numeracy and combined test scores by population subgroup for a May to December window of analysis. It is possible the initial failure to find significant effects for the placebo policy only held for the separate tests and/or the full population, yet none of the estimates are statistically significant, which provides some support for the validity of our design.

5.2 Effect of combined treatment on child outcomes

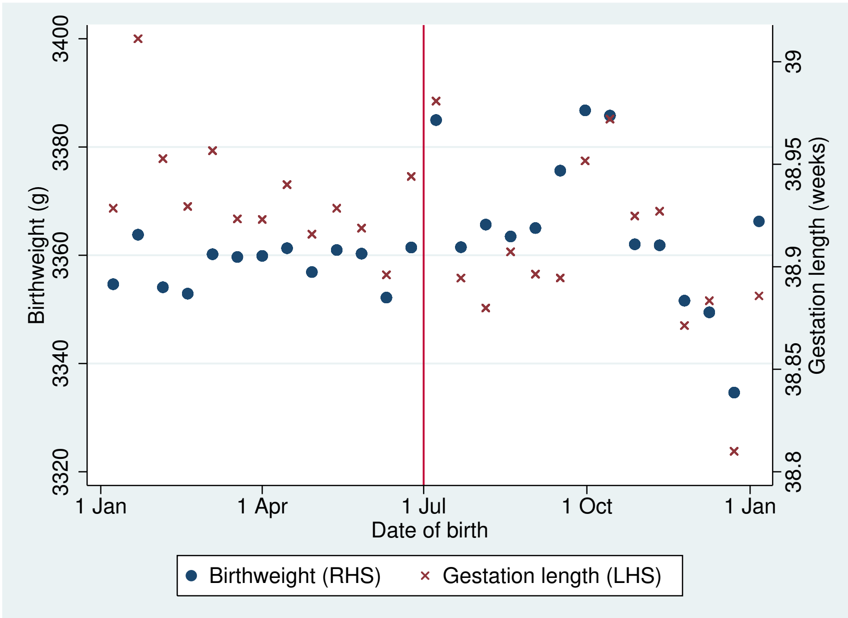
5.2.1 The unintended treatment

Policy-induced birth shifting can change a child's start in life through hospital crowding and infant health. Birth shifting may also result in differences in family background either side of the policy threshold. These differences and any long run repercussions shine a light on potential unintended harms from birth shifting and, perhaps more benignly, any inequities resulting from differing abilities or propensities to schedule births.

Large birth shifting in response to the Baby Bonus has already been documented by Gans and Leigh (2009) — there were more births on 1 July 2004 than any prior day in Australian history. As reported in Table 2, it is estimated that some 10% and 6% of births were shifted across the fortnights either side of the introduction and increase to the Baby Bonus respectively. For the reintroduction of Maternity Allowance, we find less birth shifting with 4% of births moved. These are all sizeable increases in demand on the health system.

Birth shifting also has direct, mechanical consequences for infant health: delays, all else being equal, lead to longer gestations and higher birthweights. Figure 4 shows mean gestation length and birthweight for all births in 2004. There is a clear spike in birthweight in the fortnight following the Baby Bonus' introduction. Interestingly, gestation length appears to rise either side of the introduction of the policy. This could be consistent with unsuccessful attempts to shift births into July, or providers delaying procedures to fill scheduling gaps left by those successfully shifting.

Figure 4: Mean birthweight and gestation length by date of birth (2004)



Note: means are calculated for fortnightly bins either side of 1 July 2004.

Table 7 presents estimates of the discontinuity in expected parent and child characteristics at 1 July 2004. Panels A and B use the demographic variables in the NAPLAN data. Panel C uses infant health variables aggregated across child date of birth cohorts. In both cases heteroskedacity-robust standard errors are reported. Standard errors are clustered by date of birth in the first two panels, while in the second the regression is weighted by the number of births in each date of birth cohort.

