I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13543

Unintended Consequences of Welfare Reform: Evidence from Birth Outcomes of Aboriginal Australians

Mary-Alice Doyle
Stefanie Schurer
Sven Silburn

JULY 2020

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13543

Unintended Consequences of Welfare Reform: Evidence from Birth Outcomes of Aboriginal Australians

Mary-Alice Doyle
London School of Economics

Stefanie Schurer
University of Sydney and IZA

Sven Silburn
Charles Darwin University

JULY 2020

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Unintended Consequences of Welfare Reform: Evidence from Birth Outcomes of Aboriginal Australians*

Australia's 'income management' policy requires benefit recipients to spend at least half of their government transfers on essentials (e.g. food, housing). We estimate income management's impact on birth outcomes by exploiting its staggered rollout. By changing parents' consumption patterns, the policy aims to improve child outcomes. We find no evidence of this. Instead, our estimates suggest it reduced average birthweight by 95 grams and increased the probability of low birthweight by 3 percentage points. We explore the mechanisms that may explain this finding. Our study demonstrates how policies that are not carefully implemented and tested can unintentionally escalate existing inequalities.

JEL Classification: D04, I14, I38

Keywords: welfare reform, Aboriginal children, birth outcomes, income management, unintended consequences

Corresponding author:
Stefanie Schurer
School of Economics
University of Sydney
Sydney, NSW, 2006
Australia

E-mail: stefanie.schurer@sydney.edu.au

* The authors thank the following people for their valuable feedback on elements of this research program: Heather d'Antoine, Victoria Baranov, Nicholas Biddle, Gawaian Bodkin-Andrews, Robert Breunig, Julie Brimblecombe, Janet Currie, David Cooper, Steven Durlauf, Denzil Fiebig, Marco Francesconi, Matthew Gray, Olga Havnen, James J. Heckman, Matthew James, Liz Moore, Patrick Nolan, Dilhan Perera and Jim Smith; participants of seminars at the University of Chicago, University of Essex, University of Melbourne, Australian National University, University of Western Australia, J-PAL, Institute for Fiscal Studies; and participants of Inaugural Conference of the Asian and Australasian Society of Labour Economists (Canberra, Dec 2017), the European Society of Population Economics (Bath, June 2019), the European Society of Labour Economics (Uppsala, September 2019), the NBER workshop on Health, Wellbeing, and Children's Outcomes for Native Americans and other Indigenous Peoples (Boston, November 2019), and of an organised session at the Southern Economics Conference (Washington, DC, Nov 2018) and of the 9th Australasian Workshop on Econometrics and Health Economics. This study uses data from the Northern Territory (NT) Early Childhood Data Linkage Project, `"Improving the developmental outcomes of NT children: A data linkage study to inform policy and practice across the health, education and family services sectors", which is funded through a Partnership Project between the National Health and Medical Research Council (NHMRC) and the NT Government. This study uses administrative data obtained from the NT Department of Health through this NHMRC Partnership Project. The analysis has followed the NHMRC Values and Ethics: Guidelines for Ethical Conduct in Aboriginal and Torres Strait Islander Health Research (2003) and the Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS) Guidelines for Ethical Research in Australian Indigenous Studies (2012) (Reciprocity, Respect, Equality, Responsibility, Survival and Protection, Spirit and Integrity). The researchers are bound by, and the research analysis complies with, the ethical standards outlined in the ethics agreement HREC Reference Number: 2016-2611 Project Title: Improving the developmental outcomes of Northern Territory children: A data linkage study to inform policy and practice in health, family services and education (Human Research Ethics Committee of the Northern Territory Department of Health and Menzies School of Health Research). The authors acknowledge funding from an Australian Research Council (ARC) Discovery Early Career Research Award DE140100463 and a University of Sydney SOAR Fellowship (2017-2018).

In 2007, the Australian Government embarked on its most contentious welfare reform in recent history. The new policy, referred to as 'income management', restricted the way Aboriginal benefit recipients in the Northern Territory[1] could spend their entitlements. Most benefit recipients in Australia receive a cash transfer. Under income management, people receiving government benefits – including unemployment, disability support, parenting payments and the pension – had half of those payments quarantined into a separate account, where it was designated to be spent only on priority needs and could not be withdrawn as cash.

The policy's official objective was to improve the welfare of Aboriginal children. It was hoped that by restricting spending choices, households would spend more money on child-centred goods, and less on goods viewed as harmful, such as alcohol, tobacco and gambling services.[2] Many countries use paternalistic approaches to welfare policy (Currie and Gahvari 2008) by providing restricted-used transfers – e.g. in-kind transfers or conditional cash transfers.[3] But Australia is unique in its application because income management affected all benefit recipients within a community and quarantined large shares of their income.[4] What is also unique about Australia's approach is that this policy was introduced with the goal of addressing persistent disadvantage in Indigenous communities. The Australian Prime Minister at the time emphasized "the parental responsibility that

[1] In this paper we refer to people of Aboriginal or Torres Strait Islander descent as `Aboriginal' because most Indigenous individuals in the Northern Territory identify as Aboriginal singularly or as both Torres Strait Islander and Aboriginal.

[2] The *Welfare Payment Reform Act 2007* stated explicitly that it aims to "promote socially responsible behavior, particularly in relation to the care and education of children" (*Welfare Payment Reform Act 2007* No. 130, 2007 123TB Objects, Section (a)). The government minister responsible for the policy change stated the aim was to "stem the flow of cash going towards substance abuse and gambling and ensure that funds meant to be for children's welfare are used for that purpose" (Brough 2007).

[3] We use the term 'restricted transfers' to incorporate in-kind transfers, as well as cash transfers that come with restrictions or conditions over how recipients should use their funds.

[4] While we do not have data on the value of household income and welfare payments, AIHW (2010) reports that around one-third of adults in the affected communities were employed, suggesting that the remaining two-thirds would be eligible to receive government transfer payments. In contrast, some other countries use similar policies, but they apply to a specific subset of the population (e.g. teenagers receiving welfare income in New Zealand, or asylum seekers in the UK), or to a smaller share of household income (e.g. SNAP in the US).

accompanies [parents'] right to welfare support" (Howard 2007, 73), reflecting the policy's aim to reduce entrenched disadvantage by changing individual behavior.

In this paper, we estimate the causal impact of the introduction of income management on birth outcomes. We focus specifically on birthweight and the probability of low birthweight. For the majority of Aboriginal women living in remote areas, government transfers are their main source of income (Venn, Biddle, and Sanders 2020). Our hypothesis is that if income management increased consumption of essentials, this would be reflected in improved nutrition of pregnant women, and therefore in increased birthweight – patterns that have been found, for example, in the US with the introduction of food stamps (Almond, Hoynes, and Schanzenbach 2011). Unlike other studies, we are able to estimate the effect of introducing spending restrictions as the value of entitlements remained unchanged.

Our analysis draws on administrative records on the universe of all births from the Northern Territory Data Linkage Study (Silburn et al. 2018). To identify the impact of the policy, we use a difference-in-difference model, exploiting the gradual policy rollout. We demonstrate that the rollout schedule of the policy was as good as random. The policy was implemented shortly after its announcement, was compulsory, and allowed almost no exemptions. This means there was no capacity to self-select into whether or when to receive the intervention. Importantly, the rollout was not linked to variations in birthweight or community characteristics.

We find that income management did not improve birth outcomes for Aboriginal children in the Northern Territory. Instead, the policy change reduced average birthweights of children who were exposed to income management in utero by 95 grams. The adverse effects are strongest at the lower end of the birthweight distribution. Income management increased the probability of low birthweight by 3 percentage points – a 20 percent increase from the pre-treatment period.

We explore the likely mechanisms for this unexpected finding. We find the negative treatment effect cannot be explained by changes in fertility, maternal risky health behaviours, or access to perinatal care. There is also no evidence that income management improved chances of survival for at-risk foetuses.

This leaves us with a more qualitative assessment of why income management worsened birth outcomes. We find evidence that the policy's implementation rules may have caused short-term income insecurity, having temporarily suspended benefit payments of up to one third of all recipients. This interpretation is consistent with tentative evidence that income management may have increased financial stress and reduced family functioning (Cobb-Clark et al. 2018).

A large international literature has studied the negative impacts of income, health behaviors, stress and environmental shocks in utero on birth outcomes (see Aizer and Currie (2014) for an overview), and the long-term consequences of those shocks to early life health (Almond, Currie, and Duque 2018). There is also strong evidence that restricted-use transfers can prevent or reduce the effects of those shocks (Currie and Rossin-Slater 2015). But little evidence exists on the unintended consequences that restricted-use transfer programs may have on early life health. Our findings are of critical significance to policymakers because they demonstrate that policies that are not carefully designed, implemented and tested may unintentionally escalate the inequities they seek to address.

I. Policy background

A. The Northern Territory Emergency Response (NTER)

The Northern Territory (NT) is a large and sparsely populated geographic area, covering approximately one-sixth of the Australian continent. Although nation-wide Aboriginal Australians make up only 3 percent of the population, around one-quarter of the people living in the NT are Aboriginal. Most Aboriginal people in the NT live in remote towns or communities. Aboriginal Australians – especially those who live in remote communities – experience substantial health disparities relative to non-Aboriginal Australians. In the NT, life expectancy at birth is 10 years lower for Aboriginal babies than for non-Aboriginal babies (Commonwealth of Australia 2020). One contributor to this disparity may be poor birth outcomes, given the association between birthweight and life expectancy (Risnes et al. 2011).

The Australian Human Rights Commission has highlighted the role of structural factors, including limited access to health services and infrastructure (housing, sanitation and food supplies), in explaining these disparities (Aboriginal and Torres Strait Islander Social Justice Commissioner 2005). However, as Dawson et al. (2020) observe, while Australian policymakers have often acknowledged these structural factors, the policy solutions they have offered tend to rely on individual behavior change.

This was exemplified in mid-2007 when the Australian Government announced the Northern Territory Emergency Response (NTER), a wide-ranging package of policies aimed at protecting the children's health and safety of children in remote Aboriginal communities.[5] The central policy was income management, though the NTER included a range of other policies such as alcohol and pornography bans, additional police presence, child health checks and housing and land reform (see Appendix A for full list).

These policies applied to residents in all 73 remote Aboriginal communities and their outstations, and in 10 town camps.[6] They did not apply to non-Aboriginal towns or communities in the NT. To facilitate the racially targeted nature of these policies, the government suspended Part II of the *Racial Discrimination Act 1975*, which proscribes equality before the law regardless of race.[7]

[5] The policies were enacted following the publication of a report by the NT Board of Inquiry into the Protection of Aboriginal Children from Sexual Abuse (Wild and Anderson 2007). The report called for immediate government action to address child sexual abuse in remote communities. It emphasized the need to consider child neglect, alcoholism and inadequate education and housing as long-term contributors to abuse. The policies did not focus directly on child protection, but instead on changes intended to 'normalize' remote Aboriginal communities.

[6] A town camp is an Aboriginal community situated in a town or city, or close to its boundaries.

[7] This meant that members of the communities affected by the NTER legislation were denied the ability to challenge legislation on the basis that it discriminates by race (Australian Human Rights Commission 2011).

B. The Income Management policy

Income management imposed restrictions on what benefit recipients could do with their payments. Before income management, recipients had 100 percent of their payments deposited into their bank account. Under income management, the total value of payments did not change, but half of each regular payment was set aside into a separate income management account. It could only be spent on priority needs such as food, housing, bills and clothing. It could not be withdrawn as cash. The remaining half was paid into recipients' bank accounts as usual.

Recipients were required to meet with a case officer to create a spending plan for their quarantined funds. Recipients could choose to have part of their income-managed funds paid directly to suppliers to cover bills, rent or debt repayments. They could have some of their funds credited in their name to a local store to purchase food and household goods. Unspent funds could accumulate as savings. Any changes to these allocations were made in consultation with a case officer.

Towards the end of the rollout period (8 September 2008), a debit card (the 'Basics Card') was introduced, which allowed participants to load their quarantined funds onto the card and use it to purchase items at any participating store.

Income management applied to all benefit recipients in the NT who lived in remote Aboriginal communities and town camps. While detailed data on welfare payment rates are unavailable, it is safe to say that income management affected most residents in remote communities. Data reported in a key governmental report (Australian Institute of Health and Welfare (AIHW) 2010) indicate that around three-quarters of the adults in affected communities were subject to income management at some point during the rollout period, with 55 percent being income managed at a point in time after the rollout was complete.[8] Women were more likely to be income managed than men (60 versus 40 percent of participants).

The number of people affected by income management was probably greater than the number of adults receiving benefits. Given the large average household sizes in affected communities – 6 people per household (see Table 1) – many residents who were not themselves recipients were likely living with somebody who was.