The results are consistent with and extend earlier findings. Gans and Leigh (2009) find only modestly significant differences in parental age and none in mean income of parental postcode. Our results suggest that, if anything, parents with more education or higher status occupations were marginally less likely to shift births. This could have gone either way. Well-off parents may be more informed about policy and better placed to negotiate with physicians. On the other hand, they may be less likely to take risks in delaying a birth for the sake of a relatively small amount of money.[23]

There are larger and more statistically significant differences in child characteristics. Those born on 1 July are more likely to be female, less likely to live in a metropolitan

[23]Given the introduction of the Baby Bonus was, as discussed earlier, not a uniform increase in cash transfers at birth in *net* terms, well-off parents also had less incentive in absolute terms to shift births.

area and less likely to have a language background other than English. Given the absence of an effect from the Baby Bonus' introduction on infant mortality (Gans and Leigh (2009)), the rise in the proportion of females may be due to sex differences in fetal health making physicians more willing to delay female births. Interestingly, those in regional areas seem to have been better able to shift births, possibly due to less hospital crowding. Perhaps less surprisingly, so too do those with an English-language background. All up, these results show a clearer set of family background differences between the birth cohorts born either side of the introduction of the Baby Bonus. A complex set of factors likely determines whether parents have the incentives, information, willingness and ability to shift their child's birth, so it is not surprising to see differences in some areas but not others.

The primary interest in birth shifting is the extent to which it influences, and potentially harms child health. Panel C shows clear effects of birth shifting on infant health. Those born immediately after the Baby Bonus had higher birthweights and were more likely to be of high birthweight, with an increase in expected birthweight of 30-60 grams and an increase in the expected probability of high birthweight of 1-2 percentage points. While there is no statistically significant increase in mean gestation length, this may reflect the fact that gestation lengths were rising prior to the introduction of the Baby Bonus. The estimates are nonetheless of the expected signs, consistent with longer gestations.

5.2.2 Combined treatment effects

There are clearly differences in the health settings, health outcomes and families of those born before and after the introduction of the Baby Bonus. We now test whether these differences have lasting repercussions. We do this by testing for a discontinuity in test scores either side of 1 July 2004. To complement this analysis, we investigate two other major episodes of birth shifting — the reintroduction of Maternity Allowance on 1 February 1996 and the increase to the Baby Bonus on 1 July 2006.

The results are foreshadowed in Figure 5 showing the size and later life outcomes of birth cohorts around each policy change. Although it resulted in the smallest share of births moved, the introduction of Maternity Allowance is associated with a relatively large discontinuity in the university offer rate — those born immediately after appear significantly less likely to be

Table 7: Parent and child characteristics

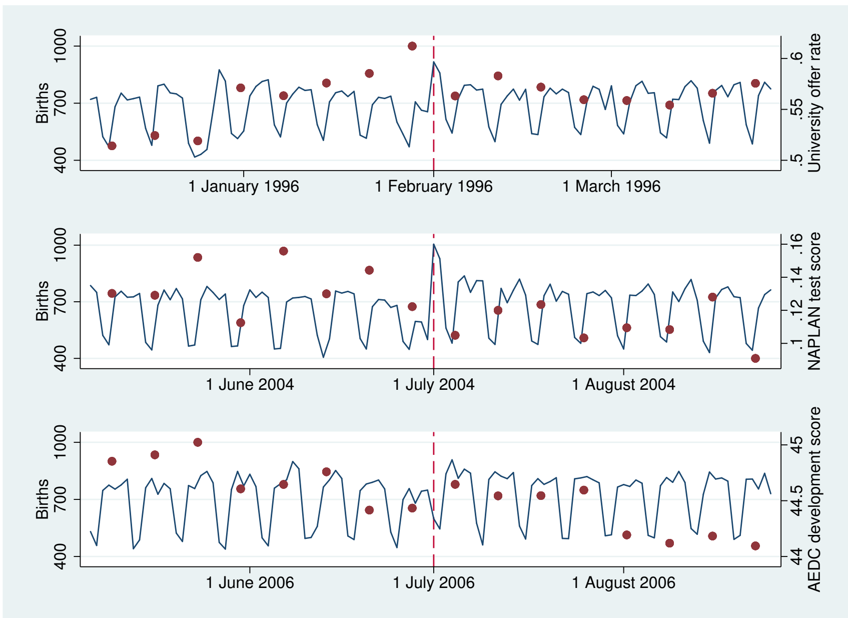
Window	± 7 days	± 14 days	± 21 days	± 28 days
	(1)	(2)	(3)	(4)
Panel A: Parent characteristics (highest education or occupation)				
Dropout	0.023	0.013	0.010	0.008
	(0.015)	(0.011)	(0.009)	(0.008)
Degree	-0.032	-0.026	-0.020	-0.023*
	(0.019)	(0.020)	(0.015)	(0.013)
Not employed	-0.012	0.003	-0.009	-0.005
	(0.015)	(0.011)	(0.009)	(0.007)
Senior manager	-0.021*	-0.006	-0.017	-0.012
	(0.012)	(0.019)	(0.015)	(0.012)
Panel B: Child characteristics				
Female	0.064***	0.047***	0.030***	0.031***
	(0.011)	(0.010)	(0.010)	(0.009)
Metropolitan	-0.038**	-0.030**	-0.033***	-0.021**
	(0.011)	(0.010)	(0.010)	(0.009)
LBOTE	-0.042***	-0.064***	-0.037***	-0.033***
	(0.014)	(0.011)	(0.009)	(0.009)
Panel C: Child cohort characteristics				
Mean birthweight (g)	62.5*	54.0**	30.4*	40.4**
	(26.0)	(20.1)	(16.8)	(15.1)
Low birthweight share	-0.003	0.004	-0.002	-0.016
	(0.009)	(0.007)	(0.006)	(0.010)
Normal birthweight share	0.005	-0.017	-0.016*	-0.008
	(0.021)	(0.012)	(0.009)	(0.007)
High birthweight share	0.011	0.019*	0.012*	0.010
	(0.020)	(0.010)	(0.007)	(0.007)
Mean gestation length (weeks)	0.025	0.031	0.044	0.044
	(0.065)	(0.056)	(0.042)	(0.041)
Pre-term share	-0.016	-0.009	-0.007	-0.006
	(0.018)	(0.012)	(0.008)	(0.007)
Term share	0.011	0.009	0.003	0.001
	(0.021)	(0.013)	(0.009)	(0.008)
Post-term share	0.005	0.002	0.003	0.004
	(0.004)	(0.003)	(0.002)	(0.002)

Robust standard errors in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

offered a place at university.[24] There is no apparent discontinuity in average NAPLAN test scores around the introduction of the Baby Bonus. And there is a potential discontinuity in average AEDC development scores around the increase to the Baby Bonus — those born immediately after appear to receive slightly higher development scores on average.

Figure 5: Birth cohort sizes (lines) and later life outcomes (dots) around three policy changes.



Note: means shown for the outcome variables in weekly bins either side of the policy change. In the first panel, 1 January birth dates are excluded from calculation of the mean due to their significant overrepresentation relative to registered births — this could be due to migrants with known birth years but unknown birth dates assuming 1 January of their relevant date of birth.

The intuitive feel for the results from Figure 5 is mostly born out by the regression discontinuity estimates in Figure 6. The introduction of Maternity Allowance is associated with a fall in the probability of receiving an offer of a university place by over 5 percentage points. The effect is significant at the 5 per cent level. There is no significant effect on average NAPLAN test scores associated with the introduction of the Baby Bonus. The effect on average AEDC development scores is not significant at the 5 per cent level.[25] It is worth noting that the significant discontinuity in university offers is entirely due to a discontinuity in university applications, rather than in the rate of offers per application, as apparent in