Benefit recipients could not avoid income management; exemptions were available, but very rare. Overall, only for 649 out of 21,763 clients who were ever income managed (3 percent), an exemption was granted (AIHW 2010). The main reason was for clients who moved away permanently from (three in five) or who had little connection to the community.

The administrative costs of income management are high. Recent estimates suggest that the policy costs around A$9,000 per participant per year, excluding the value of the benefit payment (Department of Social Services 2017).

C. Empirical evidence on income management

There are two studies which consider the effect of the introduction of income management in the Northern Territory on spending patterns, with mixed findings. Brimblecombe et al. (2010) find no evidence of a change in spending on food in a small number of remote communities after income management was introduced. Conversely, Lamb and Young (2011) find suggestive evidence of a decrease in gambling expenditure. But both studies relate to specific locations, so their findings may not be representative of the effect of the policy across the whole of the NT.[9]

[8] The lower share of residents affected at the end of the period reflects residents moving onto and off income support payments, for example, due to changes in employment status or eligibility.

[9] Brimblecombe et al. (2010) use data from 10 community stores. Before the rollout, those stores already provided a voluntary 'Food Card' system to residents, which restricted purchases to nutritious items. The 'Food Card' program was in use before income management was introduced and was subsequently provided to welfare recipients as an optional way of accessing income managed funds before the Basics Card was rolled out. The authors also note most of the ten communities had pre-existing alcohol bans. Lamb and Young (2011) use data from a single expenditure at one venue in each of two towns, and they caution that a decrease in formal gambling expenditure may be offset by informal gambling.

Apart from these two studies, government departments have written or commissioned evaluation reports on income management. Two reports cover the initial introduction of income management in the NT, which is the focus of this paper (AIHW 2010; Department of Families, Hosing, Community Services and Indigenous Affairs (FaHCSIA 2011)). The AIHW report concludes that there is consistent evidence that the policy led to increased spending on essentials, and FaHCSIA reports that while income management was perceived negatively at first, it is "now seen as beneficial by many people, especially women" (p. 11).

Although informative, the conclusions of the two reports rely mainly on surveys or focus groups with community residents and staff involved in administering the program. These surveys were conducted with small, non-random samples so are unlikely to be representative of the treated population. No baseline data were collected before income management was rolled out, so benefits found in the evaluation reports relate to the perceptions of a selected group of respondents.

We build on the existing evidence, offering a causal analysis of the mechanics of the policy. Our unique dataset allows us to estimate the average policy impact across all affected communities and to focus on the policy's impact on children's health outcomes, a key policy objective.

D. Linking income management with birthweight: nutrition

Mechanically, there are two potential reasons for low birthweight: the baby may be born prematurely (short gestational length), or the baby may be born full term, but smaller than expected (intrauterine growth restriction). The determinants of each are different and complex, although maternal nutrition is most likely to have its impact through intrauterine growth (Kramer 1987).

For pregnant women, there is strong evidence from the economic literature that increased food consumption can increase birthweight through the intrauterine growth channel, and further, that this can be achieved through transfer programs (Barber and Gertler 2008; Almond, Hoynes, and Schanzenbach 2011; Hoynes, Page, and Stevens 2011). Outside of the economic literature, studies of nutrition-focused interventions have yielded similar conclusions. A meta-analysis by Gresham et al (2014) finds that, on average, randomized trials that provide food or fortified food during pregnancy increase birthweight by 125 grams.

Based on this evidence, we expect that if income management affects food consumption, this should be evident through higher birthweight and reduced incidence of low birthweight. While food is just one of the 'essentials' that income managed funds are intended to be spent on, we know that most (65 percent) income managed funds were spent on food during the rollout period (AIHW 2010).

II. Conceptual framework

Two conditions are required for income management to have its intended positive effects. First, the restriction on spending choices must make recipients consume more essentials. Second, this change in consumption must affect health outcomes.

A. Economic theory: extramarginal households

If it is to have an impact, income management must affect household consumption differently from an equivalent cash transfer. A simple model first developed by Southworth (1945) describes the potential for such change.

Southworth defines 'inframarginal' and 'extramarginal' households, as illustrated in Figure 1. Consumers tradeoff between consumption bundles involving 'priority goods' (e.g. food, X-axis) and all other goods (e.g. alcohol, Y-axis). Income management introduces a kink in the budget constraint (at IM in Figure 1), requiring that a maximum of 50 percent of benefit income can be spent on 'other goods.' There are no restrictions on the consumption of priority goods.

In the context of income management, inframarginal households are those who already spend more than half of their benefit income on essentials, and therefore the spending restriction does not require any change in behavior. It is only if the household initially spends less than half of the value of their benefit income on priority goods that the restriction will affect consumption; these would be 'extramarginal' households. Within this framework, we would expect income management to improve birth outcomes through increased food consumption, but only if a non-negligible share of households are 'extramarginal'.

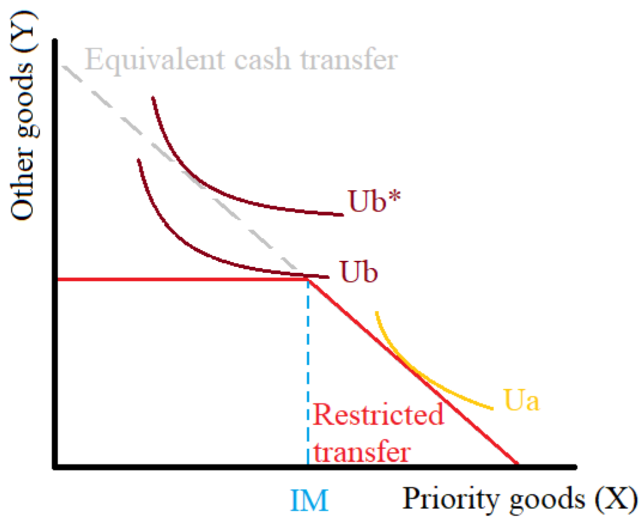

FIGURE 1: BUDGET CONSTRAINT WITH AND WITHOUT INCOME MANAGEMENT

Notes: Figure adapted from Southworth (1945). For extramarginal households, indifference curve moves from Ub* to Ub with the introduction of the requirement to spend minimum value 'IM' on priority goods. Ua indicates indifference curve for inframarginal households, whose spending patterns are unchanged.

B. Are households extramarginal?

Community-level spending data are not available, but aggregate pre-rollout data from the Australian Bureau of Statistics provide a proxy. They show that in the NT, low-income households and households that rely on government transfers as their main source of income spent most of their money on essentials (Figure 2).[10]

However, these data do not tell us the distribution of expenditures across households and are not available for residents in very remote communities. Even if the average household is inframarginal, a significant portion of households may have been extramarginal, meaning that we may still observe an impact of the policy on average outcomes.[11]

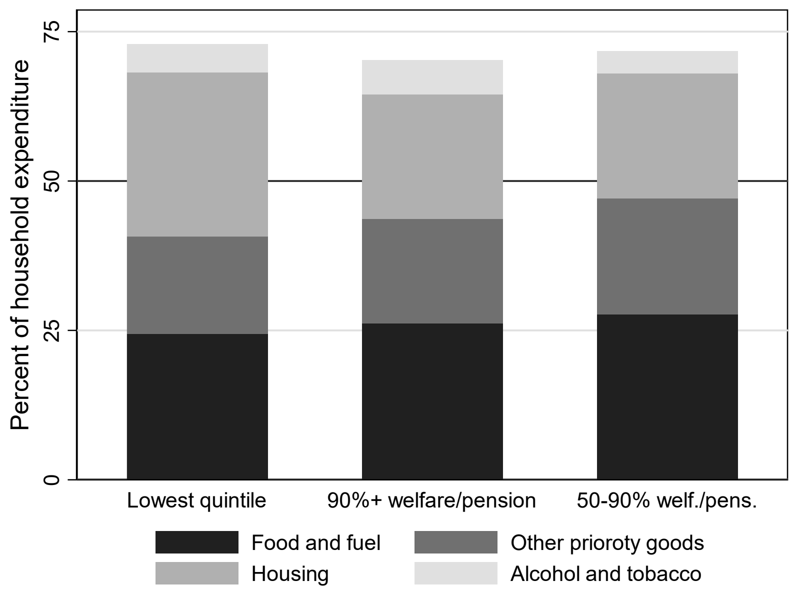

FIGURE 2. HOUSEHOLD EXPENDITURES IN THE NORTHERN TERRITORY

Notes: The figure shows spending on food, fuel, housing and 'other priority goods' (clothing and footwear, household furnishing, medical care expenses, transport), and spending on alcohol and tobacco, as a share of total household expenditures. The data do not relate specifically to residents of remote or Aboriginal communities. Lowest quintile refers to households in the lowest 20 per cent of the distribution of household income in the Northern Territory; 90%+ (50-90%) welfare pension refers to households where 90%+ (50-90%) of household income comes from government welfare payments or pensions.
Source: Australian Bureau of Statistics *Household Expenditure Survey* 2003-04.

[10] Total household income is likely to be higher than the value of welfare income. But the savings rate unlikely to affect this conclusion as recipients had the option of saving in income managed funds or in cash.

[11] The ABS data may over-estimate housing expenditure for our population. Many residents in remote NTER communities had low or no housing costs.

C. *Inputs and outcomes*

The theoretical framework explains the potential for income management to affect household consumption. But changes in household consumption are only relevant to policymakers if they translate to changes in outcomes such as improved health, education and wellbeing. As Cunha (2014) shows in analysing an in-kind transfer program in Mexico, this is not always the case. When recipients are limited to consuming certain items, they may consume these items instead of close substitutes, with no resulting change in health outcomes.

A similar dynamic may be relevant in our context. Critics of income management have argued that requirement to shop at specific food stores[12] may reduce purchases through less formal channels. If these less formal channels are lower cost, income management may actually reduce households' purchasing power and consumption (Australian Institute of Health and Welfare 2010). There is also some evidence that 'extramarginal' recipients found ways to circumvented the policy, e.g. trading items purchased with income managed funds for cash (Marston et al. 2020).

It is therefore possible that even if a household is extramarginal and income management brings about a change in purchasing patterns, the change may not flow through to the outcomes that policymakers care about.

III. International literature

There is an abundance of literature on the effect of restricted transfer programs – including the US SNAP program and conditional cash transfers in many developing countries. However, only a small number of studies directly compare restricted transfers to unrestricted cash transfers.

[12] For local stores to accept income management funds, they had to receive a license, indicating that they met certain minimum standards (in terms of their stock of fresh and nutritious foods, for example). In many cases, licenses were provided provisionally at the time of the introduction of income management, with any changes required to meet the minimum standards occurring after the rollout (see Appendix A).

In the US, researchers have reached no clear consensus on the impact of food stamps/SNAP relative to cash transfers. Several studies conclude that food stamp income is treated very similarly to cash income (Hoynes and Schanzenbach 2009; Cuffey, Beatty, and Harnack 2016). But others have identified a "cash-out puzzle", finding that food stamps increase food consumption, even though the vast majority of recipient households were inframarginal – that is, they spend more on food than the value of their food stamps. Some studies suggest mental accounting (Hastings and Shapiro 2018), or changes in intrahousehold decision making (Breunig and Dasgupta 2005) may be driving this finding.

Evidence from low- and middle-income countries also shows no clear evidence that either transfer type is more effective. Gentilini (2016) surveys ten studies that use either randomized controlled trials or natural experiments to compare transfers of cash to transfers of food. In some cases, food was given directly to participants, while in others, participants were given food vouchers. Most studies found no significant difference between the impact of the cash transfer and the restricted transfer. Both cash and restricted transfers increased food consumption and dietary diversity and reduced the incidence of malnutrition.

In summary, the literature does not point to any consistent findings on how restricted transfers affect spending and health outcomes. This may be because the existing studies have a common challenge: they provide only a low-powered test of the impact of restricted transfers, because they relate to programs worth a relatively small share of household expenditure (2.5-30 percent among those studied by Gentilini). It is therefore likely that these programs are inframarginal for most recipients. Our analysis overcomes this challenge because we study a restricted transfer program that affects a larger share of household income.

IV. Data and definitions

Our analysis is conducted with data from the NT Data Linkage Study (*NT-DLS*), which is funded through a Partnership Project between the Australian National Health and Medical Research Council (NHMRC) and the NT Government (Silburn et al. 2018). We extract from the *NT-DLS* the Perinatal Trends files (custodian: NT Chief Health Officer), which include demographic variables, information on maternal health, and birth outcomes. These files contain information on 74,425 children who were born in the Northern Territory between 1994 and 2013. For this study, we have linked in rainfall data from the Australian Bureau of Meteorology and Census community profile data from the Australian Bureau of Statistics.

A. Definition of treatment

Income management was rolled out in stages in all 73 remote Aboriginal communities (and associated outstations) in the NT, and 10 town camps. We refer to these collectively as NTER communities. In our data, NTER communities are separated into 88 locations, 85 of which had at least one birth in the rollout period.[13]

To identify newborns in NTER communities, we use information on the mother's suburb of residence at the time of birth, as recorded in the Perinatal Trends files.[14] We identify suburbs that are located in NTER communities, and link these observations to the date income management was introduced in that community (the schedule is available in Appendix A of the AIHW (2010)).

We define a child as being treated if income management was introduced in her community before the start of the third trimester in utero – that is, up to 28 weeks

[13] The number of communities in the dataset is greater than the number of NTER communities because some outstations were treated on a different timeline to their closest large community and are therefore listed separately.

[14] We use a range of sources, including www.bushtel.nt.gov.au and *Social Security (Administration) (Declared Relevant Northern Territory Areas – Various) Determination 2010 No. 8* to identify aliases, outstations and alternative spellings for each community, to map the mother's self-reported place of residence to the correct NTER community.

after the estimated date of conception. This definition is based on Almond, Hoynes, and Schanzenbach (2011) who find that the introduction of SNAP in the US significantly increased birthweight if it was in place for the full third trimester, but with no additional impacts if it was introduced earlier.

B. Sample selection

In our analysis, we use the subset of births to mothers who resided in an NTER community. In estimating the treatment effect, we limit our sample to babies who were in their third trimester during the rollout period. This gives a total sample of 1,083 babies conceived 1 January 2007-30 April 2008.[15] We choose this narrow sample period around the dates of the rollout to reduce the potential for confounding time trends and policy changes – including other NTER policies. However, to improve our estimation of other coefficients in our model, we include one year of pre-rollout period data (conceived 1 January 2006 – 31 December 2006). We exclude 39 babies born before the beginning of their third trimester (14 during the rollout period, of which eight were stillbirths).

C. Outcome variables

The outcome variables of interest are birthweight and the probability of low birthweight, which we derive from the NT Perinatal Trend files. Low birthweight is common in NTER communities – with around 14 percent of infants born with low birthweight in the year before income management was introduced – compared with 7 percent in other parts of the NT.

We focus on birthweight for two reasons. First, improved maternal nutrition during pregnancy can increase birthweight (Gresham et al. 2014). This means that

[15] We define our sample based on date of conception instead of date of birth to avoid the 'fixed cohort bias' (Strand, Barnett, and Tong 2011)

if income management was successful in increasing food consumption, we would expect an increase in birthweight. Second, birthweight is an important outcome measure in its own right, given its causal influence on child mortality, cognitive development, educational attainment and labor market outcomes (see Almond, Currie and Duque (2018) for an overview).

V. Empirical framework

A. Identification strategy

To identify the causal impact of income management, we exploit its staggered rollout. As shown in Figure 3, the policy was introduced over a period of 13 months. Provided that the rollout timing is exogenous, we can use it to estimate the causal effect of the introduction using a generalized difference-in-difference approach. In the remainder of this section, we present evidence that the rollout schedule was as good as random, and that before the rollout, there were parallel trends in – and levels of – birth outcomes between the earlier- and later-adopting communities.

First, income management was rolled out on a pre-defined timeline.[16] As shown in Appendix A, that timeline was different from the rollout of other policies that were part of the Northern Territory Emergency Response, meaning that our results are not confounded by concurrent policy changes.

Second, the rollout did not follow any clear geographic pattern (Figure 4). Income management was rolled out in parallel in two 'clusters' (north and south), but with no apparent pattern as to whether very remote communities, larger communities or town camps received treatment first within each cluster.

[16] The implementation was delayed for 13 communities, outstations and town camps. Delays were substantial (that is, more than a few weeks) for only four communities (Australian Institute of Health and Welfare 2010). Our identification is based on the actual, not planned, rollout dates.

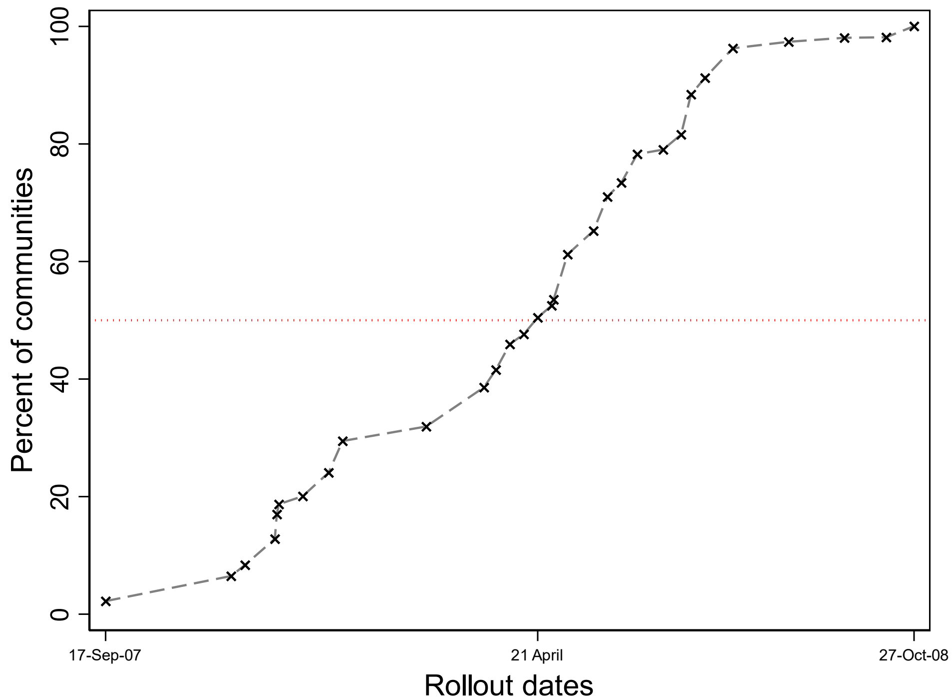

FIGURE 3. ROLLOUT OF THE INCOME MANAGEMENT POLICY

Note: The graph shows the cumulative share of NTER communities that were covered by income management on each of the implementation dates (indicated by crosses). Data are weighted by number of births in each community. For full details on the rollout schedule, see AIHW (2010).

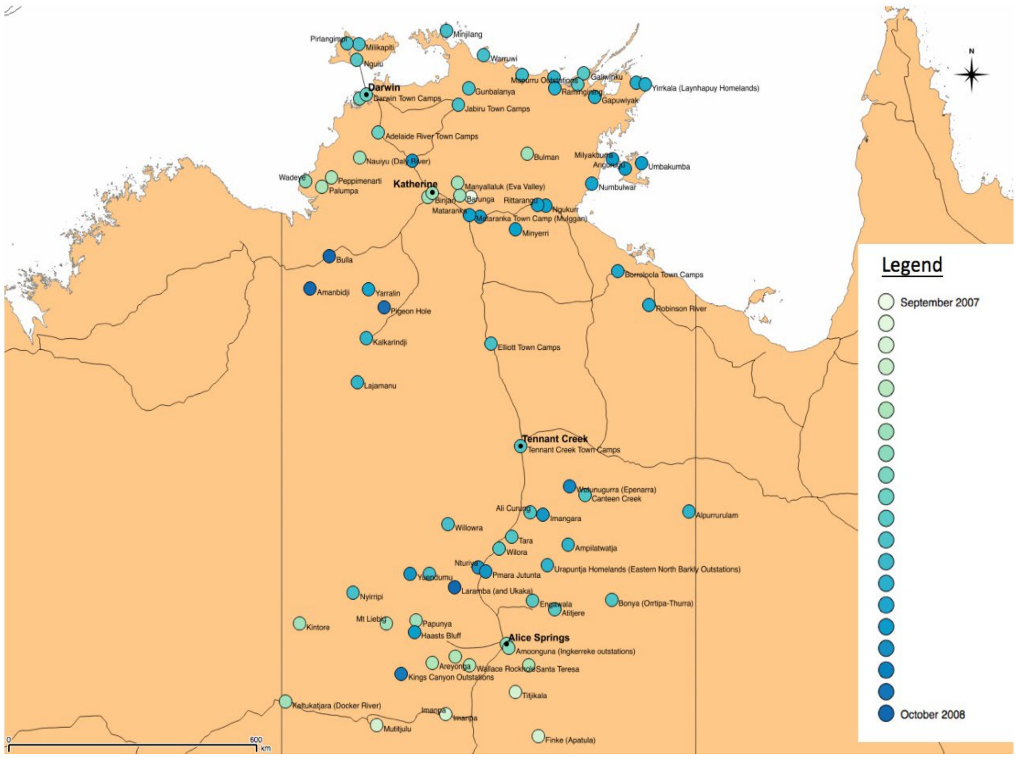

FIGURE 4. GEOGRAPHIC PATTERN OF THE INCOME MANAGEMENT ROLLOUT

Notes: Color coding reflects the date income management started in the relevant community, as indicated in the Legend. Major cities or towns in the Northern Territory are displayed for reference. People living in the municipal parts of those towns were not subject to IM, but people living in associated town camps were. Lines connecting communities represent major roads.

Third, in the year before the rollout, birth outcomes in communities that received income management early were no different from those that received it later. Constructed from our administrative data, Figure 5 suggests no apparent trend in birthweight in either group prior to the rollout.[17] We will show in a later section more formally (event study Figure 8) that the common trend assumption is valid.

[17] The dip in January 2007 represents a seasonal pattern (with generally worse birth outcomes during the wet season). This seasonality is controlled for in our econometric estimation.

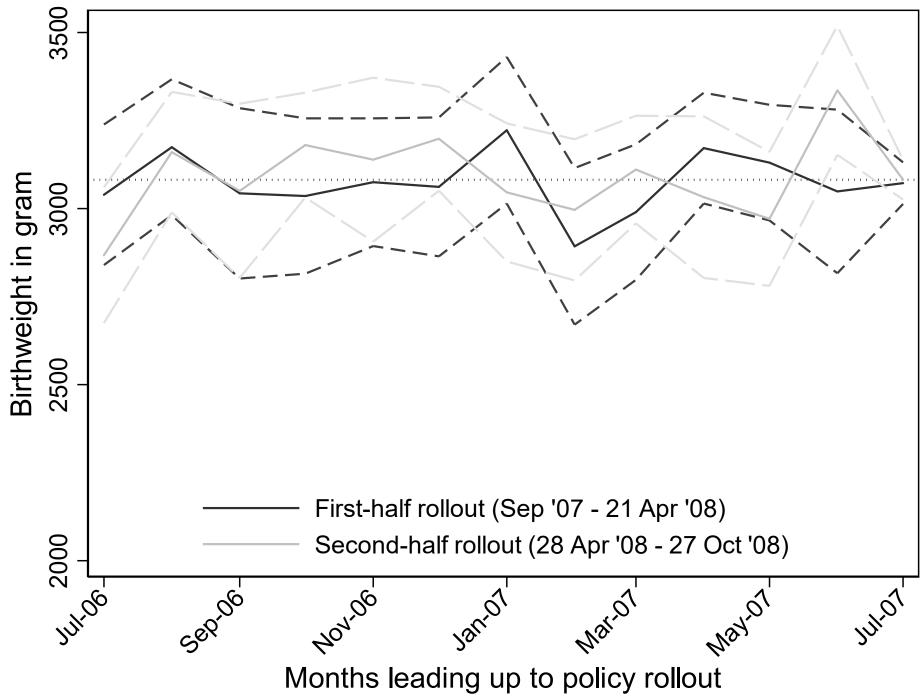

FIGURE 5. BIRTHWEIGHT TRENDS BY TIMING OF THE ROLLOUT

Notes: The graphs displays the unadjusted mean birthweight in NTER communities in each month leading up to the Income management policy, separately for communities that received IM in the first half of the rollout (weighted by the number of births) and the half of all communities. Dashed lines indicated 95% confidence intervals. There were no births between 21-28 April 2008.

Fourth, we see no pre-existing level differences in terms of average birth outcomes, birth complications, or community characteristics in the year before the rollout. Table 1 reports mean differences between early- and late-adopting communities,[18] using data from the NT Perinatal files and the 2006 Australian Bureau of Statistics Census. Table B1 (Appendix) show these same variables, but for the first and last ten communities to receive income management.[19]

[18] Early and late adopters are defined as communities where income management was implemented between September 2007 and 21 April 2008, and between 22 April 2008 and October 2008, respectively.