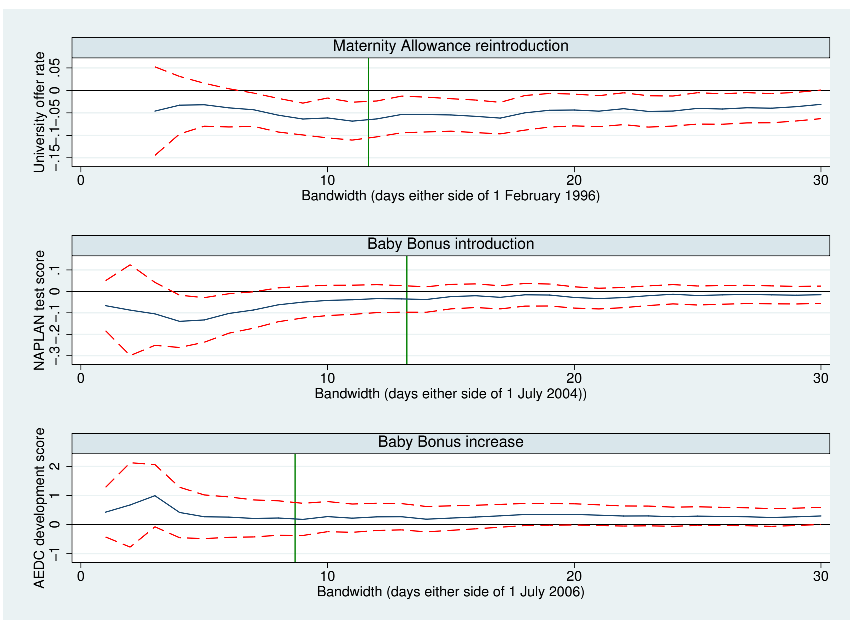
[24] The university offer rate is, more precisely, the number of university offers for those born on a given day as a proportion of the original size of that birth cohort. We are unable to adjust for migration.

[25] We also tested for a discontinuity in AEDC development scores separately across the five developmental domains, and again failed to find clear evidence of a discontinuity.

Appendix A, Figure A.6. The absence of a discontinuity in offers per application suggests no change in ability relative to aspirations at the time of applying to university.

These mixed results are not easy to interpret, particularly as they relate to different policy changes. Unfortunately, we do not have a common outcome measure that we can employ for all three policy changes. Also, it is not possible to disentangle the effects of differences in family background and birth shifting. One possibility is that any early leads in development associated with higher birthweights quickly dissipate such that by the time children enter adulthood, differences in family background arising from birth shifting may well dominate. For example, if well-off parents are less likely to shift births to qualify for Maternity Allowance, as was the case for the Baby Bonus, then the discontinuity in university offers for their young adult children may simply reflect the intergenerational transmission of ability and aspirations.

Figure 6: Combined treatment effects for three policy changes



Note: vertical line indicates rule of thumb bandwidth.

6 Conclusion

Cash transfers to families are a central part of many governments' responses to poverty and an important component of social welfare expenditure. These benefits have a variety of motives, including ameliorating current economic hardship, but also improving child outcomes. However, it is hard to assess the efficacy of cash transfers in improving child outcomes in the absence of large-scale experiments. This makes the identification and exploitation of natural experiments a crucial source of evidence.

In this paper, we use a difference-in-differences design to assess the effect of the additional family income provided by the Baby Bonus on child outcomes, here measured by a combined reading and numeracy test score in Year 3. We find no evidence that this additional family income improved outcomes in aggregate. There is some evidence for a modest effect on children from disadvantaged backgrounds — sufficient to lift the typical child around 1.5 percentile rankings — but this effect is not apparent across all specifications. This research complements existing literature finding relatively modest aggregate effects of family income on child outcomes, that are largest among children from disadvantaged backgrounds, as in Dahl and Lochner (2012), and when provided early in life, as in Morris et al. (2005).

There are a variety of explanations for these findings. The first is simply that the boost to family income was small — a one-off increase of around 1 to 2 weeks average earnings. Despite the large body of evidence on the importance of early life experiences to later outcomes (see Currie and Almond (2011)), in a wealthy country with high quality public services, such as Australia, the *average* returns to *additional* investments in early childhood may be relatively low. Alternatively, even if such investment opportunities exist, parents may be unaware, unwilling or unable to make them.

Despite the large birth shifting that accompanied the introduction, which triggered significant academic and media concern, there is no evidence those born immediately after the introduction of the Baby Bonus did any worse than those born immediately beforehand. This is consistent with the birth shifting resulting in changes in infant health that while highly statistically significant, were relatively small. Examining outcomes for children born either side of other major policy-induced birth shifting events provides a more mixed picture, with marginally better early development indicators but significantly lower rates of university offers for those born immediately afterwards, though these could come about due to differences in family background. More research on the long term consequences of birth shifting is warranted, but would remain challenging even with longitudinal unit record data given the complex set of influences and possible effects illustrated in this paper.