[19] These tables test for the possibility that the rollout schedule was intended to target the most in-need communities first, and the least in-need communities last, which would downwardly bias our estimated treatment effect. If true, we would expect the very first communities to have below-average pre-intervention outcomes, and the very last communities to be above-average. Yet, the very first-adopting communities had slightly higher pre-intervention birthweight and similar probability of low birthweight to other NTER communities

TABLE 1: SUMMARY STATISTICS, THE YEAR BEFORE INCOME MANAGEMENT
(1 JULY 2006 – 30 JUNE 2007)

	NTER communities			Rest of NT
Outcome variables	Communities in first half of rollout	Communities in second half of rollout	Difference	
Birthweight (grams)	3072	3082	9.89	3354
	(30)	(29)	(41.67)	(11)
Low birthweight (%)	14.44	13.95	-0.49	6.83
	(1.62)	(1.61)	(2.28)	(0.48)
Obstetric complications				
Premature (%)	15.5	15.02	-0.48	7.91
	(1.67)	(1.66)	(2.35)	(0.51)
Due to intrauterine growth restriction (%)	5.1	3.22	-1.88	1.44
	(1.01)	(0.82)	(1.3)	(0.23)
Due to anemia (%)	9.55	9.87	0.317	2.01
	(1.35)	(1.38)	(1.94)	(0.27)
Due to gestational diabetes (%)	7.22	9.01	1.79	6.83
	(1.19)	(1.33)	(1.78)	(0.48)
Any complication (%)	43.52	45.92	2.4	24.19
	(2.28)	(2.31)	(3.25)	(0.81)
Other characteristics				
Age of mother	23.88	23.74	-0.14	28.56
	(0.28)	(0.28)	(0.4)	(0.12)
APGAR 5	8.79	8.88	0.09	8.97
	(0.07)	(0.07)	(0.1)	(0.02)
Community characteristics				
Community size	388.84	486.45	97.61	na
	(55.96)	(67.09)	(88.77)	
Female share of population (%)	50.83	50.96	0.13	48.49
	(0.6)	(0.75)	(0.98)	
Median age	22.81	22.14	-0.67	31
	(0.38)	(0.41)	(0.58)	
People per household	5.39	6.53	1.14	2.9
	(0.23)	(0.22)	(0.32)	
Median personal income ($)	214.62	206.61	-8.01	549
	(10.45)	(3.1)	(10.77)	
Median rent payments ($)	43.91	42.21	-1.69	140
	(3.17)	(6.15)	(7.22)	
Labor force share of population (%)	39.85	36.43	-3.42	47.27
	(3.24)	(2.81)	(4.36)	

Notes: Standard errors in parentheses. First half of rollout defined as communities where income management was introduced from 17 September 2007 to 21 April 2008, second half defined as communities where income management was introduced from 28 April 2008 to 27 October 2008. APGAR 5 stands for Appearance, Pulse, Grimace, Activity and Respiration measured 5 minutes after birth. Each of the five categories is scored 0, 1 or 2, for a maximum total score of 10.

Most birth outcome measures, including obstetric complications, characteristics of the mother and APGAR scores[20] were similar between the two groups before the rollout. The only notable differences are observed for some community-level characteristics. Early-adopting communities were smaller on average by 100 community members, and families were smaller by one household member (5.4 versus 6.5). Community composition and median age were not significantly different between early and late adopters, nor were local economic conditions (as proxied by the median personal income and the labor force-to-population ratio).

A separate regression model to predict policy implementation timing confirms that community characteristics are not predictive, explaining only around 10 percent of the variation in timing (Appendix Table B.2).[21]

Finally, some residents in early-adopting communities may have wished to move to late-adopting communities to delay participating in income management. But the scope for this was very limited, as eligibility was determined based on place of residence one week after the policy was announced. Cobb-Clark et al (2018) show empirically that income management did not impact short-term mobility.

For these reasons, we conclude that the rollout timeline can be considered exogenous to our outcome measure.

B. Econometric model

We estimate the causal effect of the introduction of income management using a generalized difference-in-differences (DID) specification. Denoting the outcome

[20] APGAR is a test (appearance, pulse, grimace, activity and respiration) given to newborns at 1 minute and 5 minutes after birth.

[21] Similar to Hoynes and Schanzenbach (2009) we estimated a regression model in which an index of the timing of the reform, indexed to 1 for 17 September 2007, was regressed on pre-treatment community characteristics, levels in birthweight, and rainfall. We find no significant association between any of the variables and the timing of the reform, except for a significant coefficient on household size. Overall, our extended set of control variables in this regression explain up to 12 per cent of the variation in the roll-out date, which suggests that most of the variation remains unexplained. This weakness in model fit is a strength for our identification strategy, and the negative coefficient on birthweight operates in the opposite direction from our treatment effects. See Hoynes and Schanzenbach (2009) for similar arguments in the context of the roll out of the Food Stamps program.

variables (birthweight and low birthweight) for baby *i* born at time *t* in community *c* by Y_{itc}, our main regression equation is given by:

$$(3) \qquad Y_{itc} = \alpha + \tau P_t + \delta IM_{itc} + \pi P_t \times IM_{itc} + \gamma X_{ic} + \eta_c + \theta_t + \rho S_{tr} + \varepsilon_{itc},$$

where η_c denotes community fixed effects, θ_t year fixed effects and S_{tr} captures controls for region-specific seasonal patterns (see Appendix Figure B1).[22] X_{ic} is a set of individual-level control variables: the sex of the baby, an indicator for whether it is the mother's first pregnancy, the mother's age (categorized into 5-year age groups), and an indicator for whether the baby was born prematurely. We include pre-term birth as a control variable because it proxies hard-to-measure seasonal variation at the time of conception (Darrow et al. 2009), which we observe in our case before treatment occurs. Furthermore, controlling for pre-term birth allows us to isolate the intrauterine growth channel, which is more likely to respond to a change in maternal nutrition (see Section *I.D* above).

Importantly, P_t is the rollout period indicator, which equals 1 during the rollout period, and 0 for the pre-rollout period. IM_{itc} is the 'treatment' indicator, which is equal to 1 if income management was in place in community *c* at the beginning of the third trimester of pregnancy, and 0 otherwise. The interaction term between the rollout period and the reform indicator $P_t \times IM_{itc}$ is of main interest.

Our treatment estimate of π captures intention-to-treat (ITT) effects. We do not observe which mothers received government transfers (and thus were directly

[22] Optimally, we would like to control for time-varying community fixed effects. Unfortunately, we ask too much of the data, given the small populations in some communities. Instead, we include an indicator for whether the baby was conceived in the wet season or the dry season, as well as a control for total region-level rainfall (in ml) in the three months to birth. We define four regions: Darwin and surrounds; East Arnhem; Katherine and Barkly; and Alice Springs and surrounds. This method of controlling for seasonality allows for the timing of the wet season to vary from year to year. Figure 7 shows that results are robust to different approaches to controlling for seasonal variation (e.g. community-level instead of region-level rainfall, using year-quarter fixed effects and interacted year-season fixed effects, or controlling for month of conception instead of season of conception).

affected by income management). But with the intervention affecting most of the population, we expect ITT effects are close to the treatment-on-treated effects.

We use ordinary least squares to estimate the impact of income management on birthweight, and a linear probability model to estimate its impact on the probability of low birthweight. In addition, we estimate quantile treatment effects in order to investigate potential heterogeneity along the birthweight distribution. Standard errors are clustered at the community level.

VI. Estimation results

We first report the ITT estimates from our benchmark model. We then report the outcomes of a series of robustness checks, and a heterogeneity analysis with respect to the distribution of birthweight and the intensity of treatment exposure.

A. Benchmark model

Table 2 reports the results of our main model. In a model without control variables (columns 1 and 4), we find that average birthweight is around 10 grams lower in the treatment group than in the untreated group, with a standard error (S.E.) of 33 grams. The probability of low birthweight is no different between groups.

Adding all controls except for community fixed effects and premature birth (columns 2 and 6) we find a reduction in birthweight of 117 grams (S.E. 45g) and an increase in the probability of low birthweight to 5.4 percentage points (S.E. 2ppt). Further controlling for community fixed effects (columns 3 and 7) increases the absolute value of the estimated treatment effects, implying a reduction in birthweight of 130 grams (S.E. 54) and an increase in the probability of low birthweight of 5 percentage points (S.E. 2.3pts). Both are statistically significant at the 5 percent level. Finally, our preferred model – which also controls for premature birth – suggests slightly smaller treatment effects. It shows that income management reduced birthweight by 95 grams (S.E. 50g) and increased probability of low birthweight by 3 percentage points (S.E. 2.1ppts)

TABLE 2 – IMPACT OF INCOME MANAGEMENT ON BIRTHWEIGHT AND PROBABILITY OF LOW BIRTHWEIGHT

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Outcome: Birthweight (grams, OLS)				Outcome: Low birthweight (LPM)			
Income	-10.41	-116.73	-129.70	-94.89	0.000	0.054	0.050	0.030
management	(32.85)	(45.39)	(54.00)	(49.81)	(0.017)	(0.020)	(0.023)	(0.021)
Pre-rollout	-55.89	55.80	42.40	-21.55	0.003	-0.055	-0.053	-0.016
period	(31.38)	(61.46)	(66.13)	(59.71)	(0.018)	(0.034)	(0.037)	(0.031)
Rainfall in ml		-0.15	-0.05	-0.07		0.000	0.000	0.000
		(0.05)	(0.05)	(0.04)		(0.000)	(0.000)	(0.000)
Conceived in		41.06	69.98	26.56		-0.032	-0.036	-0.011
wet season		(40.82)	(43.40)	(33.21)		(0.019)	(0.020)	(0.019)
Sex (male = 0)		-72.98	-70.33	-75.91		0.038	0.038	0.040
		(28.97)	(31.28)	(25.66)		(0.017)	(0.018)	(0.014)
Mother's first		-14.80	-13.41	-1.49		-0.008	-0.009	-0.016
pregnancy		(29.18)	(30.45)	(25.68)		(0.018)	(0.018)	(0.012)
Mother's age (base cat. = under 20)								
20-24		64.19	64.20	60.89		-0.004	-0.000	0.002
		(35.52)	(36.93)	(27.89)		(0.023)	(0.024)	(0.018)
25-29		184.48	200.51	184.06		-0.020	-0.020	-0.010
		(44.06)	(46.62)	(33.32)		(0.022)	(0.024)	(0.016)
30-34		249.66	262.10	258.33		-0.035	-0.035	-0.033
		(57.42)	(61.24)	(50.85)		(0.028)	(0.029)	(0.022)
35+		158.43	154.06	165.99		0.027	0.030	0.020
		(71.72)	(74.45)	(54.53)		(0.041)	(0.043)	(0.031)
Premature				-945.11				0.565
				(35.91)				(0.028)
Constant	3,149.88	2,850.91	3,328.27	3,501.12	0.128	0.254	0.120	0.020
	(24.26)	(87.10)	(80.66)	(82.82)	(0.013)	(0.050)	(0.050)	(0.051)
Year fixed effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Community fixed effects	No	No	Yes	Yes	No	No	Yes	Yes
Number of communities	85	85	85	85	85	85	85	85
Observations	1,983	1,982	1,982	1,982	1,984	1,983	1,983	1,983
R-squared	0.00	0.04	0.09	0.37	0.000	0.018	0.058	0.392

Notes: Community-clustered standard errors in parentheses Low birthweight is defined as less than 2500 grams. Income management is estimated as the interaction between treatment and income management rollout period. As the base category in this interaction is the untreated group within the rollout period, this coefficient can be directly interpreted.

B. Robustness tests

Our finding – of a large negative and statistically significant effect of income management on birth outcomes – is robust to changes in research design (matching methods), sample selection criteria and restrictions, and model specifications, which allow for community specific seasonal trends or alternative time and seasonal trend definitions. These robustness tests are summarized in Figure 6, with full details of each model in Appendix C. The treatment effects vary between -60 and -200 grams. In some cases, where sample sizes become small or many additional parameters are estimated, treatment effects are estimated less precisely.

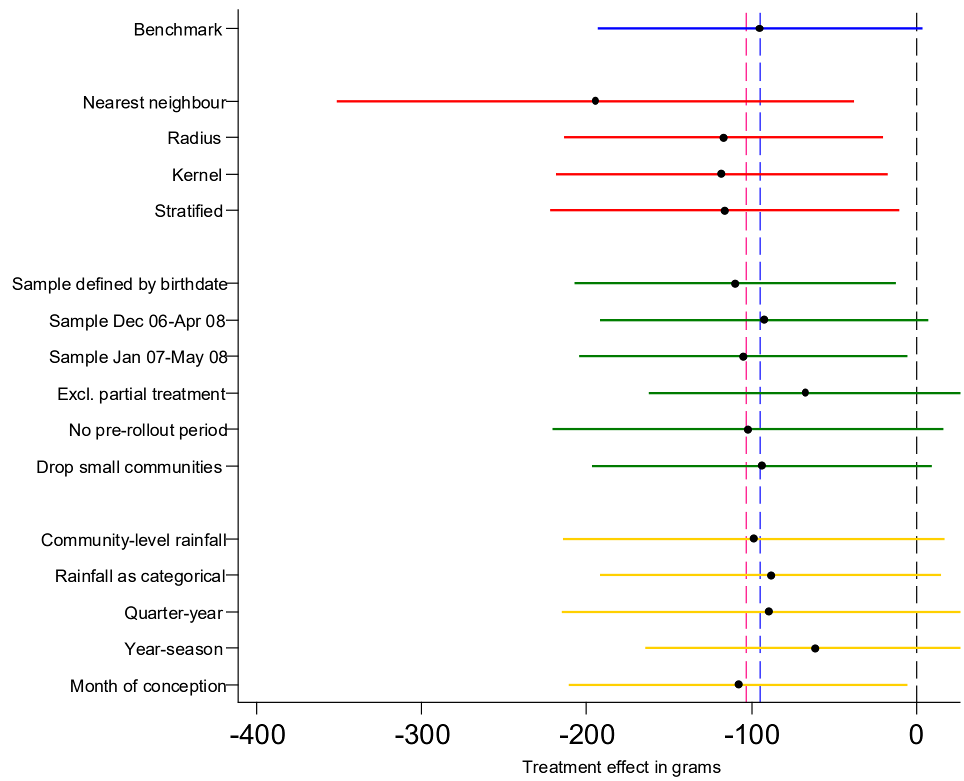

FIGURE 6. ROBUSTNESS CHECKS ON BENCHMARK AVERAGE TREATMENT EFFECT

Note: The figure depicts the treatment effect of income management on average birthweight obtained from our benchmark model, compared with estmates obtained from a series of robustness tests,with 95% confidence intervals. Each row is the estimated treatment effect (dot point) obtained from a separate regression model (results tables available in Appendix C).

To test the validity of our identification strategy, we tested for placebo reform effects (Table 3). We run our main specification with a one- to three-year lead on treatment timing, using 'placebo' rollout periods before income management was introduced (2004-2006; 2005-2007, 2006-2008). It reveals small and statistically insignificant placebo effects on both outcomes. These estimates increase our confidence that our main treatment estimate captures the effect of the treatment, and not simply unobserved trends in birth outcomes that were present before income management.

TABLE 3 – PLACEBO TESTS

Years lead	Rollout date range (by conception date)			Birthweight	Low birthweight
Actual sample	17-Jan-07	to	30-Apr-08	-102.98	0.040
				(45.79)	(0.020)
1	17-Jan-06	to	30-Apr-07	-15.51	0.002
				(41.91)	(0.024)
2	17-Jan-05	to	30-Apr-06	-39.11	0.025
				(38.56)	(0.023)
3	17-Jan-04	to	30-Apr-05	-18.58	-0.003
				(36.06)	(0.026)

Notes: Community-clustered standard errors in parentheses. Regressions include community and year fixed effects and controls for: premature birth, rainfall, conceived in wet season, sex of baby, whether mother's first pregnancy, and mother's age.

C. *Heterogeneity in treatment effect*

Across the birthweight distribution – Our finding of a decline in average birthweight could represent better infant health if concentrated among heavier babies. However, quantile regressions estimates (Figure 7), show the effect was present throughout the distribution, and largest at the bottom of the distribution.

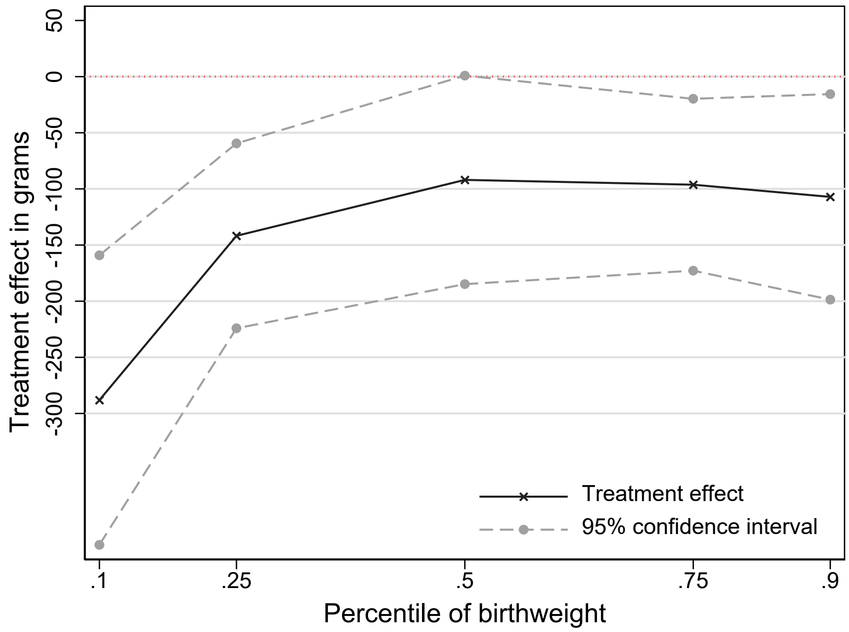

FIGURE 7. TREATMENT EFFECT BY BIRTHWEIGHT QUANTILE

Notes: Estimated coefficients obtained from quantile regression models, where conditional treatment effect is estimated with community and year fixed effets, and controls for: premature birth, rainfall, conceived in wet season, sex of baby, whether mother's first pregnancy, and mother's age.

By length of exposure – To test whether effects differ by length of exposure, we conduct an event study version of our main model, allowing the treatment effects to vary by the length of time the baby was exposed to income management in utero.

Almond, Hoynes and Schanzenbach (2011) found an effect of exposure to SNAP during the third trimester, but no additional impact if it was introduced before the third trimester. Their finding is consistent with the third trimester being a key period for nutrition and intrauterine growth (Kramer 1987). Based that evidence, we had expected a similar pattern in our context. But contrary to our expectations, newborns who were exposed to income management only for part of the pregnancy did not experience significantly adverse birth outcomes. It was those exposed for the full pregnancy for whom the effect was largest and statistically significant. This pattern suggests that the treatment effect may have come through some channel other than – or in addition to – the effect of nutrition on intrauterine growth. Another interpretation may be that the dose of exposure to income management throughout gestation mattered more than exposure during any critical period.

Figure 8 also shows that there were no significant pre-treatment trends in birthweight (before time period *t*). This gives us more certainty that the identification strategy is valid.

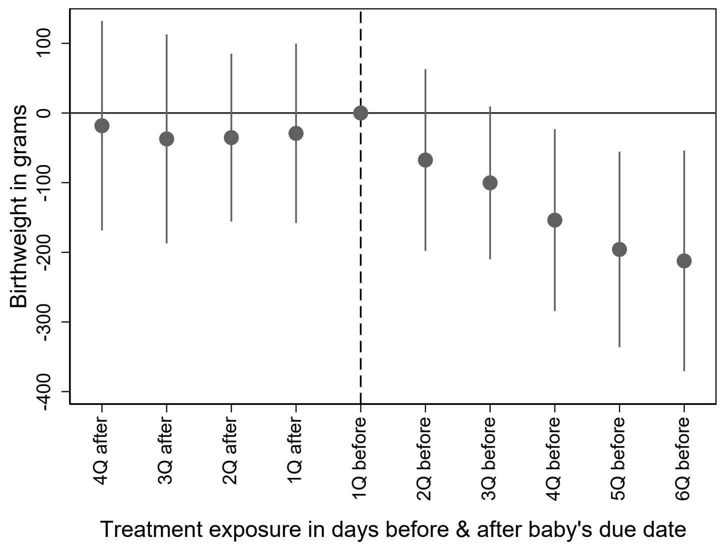

FIGURE 8. TREATMENT EFFECT BY DURATION OF EXPOSURE

Note: This graph shows an event study version of our main treatment effect with 95% confidence intervals. Time periods are based on the length of exposure to income management relative to the baby's full-term due date, in quarters. The base category '1Q before' indicates income management was introduced between 0 and 89 days before the baby's due date (i.e. during in the final trimester). Babies born in 'after' periods were not exposed to the policy in utero. We use the same controls as the benchmark model, and standard errors are clustered by community.

VII. Why did income management reduce birthweight?

Why did the introduction of income management worsen health outcomes? We explore four channels that may explain this effect: changes in fertility and maternal characteristics, maternal risky health behaviours, better access to quality care, and temporary or extended income shocks.

A. Fertility and maternal characteristics

Income management may have changed fertility decisions and the composition of women who planned to be pregnant during this period. Healthier mothers, expected to have healthier babies, may have opted to postpone pregnancy, potentially reducing average birthweight. We find no evidence for this hypothesis. The policy had no impact on community-level fertility rates, as measured by the number of births per resident (Table 4, Panel A). We also do not see a significant difference in the medical history of women who gave birth after income management was introduced, either in terms of their previous pregnancies (Panel B), or their history of medical complications (Panel C). Therefore, not surprisingly, controlling for these additional maternal characteristics does not change the treatment effect of income management (Appendix Table C.5 cols 1 and 4).

TABLE 4 – FERTILITY AND MOTHER'S MEDICAL HISTORY

	Treated	Untreated	Difference (treated-untreated) No controls	Difference (treated-untreated) With controls	Observations
Panel A: Fertility rate (community-level)					
Births per 1000 women per quarter	15.26	15.25	-0.04	-0.99	290
			(2.99)	(1.6)	
Panel B: Previous pregnancies					
First pregnancy (%)	26.61	25.86	0.75	-0.64	992
			(3.21)	(3.97)	
Total number of previous pregnancies	2.92	2.88	0.04	-0.01	992
			(0.11)	(0.15)	
Panel C: Mother's history of medical complications					
Any complication (%)	57.08	53.99	3.09	-1.41	992
			(3.47)	(5.66)	
Number of previous complications	0.80	0.70	0.09	0.04	992
			(0.06)	(0.11)	
Medical history unknown (%)	8.45	8.04	0.41	0.03	1081
			(1.55)	(2.53)	

Notes: Community-clustered standard errors in parentheses. 'Difference' columns report coefficients for OLS or LPM regressions. 'With controls' includes community and year fixed effects and controls for: premature birth, rainfall, conceived in wet season, sex of baby, whether mother's first pregnancy, and mother's age. In Panel A, regressions are conducted on data averaged at the community-quarter level. Because income management may be introduced during the quarter, the treatment variable can be nonbinary. Treatment status is rounded for the treatment and control columns. In Panel C, complications are: anemia, cardiac disease, epilepsy, pre-existing hypertension, pre-existing diabetes, gestational diabetes, syphilis, and urinary tract infections.

B. *Maternal health behaviors*

Alternatively, the negative treatment effect may be the result of a change in maternal health behaviours. Income management intended to create a healthier consumption environment, but it may have in fact increased maternal risky health behaviours. For instance, income management could have created a new – lower – mental anchor on how much money should be spent on priority goods, potentially increasing spending on alcohol and tobacco. We do not have consumption data available, but our perinatal data include self-reported information on whether the mother was drinking or smoking at the time of the first antenatal visit.[23]

TABLE 5 – IMPACT OF INCOME MANAGEMENT ON SMOKING AND DRINKING AT FIRST ANTENATAL VISIT

			Difference (treated - untreated)		
	Treated	Untreated	No controls	With controls[(a)]	Pr missing
Panel A: Smoking					
Smoking at first antenatal visit	49.36	51.87	-2.51	4.39	0.05
			(4.46)	(4.45)	(0.03)
Constant			51.9	58.9	
			(3.56)	(12.7)	
Observations	393	347	1,506	1,506	949
Panel B: Drinking					
Drinking at first antenatal visit	12.68	10.24	2.43	-0.64	0.04
			(2.26)	(2.66)	(0.02)
Constant			10.2	14.1	
			(1.72)	(8.73)	
Observations	426	371	1,541	1,541	949

Notes: Columns 3-5 are linear probability models with community-clustered robust standard errors. In this table, treatment timing is defined relative to the date of the first antenatal visit, not relative to the date of birth (as in the main analysis). An observation is defined as 'treated' if the first antenatal visit occurred on or after the date that Income Management was introduced in the mother's community. 'With controls' includes community and year fixed effects and controls for: premature birth, rainfall, conceived in wet season, sex of baby, whether mother's first pregnancy, and mother's age. The sample size is larger for the regression estimates, as the regression includes the 1 year before income management was introduced, with the treatment effect estimated as the interaction between the rollout period and treatment.

[23] Treatment here is defined to take the value 1 if the first antenatal visit occurred after income management was introduced into the mother's community, and 0 otherwise. The sample of pregnancies covered is therefore different from our main analysis sample.