Conflict of Interest Statement

The authors declare that they have no conflict of interest.

References

Australian Bureau of Statistics (2014). *Births, Australia, 2013*, Australian Bureau of Statistics, Canberra. Cat. no. 3301.0.

Australian Curriculum, Assessment and Reporting Authority (2015). *National Assessment Program — Literacy and Numeracy 2014: Technical Report*, Sydney.

Australian Government (2013). *A Snapshot of Early Childhood Development in Australia 2012 - AEDI National Report*, Australian Government, Canberra.

Australian Institute of Health and Welfare (2014). Australia's mothers and babies 2012, *Perinatal statistical series no. 30. cat. no. PER 69*, Australian Institute of Health and Welfare, Canberra.

Baker, M. and Milligan, K. (2010). Evidence from maternity leave expansions of the impact of maternal care on early child development, *Journal of Human Resources* **45**(1): 1–32.

Carneiro, P., Løken, K. V. and Salvanes, K. G. (2015). A flying start? Maternity leave benefits and long-run outcomes of children, *Journal of Political Economy* **123**(2): 365–412.

Currie, J. and Almond, D. (2011). Human capital development before age five, *Handbook of Labor Economics* **4**: 1315–1486.

Dahl, G. B. and Lochner, L. (2012). The impact of family income on child achievement: Evidence from the earned income tax credit, *The American Economic Review* **102**(5): 1927–1956.

Daniels, D. (2009). Social security payments for people caring for children, 1912 to 2008: a chronology, *Parliament of Australia, Department of Parliamentary Services, Parliamentary Library, Background Note* .

Department of Education and Training (2016). *Higher Education Statistics*.
URL: *https://www.education.gov.au/higher-education-statistics*

Dustmann, C. and Schönberg, U. (2012). Expansions in maternity leave coverage and children's long-term outcomes, *American Economic Journal: Applied Economics* pp. 190–224.

Family Assistance Guide (2016).

URL: *http://guides.dss.gov.au/family-assistance-guide/3/6/4*

Gans, J. S. and Leigh, A. (2009). Born on the first of July: An (un) natural experiment in birth timing, *Journal of Public Economics* **93**(1): 246–263.

Gans, J. S. and Leigh, A. (2012). Bargaining over labour: Do patients have any power?, *Economic Record* **88**(281): 182–194.

Gans, J. S., Leigh, A. and Varganova, E. (2007). Minding the shop: The case of obstetrics conferences, *Social Science & Medicine* **65**(7): 1458–1465.

Hutchinson, D. (2016). *Weekly wages, average compensation and minimum wage for Australia from 1861-Present.*

URL: *https://www.measuringworth.com/auswages/*

Kewley, T. H. (1973). *Social security in Australia, 1900-72*, Sydney University Press.

Lee, D. S. and Lemieux, T. (2010). Regression discontinuity designs in economics, *Journal of Economic Literature* **48**: 281–355.

Liu, Q. and Skans, O. N. (2010). The duration of paid parental leave and children's scholastic performance, *The BE Journal of Economic Analysis & Policy* **10**(1).

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* pp. 175–205.

Morris, P., Duncan, G. J. and Clark-Kauffman, E. (2005). Child well-being in an era of welfare reform: the sensitivity of transitions in development to policy change., *Developmental Psychology* **41**(6): 919.

Rasmussen, A. W. (2010). Increasing the length of parents' birth-related leave: the effect on children's long-term educational outcomes, *Labour Economics* **17**(1): 91–100.

Rossin, M. (2011). The effects of maternity leave on children's birth and infant health outcomes in the United States, *Journal of Health Economics* **30**(2): 221–239.

Schulkind, L. and Shapiro, T. M. (2014). What a difference a day makes: quantifying the effects of birth timing manipulation on infant health, *Journal of Health Economics* **33**: 139–158.