Table 5 shows that if income management was introduced before the first antenatal visit, women were 4.4 percentage points more likely to report smoking at that time (a 7.5 percent increase), but the S.E. is so large that we have no certainty that this effect is not due to random variation (Panel A). There was no difference in the probability of drinking alcohol (Panel B). We thus conclude that changes in smoking and drinking behaviour are unlikely to explain the treatment effect. Controlling for smoking and drinking behaviour directly in the regression model does not change the treatment effect (see Appendix Table C.5, cols 2-3, 5-6).

C. Access to quality care

A reduction in birthweight or an increase in the probability of low birthweight does not necessarily indicate a worsening in birth outcomes. The introduction of income management could have increased the likelihood of receiving earlier or more comprehensive antenatal care through more frequent contact with government staff. This could have led to better monitoring of fetal and maternal health, earlier detection of serious complications, and thus referral to emergency C-sections. This could have increased survival probabilities of at-risk babies. Despite lower birthweights, it would have been a preferred health outcome.

But we do not find evidence for this hypothesis. Table 6 demonstrates that treated and untreated babies did not differ in the probability of receiving antenatal care and did not have significantly different probabilities of being born in a major hospital or by emergency delivery. Consistent with this finding, the treatment effect is robust to controlling for these variables (Table C.6, Appendix).

TABLE 6 – INDICATORS OF ANTENATAL CARE AND HOSPITAL CARE

	Level		Difference (treated-untreated)		Observations	
	Treated	Untreated	No controls	With controls	N	Pr missing (T-U)
Panel A: Antenatal care						
N. antenatal visits	8.68	8.95	-0.352	-0.234	987	-0.60
			(0.34)	(0.47)		(0.55)
Gest. age at first visit (weeks)	16.24	16.63	-0.430	-0.773	983	-1.05
			(0.45)	(0.61)		(0.83)
Panel B: Care at birth						
Born in main hospital (%)	71	65	6.15	-0.371	1,086	na
			(4.23)	(3.64)		
Emergency delivery (%)	20	19	1.13	1.82	1,086	na
			(2.39)	(4.69)		
Days' stay in hospital	6.47	6.56	-0.089	1.140	1,072	0.90
			(0.68)	(0.7)		(0.71)
Special care nursery (%)	23	22	0.983	0.166	1,085	-0.17
			(2.55)	(3.87)		(0.18)

Notes: Community-clustered standard errors in parentheses. "With controls' includes community and year fixed effects and controls for: premature birth, rainfall, conceived in wet season, sex of baby, whether mother's first pregnancy, and mother's age. Main hospitals are Royal Darwin Hospital, Darwin Private or Alice Springs Hospital.

TABLE 7: IMPACT OF INCOME MANAGEMENT ON MORTALITY AND VERY LOW BIRTHWEIGHT

	(1)	(2)	(3)	(4)
Panel A: Linear probability models on stillbirth				
Income management	0.003	0.001	0.005	-0.002
	(0.006)	(0.005)	(0.007)	(0.008)
Pre-rollout period	0.010	0.005	-0.012	-0.008
	(0.006)	(0.005)	(0.013)	(0.009)
Born before third trimester		0.379		0.318
		(0.083)		(0.088)
Constant	0.010	0.006	0.039	0.034
	(0.005)	(0.004)	(0.023)	(0.018)
Controls and FE	no	no	yes	yes
Observations	2,020	2,020	2,019	2,019
Panel B: Linear probability model on very low birthweight (<1500)				
Income management	0.008	0.042	0.030	
	(0.008)	(0.012)	(0.013)	
Pre-rollout period	0.031	-0.035	-0.018	
	(0.010)	(0.029)	(0.026)	
Premature birth			0.215	
			(0.023)	

Constant	0.019	0.102	0.059
	(0.006)	(0.051)	(0.047)
Controls and FE	no	yes	yes
Observations	2,020	2,019	2,019

Notes: Community-clustered standard errors in parentheses. The sample in this table includes infants born before the beginning of the third trimester of pregnancy, who are dropped from our main analysis.

D. Short term income shocks

Having ruled out medical and behavioral channels as cause of the worsening of birth outcomes, we are left with a more qualitative assessment of the policy's impact. The negative effects of the policy could have come from a temporary income shock that was caused by the new procedures for accessing benefit funds. Two institutional sources of disruption were observed (AIHW 2010). One came through a rule that required benefit recipients to allocate the quarantined part of their entitlements to specific spending categories. If a recipient failed to meet with a government welfare agency caseworker to discuss those allocations, the money was 'auto-income managed'. This meant that the money was quarantined, but the recipient had no way of accessing it until they met with a caseworker.

Auto income management happened frequently at the beginning of the rollout: two months into the rollout, about 50 percent of recipients could not access their quarantined benefits (AIHW 2010, p. 30). Up until 26 September 2008 (one year into the rollout), about 20 percent of quarantined funds were still not allocated and thus could not be spent. Although this problem was fully resolved after the rollout was completed, the implication is that for most of the period, about one in five dollars earmarked for the consumption of essentials was not available.

A second source of disruption was payment suspensions. If a recipient's account was suspended, then all benefit payments were missing in the account. This happened when recipients made administrative errors, remained on auto-income management for 13 weeks or more, or went to jail.

According to the AIHW (2010) report, the number of affected community members was large (p. 26-27, Table 9). About one-third of all 21,763 clients had at least one payment suspended. With an average community size of 440 members (Table 1, average of cols 1 and 2), on average 107 clients per community were missing at least one full payment. During the rollout period, in total 9,846 payments were missing. Assuming that one payment was equivalent to AU$210 per week (AU$105 quarantined),[24] then in total AU$2,067,660 were missing. Although most payments were restored, by March 2009, 3,020 payments were still missing. We do not know how many of these long-term payment suspensions were due auto-income management, but we know that this was the most common reason for short-term suspensions (AIHW 2010, Table 9).

During the rollout period, 1,833 clients had at least one payment suspended after they failed to contact the government welfare agency within 13 weeks of being auto income-managed. For these clients, only half of their benefit payments were available for 13 weeks in a row – they could only access the cash portion of their benefits. The other half of their entitlements were accumulated in their quarantined account as unallocated funds. If by week 14 the client still did not contact a caseworker, then all payments, including the cash payout, were suspended.

If we assume that suspensions due to prolonged auto-income management were equally distributed across the 83 communities and town camps, then 22 individuals per community lived for three months on half of their entitlements and lost a full entitlement in week 14. Independent of whether the suspended payment was restored in week 14, those clients would have been short by at least AU$1,365 (that

[24] This is based on Australian Bureau of Statistics estimates of Household Income and Income Distribution for 2005-06, Table 6523.0.55.001. As a conservative estimate of the typical value of welfare payments, we take the estimate of income per week for households at the 10th percentile of the national income distribution, for whom $213 per week on average comes from government pensions and allowances (out of a total household income of $274).

is, $105 quarantined over 13 weeks). These shortfalls would represent a severe temporary income shock.[25]

VIII. Discussion and conclusion

Over the past 13 years, Australia has continued to expand income management, moving many more benefit recipients from unconditional cash transfers to restricted transfers. Although the welfare of children has taken centre stage in the policy debate, there is little evidence on how income management affects children. This study is one of the first attempts to quantify the impact on children's welfare.

We conclude that the introduction of income management did not improve infant health. Unexpectedly, it appears to have had a negative effect. Our benchmark estimate suggests a reduction in birthweight of 95g and an increased probability of low birthweight of 3 percentage points. Our estimates are robust.

These effects are large, but they are within the range of estimates from other contexts. For instance, Savitri et al (2014) find that mothers who fasted during pregnancy gave birth to babies that were on average 200 grams lighter. The effect was largest for babies exposed to fasting during the first trimester – a finding consistent with our estimates. Stein and Susser (1975) found that famine exposure during pregnancy reduced birthweight by around 150 grams.

While we do not observe consumption directly, our findings suggest it is highly unlikely that income management led to the intended increase in consumption of food and other essentials – at least among pregnant women. Our findings stand in contrast to the main government report on the introduction of income management (AIHW 2010), which relied on focus groups and a small, non-randomly sampled survey. It suggested a broadly positive effect of income management on children’s

[25] There were other reasons for which clients’ accounts were suspended. For one in six suspensions, it was because clients were in jail. For another third of suspensions, it was because of administrative errors (failure to respond to correspondence, failure to sign an activity statement and failure to attend an interview).

food consumption and health. This contrast underscores the importance of appropriate methods when evaluating the overall effects of major policy changes.

While we are uncertain of the reason for the negative effect, we can eliminate three channels: income management did not have its effect through a change in fertility, a change in mothers' risky health behaviors, or a change in quality of perinatal care and survival of at-risk fetuses.

Ruling out these explanations, we propose that the effect is due both to a reduction in consumption, and to an increase in maternal stress. Stress experienced by pregnant women can affect development in utero, as cortisol is passed on to the fetus (Aizer, Stroud, and Buka 2016). Three pieces of evidence suggest pregnant women were exposed to a high-stress environment during the rollout of income management – and that part of this was because of difficulties accessing quarantined funds, which is likely to have reduced their consumption.

First, when benefit recipients become income managed, they cannot access their quarantined funds before discussing with a case worker how the money will be spent. This administrative hurdle was the most common reason for payment suspensions, which led to a stop of all payments to which a beneficiary was entitled. This happened to one-third of all income-managed recipients during the rollout period (AIHW 2010). Returning to our theoretical framework model described in section II.A, it implies that we cannot assume that payments de facto remained the same when moving onto income management. Many participants experienced income shortfalls that lasted between one to 14 weeks, or more.

Second, even after gaining access to quarantined funds, the new way of managing money reportedly caused stress for many recipients, as summarized in the NTER Review Board's report: "People were required to master new, complex and often challenging procedures with a minimum of information or explanation. This led to confusion and anxiety, especially because the vast majority of recipients speak English as a second or third language" (2008, p. 20).

Third, it is likely that income management also disrupted existing intra-household spending patterns, which may have affected both consumption and stress levels. For instance, restrictions on spending make it more difficult for families to pool resources if one parent is away from home. It may also cause stress and conflict within the household. Previous research has suggested tentative evidence that the number of household arguments that affected children doubled after the introduction of income management (Cobb-Clark et al 2018).

With our data, we cannot disentangle the relative contribution of these factors. And we do not know whether outcomes improved after the initial policy rollout – in effect our estimates measure the combined impact of the policy itself, and the way it was implemented. But it is likely that the way the policy was first introduced affects its continued operation. The introduction of income management is reported to have led to a sense of loss of freedom, disempowerment, and reduced community control. The government evaluation report stated that some survey respondents perceived the program as 'patronising and dehumanising', with many highlighting the suspension of Part II of the *Racial Discrimination Act* as a contributing factor to this perception (AIHW 2010). These attitudes can be persistent. As Dalley (2020) describes based on research with an Aboriginal community in Western Australia, the racially targeted nature of income management (now named 'Cashless Debit Card') has continued to affect the way that many participants engage with it, even a decade later.

While our findings raise additional questions, they also convey a clear message. The unexpected negative effect of the introduction of income management highlights the importance of careful design, testing and consultation for major social policy changes. As Gracey and King (2009) describe, there are demonstrated and persistent inequalities between indigenous and non-indigenous peoples worldwide. In advanced economies, these inequalities have been slow to change across many measures of human development (Cooke et al. 2007). Income management represents one attempt to reduce these disparities, but we find evidence it exacerbated them – at least in the short term.

Whether or not income management has continued to have the same negative effects for subsequent cohorts, our findings indicate that the cohort of children born during this period were negatively impacted by the introduction of income management. For those individuals, the international evidence suggests that this initial early life shock may persist throughout their lives, weighing on their health, education and career prospects (Almond, Currie, and Duque 2018).

REFERENCES

Aboriginal and Torres Strait Islander Social Justice Commissioner. 2005. 'Social Justice Report 2005'. 3. Sydney, Australia: Human Rights and Equal Opportunities Commission.

Aizer, Anna, and Janet Currie. 2014. 'The Intergenerational Transmission of Inequality: Maternal Disadvantage and Health at Birth'. *Science* 344 (6186): 856–61. https://doi.org/10.1126/science.1251872.

Aizer, Anna, Laura Stroud, and Stephen Buka. 2016. 'Maternal Stress and Child Outcomes: Evidence from Siblings'. *Journal of Human Resources* 51 (3): 523–55. https://doi.org/10.3368/jhr.51.3.0914-6664R.

Almond, Douglas, Janet Currie, and Valentina Duque. 2018. 'Childhood Circumstances and Adult Outcomes: Act II'. *Journal of Economic Literature* 56 (4): 1360–1446. https://doi.org/10.1257/jel.20171164.

Almond, Douglas, Hilary W. Hoynes, and Diane Whitmore Schanzenbach. 2011. 'Inside the War on Poverty: The Impact of Food Stamps on Birth Outcomes'. *Review of Economics and Statistics* 93 (2): 387–403. https://doi.org/10.1162/REST_a_00089.