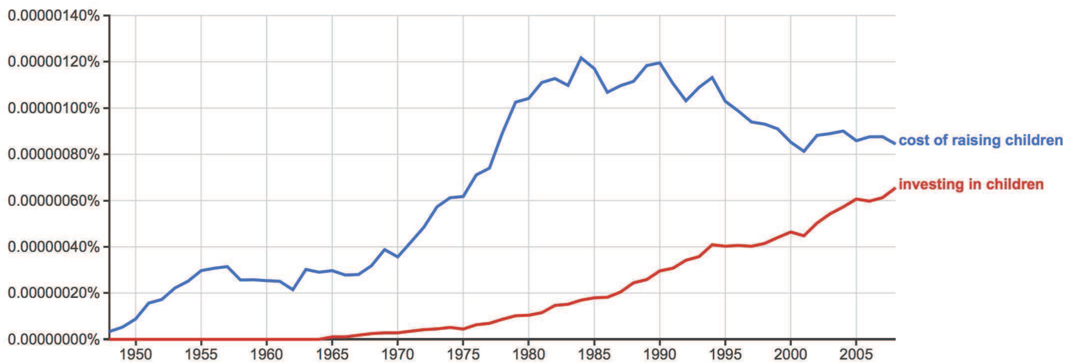
Tamm, M. (2013). The impact of a large parental leave benefit reform on the timing of birth around the day of implementation, *Oxford Bulletin of Economics and Statistics* **75**(4): 585–601.

Thorngren-Jerneck, K. and Herbst, A. (2001). Low 5-minute Apgar score: A population-based register study of 1 million term births, *Obstetrics & Gynecology* **98**(1): 65–70.

APPENDIX

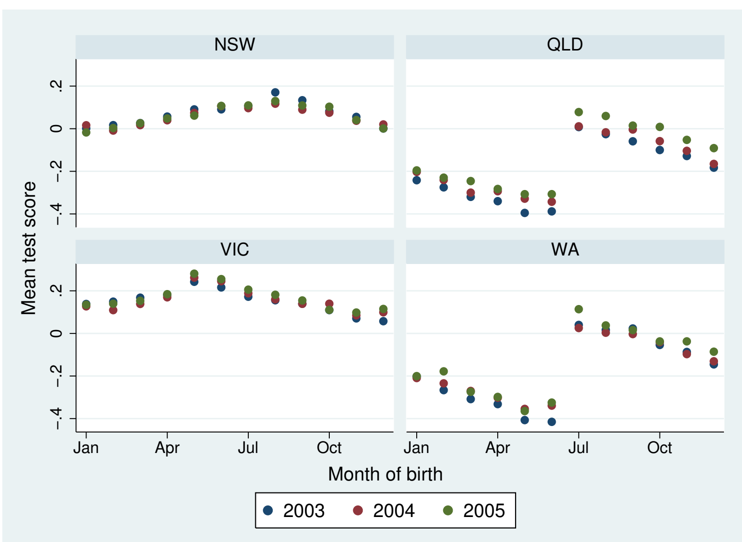
A Additional charts

Figure A.1: Are children an expense or investment? Written-word prevalence of potential rationale for family benefits.



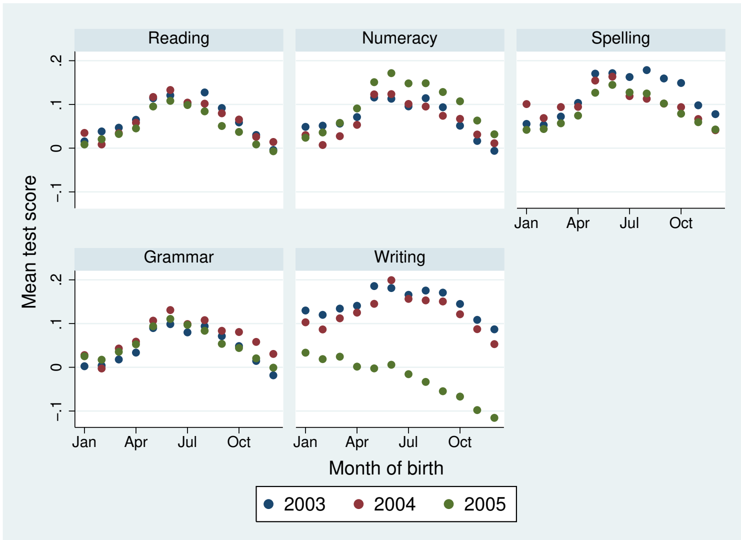
Source: Google ngram viewer.