Australian Human Rights Commission. 2011. 'The Suspension and Reinstatement of the RDA and Special Measures in the NTER'. 2 November 2011. https://www.humanrights.gov.au/our-work/suspension-and-reinstatement-rda-and-special-measures-nter-0.

Australian Institute of Health and Welfare, ed. 2010. *Evaluation of Income Management in the Northern Territory*. Occasional Paper / Department of Families, Housing, Community Services and Indigenous Affairs 34. Canberra.

Barber, Sarah L., and Paul J. Gertler. 2008. 'The Impact of Mexico's Conditional Cash Transfer Programme, Oportunidades, on Birthweight'. *Tropical*

Medicine & International Health 13 (11): 1405–14. https://doi.org/10.1111/j.1365-3156.2008.02157.x.

Breunig, Robert, and Indraneel Dasgupta. 2005. 'Do Intra-Household Effects Generate the Food Stamp Cash-Out Puzzle?' *American Journal of Agricultural Economics* 87 (February): 552–68. https://doi.org/10.1111/j.1467-8276.2005.00747.x.

Brimblecombe, Julie K., Joseph McDonnell, Adam Barnes, Joanne Garnggulkpuy Dhurrkay, David P. Thomas, and Ross S. Bailie. 2010. 'Impact of Income Management on Store Sales in the Northern Territory'. *The Medical Journal of Australia* 192 (10): 549–54. https://doi.org/10.5694/j.1326-5377.2010.tb03632.x.

Brough, Mal. 2007. 'Social Security And Other Legislation Amendment (Welfare Payment Reform) Bill 2007 – Second Reading Speech'. Presented at the Australian House of Representatives, Canberra, Australia, August 7.

Cobb-Clark, Deborah, Nathan Kettlewell, Stefanie Schurer, and Sven Silburn. 2018. 'The Effect of Quarantining Welfare on School Attendance in Indigenous Communities'. IZA Discussion Paper Series 11514. Bonn, Germany: IZA Institute of Labor Economics.

Commonwealth of Australia, Department of the Prime Minister and Cabinet. 2020. 'Closing the Gap Report 2020'.

Cooke, Martin, Francis Mitrou, David Lawrence, Eric Guimond, and Dan Beavon. 2007. 'Indigenous Well-Being in Four Countries: An Application of the UNDP'S Human Development Index to Indigenous Peoples in Australia, Canada, New Zealand, and the United States'. *BMC International Health and Human Rights* 7 (1): 9. https://doi.org/10.1186/1472-698X-7-9.

Cuffey, Joel, Timothy KM Beatty, and Lisa Harnack. 2016. 'The Potential Impact of Supplemental Nutrition Assistance Program (SNAP) Restrictions on Expenditures: A Systematic Review'. *Public Health Nutrition* 19 (17): 3216–31. https://doi.org/10.1017/S1368980015003511.

Cunha, Jesse M. 2014. 'Testing Paternalism: Cash versus In-Kind Transfers'. *American Economic Journal: Applied Economics* 6 (2): 195–230. https://doi.org/10.1257/app.6.2.195.

Currie, Janet, and Maya Rossin-Slater. 2015. 'Early-Life Origins of Life-Cycle Well-Being: Research and Policy Implications'. *Journal of Policy Analysis and Management* 34 (1): 208–42. https://doi.org/10.1002/pam.21805.

Dalley, Cameo. 2020. 'The "White Card" Is Grey: Surveillance, Endurance and the Cashless Debit Card'. *Australian Journal of Social Issues* 55 (1). https://doi.org/10.1002/ajs4.100.

Darrow, Lyndsey A., Matthew J. Strickland, Mitchel Klein, Lance A. Waller, W. Dana Flanders, Adolfo Correa, Michele Marcus, and Paige E. Tolbert. 2009. 'Seasonality of Birth and Implications for Temporal Studies of Preterm

Birth'. *Epidemiology* 20 (5): 699–706. https://doi.org/10.1097/EDE.0b013e3181a66e96.

Dawson, Jessica, Martha Augoustinos, David Sjoberg, Kootsy Canuto, Karen Glover, and Alice Rumbold. 2020. 'Closing the Gap: Examining How the Problem of Aboriginal and Torres Strait Islander Disadvantage Is Represented in Policy'. *Australian Journal of Social Issues* forthcoming. https://doi.org/10.1002/ajs4.125.

Department of Social Services. 2017. 'Freedom of Information Request Number 16/17-123: Costs of the Cashless Debit Card Trials in Ceduna and Kununurra.' Australian Government. https://www.dss.gov.au/sites/default/files/documents/04_2017/attachment_a_cashless_debit_card_trials_costs.pdf.

Gracey, Michael, and Malcolm King. 2009. 'Indigenous Health Part 1: Determinants and Disease Patterns'. *The Lancet* 374 (9683): 65–75. https://doi.org/10.1016/S0140-6736(09)60914-4.

Gresham, Ellie, Julie E Byles, Alessandra Bisquera, and Alexis J Hure. 2014. 'Effects of Dietary Interventions on Neonatal and Infant Outcomes: A Systematic Review and Meta-Analysis'. *The American Journal of Clinical Nutrition* 100 (5): 1298–1321. https://doi.org/10.3945/ajcn.113.080655.

Hastings, Justine, and Jesse M. Shapiro. 2018. 'How Are SNAP Benefits Spent? Evidence from a Retail Panel'. *American Economic Review* 108 (12): 3493–3540. https://doi.org/10.1257/aer.20170866.

Howard, John. 2007. 'To Stabilise and Protect: Little Children Are Sacred'. *Sydney Papers, The* 19 (3): 68.

Hoynes, Hilary W., Marianne Page, and Ann Huff Stevens. 2011. 'Can Targeted Transfers Improve Birth Outcomes? Evidence from the Introduction of the WIC Program'. *Journal of Public Economics* 95 (7–8): 813–27. https://doi.org/10.1016/j.jpubeco.2010.12.006.

Hoynes, Hilary W., and Diane Whitmore Schanzenbach. 2009. 'Consumption Responses to In-Kind Transfers: Evidence from the Introduction of the Food Stamp Program'. *American Economic Journal: Applied Economics* 1 (4): 109–39. https://doi.org/10.1257/app.1.4.109.

Kramer, M S. 1987. 'Determinants of Low Birth Weight: Methodological Assessment and Meta-Analysis'. *Bulletin of the World Health Organization* 65 (5): 663–737.

Lamb, David, and Martin Young. 2011. ''Pushing Buttons': An Evaluation of the Effect of Aboriginal Income Management on Commercial Gambling Expenditure'. *Australian Journal of Social Issues* 46 (2): 119–40.

Marston, Greg, Philip Mendes, Shelley Bielefeld, Michelle Peterie, Zoe Staines, and Steven Roche. 2020. 'Hidden Costs: An Independent Study into Income Management in Australia'.

Northern Territory Emergency Response Review Board. 2008. 'Northern Territory Emergency Response: Report of the NTER Review Board'. Canberra: Commonwealth of Australia. https://apo.org.au/node/551.

Risnes, Kari R., Lars J. Vatten, Jennifer L. Baker, Karen Jameson, Ulla Sovio, Eero Kajantie, Merete Osler, et al. 2011. 'Birthweight and Mortality in Adulthood: A Systematic Review and Meta-Analysis'. *International Journal of Epidemiology* 40 (3): 647–61. https://doi.org/10.1093/ije/dyq267.

Savitri, AI, N Yadegari, J Bakker, RJ van Ewijk, DE Grobbee, RC Painter, CS Uiterwaal, and TJ Roseboom. 2014. 'Ramadan Fasting and Newborn's Birth Weight in Pregnant Muslim Women in The Netherlands'. *British Journal of Nutrition* 112 (9): 1503–9. https://doi.org/10.1017/S0007114514002219.

Silburn, Sven, Steven Guthridge, John McKenzie, Jiunn-Yih Su, Vincent He, and Sharon Haste, eds. 2018. *Early Pathways to School Learning: Lessons from the NT Data Linkage Study.* Darwin: Menzies School of Health Research. https://www.nintione.com.au/?p=13869.

Southworth, Herman M. 1945. 'The Economics of Public Measures to Subsidize Food Consumption'. *American Journal of Agricultural Economics* 27 (1): 38–66. https://doi.org/10.2307/1232262.

Stein, Zena, and Mervyn Susser. 1975. 'The Dutch Famine, 1944–1945, and the Reproductive Process. I. Effects on Six Indices at Birth'. *Pediatric Research* 9: 70–76.

Strand, Linn Beate, Adrian G. Barnett, and Shilu Tong. 2011. 'Methodological Challenges When Estimating the Effects of Season and Seasonal Exposures on Birth Outcomes'. *BMC Medical Research Methodology* 11 (1): 49. https://doi.org/10.1186/1471-2288-11-49.

Venn, Danielle, Nicholas Biddle, and William Sanders. 2020. 'Trends in Social Security Receipt among Indigenous Australians: Evidence from Household Surveys 1994-2015'. Center for Aboriginal Economic Policy Research 135. Working Paper. ANU. https://openresearch-repository.anu.edu.au/handle/1885/204834.

FOR ONLINE PUBLICATION – APPENDICES FOR:

Unintended consequences of welfare reform: evidence from birth outcomes of Aboriginal Australians, by Mary-Alice Doyle, Stefanie Schurer and Sven Silburn

Appendix A: Other NTER measures

TABLE A.1 OTHER NTER POLICIES: CUMULATIVE NUMBER OF COMMUNITIES COVERED BY EACH POLICY, JULY 2007-JULY 2008

Policy measure	Jul-Sept 07	Oct-Dec-07	Jan-Mar 08	Apr-Jul 08	Target communities
Income management	4	23	33	78	83
Child health checks	22	48	69	81	83
School nutrition	3	7	25	68	73
Accelerated literacy	0	0	0	30	73
Quality teacher package	0	0	0	34	73
Leases	27	27	65	68	68
Store license	2	8	18	54	54
Safe house	0	0	0	10	73
Night patrols	0	0	1	14	43
Extra police	6	12	16	17	73
THEMIS police station	6	12	16	17	73
Remote Area Exemptions lifted	15	65	65	65	65
Community Development Employment Projects transition	3	30	30	30	83
Community Employment Brokers	25	38	54	69	83
Banning alcohol	73	83	83	83	83
Banning pornography	73	83	83	83	83
Remote Aboriginal Family and Community Workers	0	0	0	12	83
Child special services	0	0	0	12	83
Make safe works	2	24	44	68	68
Minor repairs	0	1	7	68	68
Asbestos survey	0	5	22	73	73
All Community Clean Up works completed	0	0	0	72	73
Government Business Managers	12	67	81	81	81

Notes: See NTER Review Board (2008) for details of each policy. Note that this table ends in July 2008, before the end of the rollout period. 5 communities received income management after July 2008.

Source: NTER Review Board, 2008

Appendix B: Descriptive statistics

TABLE B.1 – PRE-ROLLOUT OUTCOMES AND COMMUNITY CHARACTERISTICS, YEAR PRIOR TO NTER (1 JULY 2006 - 30 JUNE 2007)

	Earliest and latest NTER communities to receive income management			Rest of NTER
Outcome variables	First 10 communities	Last 10 communities	Difference	
Birthweight (grams)	3193.54	3162.84	-30.70	3056.83
	(71.3797)	(60.5082)	(95.17)	(23.0612)
Low birthweight (%)	14.63	9.46	-5.17	14.6
	(3.91)	(3.4)	(5.26)	(1.26)
Obstetric complications				
Premature (%)	14.63	6.76	-7.88	16.13
	(3.91)	(2.92)	(4.99)	(1.32)
Due to intrauterine growth restriction (%)	4.88	2.7	-2.18	4.23
	(2.38)	(1.89)	(3.1)	(0.72)
Due to anaemia (%)	2.44	4.05	1.62	11.01
	(1.7)	(2.29)	(2.84)	(1.12)
Due to gestational diabetes (%)	12.2	8.11	-4.09	7.68
	(3.62)	(3.17)	(4.88)	(0.95)
Any complication (%)	50	54.05	4.05	43.28
	(5.52)	(5.8)	(8.06)	(1.77)
Other characteristics				
Age of mother	22.61	23.65	1.039	23.96
	(0.6)	(0.64)	(0.88)	(0.22)
APGAR 5	8.98	8.81	-0.165	8.82
	(0.12)	(0.19)	(0.22)	(0.05)
Community characteristics[(a)]				
Community size	318.80	277.33	-41.47	485.50
	(68.85)	(58.75)	(96.73)	(55.65)
Female share of population (%)	49.48	52.81	3.33	0.51
	(1.11)	(1.29)	(2)	(0.01)
Median age	22	22.75	0.750	22.5116
	(0.73)	(0.82)	(1.27)	(0.33)
People per household	5.4216	5.95	0.528	6.057
	(0.39)	(0.22)	(0.63)	(0.21)
Median personal income	209.14	211.00	1.857	210.63
	(2.71)	(7.27)	(7.26)	(6.62)
Median rent payments	41.79	32.50	-9.292	44.21
	(6.75)	(8.39)	(12.11)	(4.2)
Labour force share of population (%)	39.86	22.75	-17.11**	39.21
	(4.35)	(4.76)	(7.51)	(2.46)

Notes: Standard errors in parentheses. APGAR 5 stands for Appearance, Pulse, Grimace, Activity and Respiration measured 5 minutes after birth. Each of the five categories is scored 0, 1 or 2, for a maximum total score of 10. Community characteristics are from the Australian Bureau of Statistics 2006 Census community profile data; most variables available for 54 NTER communities; rest of NTER is the average across communities that were not the first or last 10 to receive income management.