Figure A.2: Binned mean test score by month of birth: four largest states



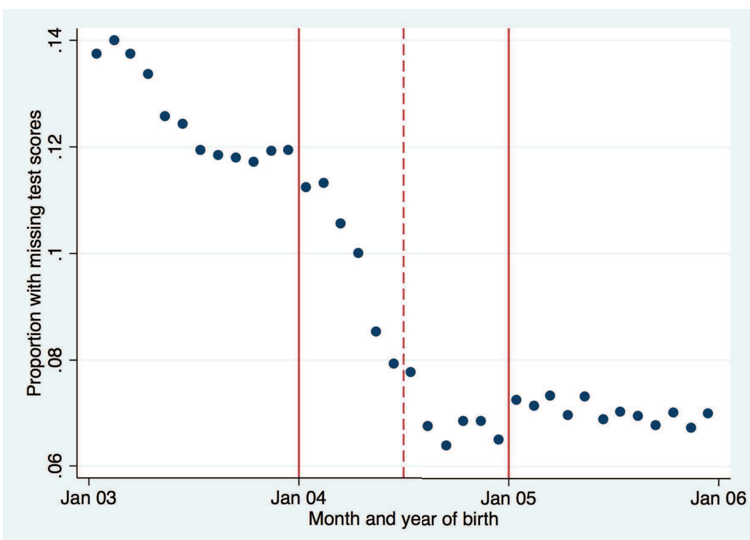
Note: Outcome variable is the summed reading and numeracy score, normalised across all years and states and territories to have mean zero and standard deviation one.

Figure A.3: Binned mean test score by month of birth: test components



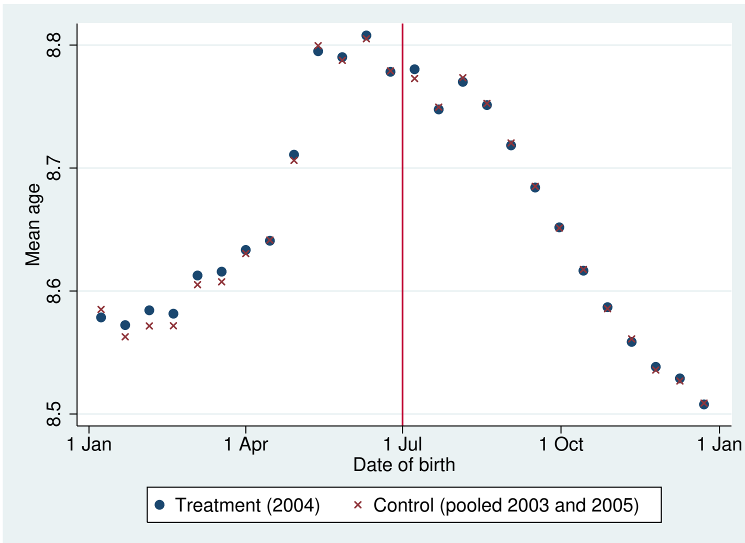
Note: Excludes those states and territories with 1 July cut off dates.

Figure A.4: Missing values by year of birth



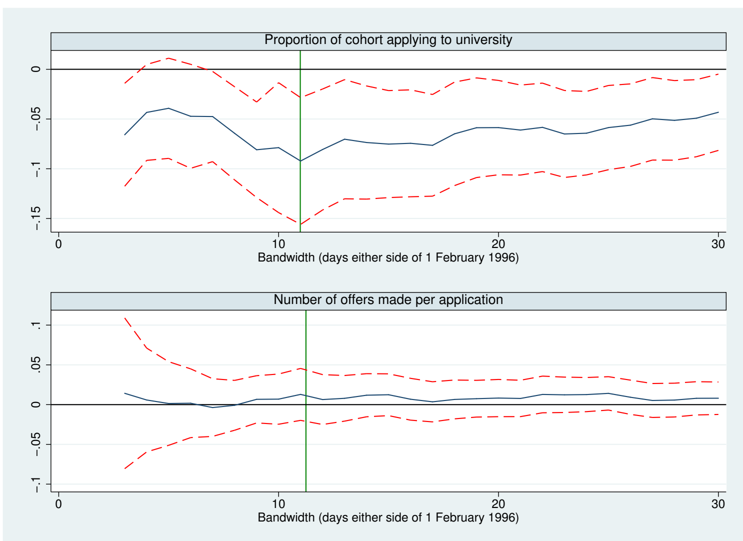
Note: excludes those states and territories with 1 July cut off dates. The higher rate of missing values for earlier months likely reflects the increasing share in earlier birth cohorts taking the test in 2011 or 2012, where we infer the Year 3 test scores from the past test scores of Year 5 students two years later. For example, for 2013 we have both direct data on Year 3 test scores and indirect data from the past test scores of Year 5 students in 2015. The rate of missing values in the direct data is 7.6% compared to 13.6% in the indirect data. The higher rate of missing values in the indirect data could stem from: a failure to link an individual to their past test scores (so past scores are recorded missing, despite actually existing); and migration resulting in the Year 5 population being larger than the actual eligible population of Year 3 students two years prior. In either case, the arguments in the text hold that the trends in missing values will if anything likely bias our estimates downwards. As a further check, we ran regression (3) from Table 6 restricting attention to those sitting the tests in the year in which they turned 9 or 10. This ensures only direct test data is used. The estimates did not change in a substantive way (results available on request).

Figure A.5: Mean age-at-test by date of birth



Note: Excludes those states and territories with 1 July cut off dates.

Figure A.6: Decomposition of effect of Maternity Allowance reintroduction



Note: vertical line indicates rule of thumb bandwidth.

B Additional tables

Table B.1: Estimated treatment effects of placebo policy on chosen test scores by population subgroups.

	All	Dropout	High school or less	Degree	Not employed	Senior manager
	(1)	(2)	(3)	(4)	(5)	(6)
Reading	-0.008	-0.039	-0.014	-0.002	-0.030	0.022
	(0.010)	(0.030)	(0.022)	(0.016)	(0.034)	(0.019)
Numeracy	-0.000	-0.008	-0.006	0.004	-0.038	0.018
	(0.010)	(0.030)	(0.023)	(0.016)	(0.034)	(0.020)
Combined	-0.005	-0.026	-0.012	0.001	-0.033	0.022
	(0.010)	(0.030)	(0.022)	(0.016)	(0.034)	(0.019)

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

C First Child Tax Rebate

The First Child Tax Refund was a refundable tax offset that was designed to pay back to mothers the tax paid on their income (up to a limit) in the year prior to having their first child. (In Australia, the unit of taxation is the individual.) The primary carer of a first child born between 1 July 2001 and 30 June 2004 could claim the refund at the end of each of the first five financial years following the birth. The base annual refund was equal to a fifth of the income tax paid in the financial year prior to the birth of the child, bounded by \$500 below and \$2,500 above. This was then scaled down in proportion to the time the child was in their care and the degree to which the primary carer's taxable income I_t remained below that in the year before the child was born I_0. Thus, the refund in year t was:

$$\text{First Child Tax Refund}_t = (\text{base annual refund}) \cdot (1 - I_t/I_0).(\text{days in claimant's care}/365) \tag{2}$$

As a result, families could receive anywhere between nothing and \$12,500 through the refund. In practice, there were 736,000 births registered in Australia in the eligibility window for the refund and an average refund paid of \$1,793 for each birth. As the refund only applied to the first birth after 1 July 2001, average refunds fell over time. In the month prior to the introduction of the Baby Bonus the mean real refund paid over the following five years was \$1,280 while the median was only \$485 (Table 1).

As refunds differed substantially according to individual circumstances, it might be expected that the average refund would increase prior to its abolition as individuals sorted across the threshold date to qualify for the Baby Bonus or the First Child Tax Refund according to whichever was most financially attractive. In fact, there is only a very modest increase in the average refund in the days immediately prior to 1 July 2004. This could be due to the refund being: less salient (and harder to calculate) than the bonus; and heavy discounting due to its delayed payment and uncertainty regarding the primary carer's return to work.