TABLE B.2: REGRESSION OF PRE-ROLLOUT COMMUNITY CHARACTERISTICS ON ROLLOUT TIMING

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Outcome: Days from first day of rollout to day community received income management				Outcome: Order in which community received income management (i.e. first, second, etc)			
Average birthweight (1 Jul 2006-30 Jun 2007)	-0.005 (0.016)	-0.004 (0.016)	0.002 (0.019)	-0.005 (0.016)	0.0002 (0.001)	0.0003 (0.001)	0.0006 (0.001)	0.0003 (0.002)
Average rainfall in region		0.294 (0.258)	0.342 (0.412)	0.295 (0.259)		0.0214 (0.0235)	0.0266 (0.0374)	0.0214 (0.0237)
Population size			-0.051 (0.053)				-0.0049 (0.0049)	
Female share of population			-273.7 (410.5)				-18.65 (37.21)	
Median age			5.437 (8.604)				0.564 (0.780)	
Share of population aged 65+			-1.198 (12.16)				-0.160 (1.102)	
Average household size			35.13* (18.35)				3.572** (1.664)	
Median personal income			0.486 (0.429)				0.0464 (0.0389)	
Labour force to population ratio			0.0996 (1.072)				0.0148 (0.0972)	
Missing Census data				16.25 (24.58)				0.954 (2.248)
Constant	218.6*** (50.48)	193.6*** (54.94)	-95.21 (399.5)	191.7*** (55.21)	14.42*** (4.595)	12.60** (5.016)	-20.03 (36.21)	12.49** (5.050)
Observations	80	80	49	80	80	80	49	80
R-squared	0.001	0.018	0.122	0.024	0.000	0.011	0.137	0.013

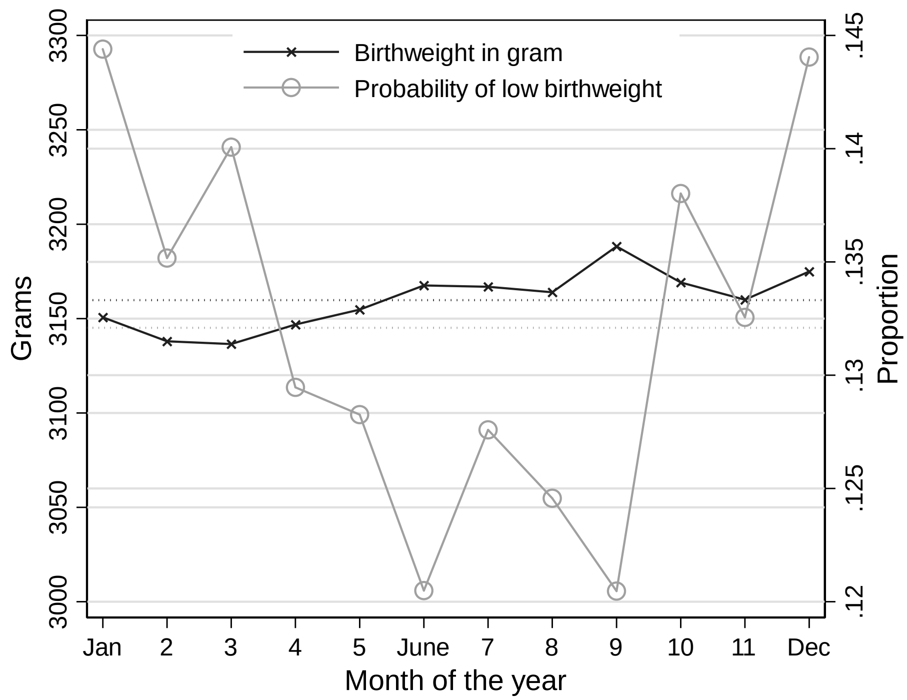

FIGURE B2. SEASONAL TRENDS IN BIRTHWEIGHT AND LOW BIRTH WEIGHT, 1996-2003

Notes: The graph reports the average birthweight (left vertical axis) and the probability of low birthweight (<2500 g) (right vertical axis) by month, averaged between 1996 and 2013. Data series are Winsorized at the 10% level to limit the influence of extreme outliers.

Appendix C: Robustness tests

Instead of exploiting the policy rollout timing, we can instead assess whether our results change when we compare treated newborns to those who most closely resemble them in terms of observable characteristics, using propensity score matching. Under various matching method approaches we find slightly larger effects than our benchmark model. We find an average reduction in birthweight of 116-194g (significant at the 5 percent level). The treatment effect on the probability of low birthweight varies between 4.8 and 8.6 percentage points (significant at the 5 or 10 percent level, Table C.1).

TABLE C.1 – PROPENSITY SCORE MATCHING – TREATMENT ON TREATED EFFECT[(A)]

	Sample size	Birthweight	Low birthweight
Matching on: year of birth, sex of baby, region, regional rainfall in 3 months to birth, mother's age, whether this is the mother's first pregnancy, and whether born in a major hospital.			
Matching method			
Nearest neighbour	1,984	-194.6 (79.53)	0.0860 (0.0346)
Radius of 0.1	1,984	-116.9 (48.96)	0.0540 (0.0242)
Kernel	1,984	-118.2 (50.94)	0.0494 (0.0246)
Stratified	1,984	-116.4 (53.61)	0.0478 (0.0273)

Notes: Bootstrapped standard errors in parentheses. (a) Treatment is defined as being in utero in a community covered by income management before the beginning of the third trimester of pregnancy.

We also test the effect of a range of changes in the way that we define the sample and the rollout period (Table C.2). Our results do not change meaningfully when we shift the rollout period earlier or later, define sample by birthdate (instead of date of conception), or drop small communities (with less than 10 births) during the rollout period from our analysis. Our results remain negative but are smaller in magnitude if we drop babies that were partially treated (that is, income management was introduced during their third trimester), or limit our sample to a heathy birthweight range (2500-5000g).

TABLE C.2 – TREATMENT EFFECT WITH DIFFERENT SAMPLE DEFINITIONS

	Sample size	Birthweight (grams)	Low birthweight
Main sample (conceived 17 Jan 2007 - 20 Apr 2008)	1,982	-94.89 (49.81)	0.030 (0.021)
Earlier sample (conceived 17 Dec 2006 to 30 Apr 2008)	2,057	-92.14 (50.50)	0.032 (0.021)
Later sample (Conceived 17 Jan 2007 to 30 May 2008)	2,074	-104.73 (50.41)	0.037 (0.020)
Sample defined by birthdate (not conception date)	2,065	-109.93 (49.44)	0.044 (0.021)
Excluding partial treatment from control group (main sample)	1,835	-67.29 (48.18)	0.018 (0.026)
Sample limited to healthy birthweight range (2500-5000g)	1,724	-71.48 (52.49)	na
Drop communities with <10 births	1,825	-93.72 (52.05)	0.035 (0.022)
Removing pre-rollout period from estimation	1,086	-102.16 (60.12)	0.030 (0.023)

Notes: Community-clustered standard errors in parentheses. Regressions include community and year fixed effects and controls for: premature birth, year of birth, rainfall, conceived in wet season, sex of baby, whether mother's first pregnancy, and mother's age. Excluding partial treatment drops the infants for whom income management was introduced during the third trimester (in main regression, these observations are included in the 'control' group).

Our results are broadly similar under a range of different methods for controlling for seasonal variation in birth outcomes. If we use rainfall data from the weather station closest to the community (instead of region-level rainfall data),[1] our estimate is very similar (Table C.3). If we use quarter-year fixed effects, year-season fixed effects, controls for the month the baby was conceived (instead of season), or measure rainfall in categories (5 categories based region-

[1] We do not do this in our main specification because rainfall data are not available for all communities, and some weather stations have missing data or data of unverified quality.

specific percentiles of 0-10, 10-40, 40-60, 60-90, 90-100, from on historical data) instead of as a continuous variable, we retain a negative treatment effect of around 100g or more (Table C.4).[2]

TABLE C.3 – IMPACT OF INCOME MANAGEMENT ON BIRTHWEIGHT AND PROBABILITY OF LOW BIRTHWEIGHT, CONTROLS FOR RAINFALL AT COMMUNITY'S CLOSEST WEATHER STATION

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Outcome: Birthweight (grams, OLS)				Outcome: Low birthweight (LPM)			
Income management	-10.41	-105.67	-123.11	-98.45	0.000	0.046	0.041	0.028
	(32.85)	(53.27)	(63.41)	(58.63)	(0.017)	(0.024)	(0.029)	(0.026)
Pre-rollout period	-55.89	32.62	19.55	-26.27	0.003	-0.054	-0.052	-0.026
	(31.38)	(70.28)	(75.66)	(69.02)	(0.018)	(0.038)	(0.041)	(0.033)
Rainfall in ml		-0.16	-0.04	-0.05		0.000	0.000	0.000
		(0.06)	(0.07)	(0.05)		(0.000)	(0.000)	(0.000)
Conceived in wet season		26.85	63.66	25.09		-0.032	-0.035	-0.013
		(39.97)	(40.80)	(34.87)		(0.022)	(0.022)	(0.022)
Sex (male = 0)		-77.44	-70.10	-70.01		0.035	0.033	0.032
		(32.87)	(35.32)	(29.77)		(0.018)	(0.019)	(0.015)
Mother's first pregnancy		-1.21	0.38	18.89		-0.011	-0.012	-0.023
		(32.93)	(34.33)	(27.26)		(0.022)	(0.022)	(0.014)
Mother's age (base cat. = under 20)								
20-24		64.98	60.54	67.29		-0.007	-0.002	-0.006
		(39.46)	(40.39)	(30.26)		(0.025)	(0.026)	(0.021)
25-29		170.31	186.28	196.30		-0.010	-0.011	-0.017
		(38.56)	(41.42)	(34.86)		(0.022)	(0.025)	(0.020)
30-34		279.99	284.37	282.27		-0.036	-0.034	-0.032
		(61.95)	(66.02)	(55.75)		(0.030)	(0.031)	(0.024)
35+		176.44	168.09	183.02		0.009	0.012	-0.001
		(76.55)	(79.88)	(60.93)		(0.045)	(0.047)	(0.034)
Premature				-930.50				0.552
				(39.33)				(0.035)
Constant	3,149.88	2,905.42	3,349.85	3,479.60	0.128	0.246	0.123	0.050
	(24.26)	(96.39)	(85.71)	(91.85)	(0.013)	(0.058)	(0.056)	(0.055)
Year fixed effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Community fixed effects	No	No	Yes	Yes	No	No	Yes	Yes
Number of communities	72	72	72	72	72	72	72	72
Observations	1,983	1,620	1,620	1,620	1,984	1,621	1,621	1,621
R-squared	0.00	0.04	0.09	0.35	0.000	0.015	0.061	0.374

Notes: This table uses R package 'Bomrang' to locate and download data from the Australian Bureau of Meteorology weather station closest to each individual community, based on its longitude and latitude coordinates. Some communities' weather stations have unreliable or missing data for all or part of the sample period, which is why there are fewer observations in columns 2-4. For this reason, our main analysis uses region-level, instead of community-level, weather controls.

[2] The key exception is the quarter-year fixed effects, where the treatment effect is 176g and 7.8ppts. As described earlier, this is likely due to overfitting.

TABLE C.4 – SENSITIVITY OF RESULTS TO DIFFERENT METHODS OF CONTROLLING FOR SEASONAL AND TIME TRENDS

	Birthweight (grams)	Low birthweight
Benchmark model	-94.89	0.030
	(49.81)	(0.021)
Without year controls	-62.88	0.015
	(38.73)	(0.017)
Quarter & year interacted	-89.63	0.016
	(63.74)	0.016
Year & season interacted	-61.57	0.016
	(52.01)	(0.021)
Without rainfall	-82.23	0.021
	(49.25)	(0.021)
Rainfall as categorical variable	-88.13	0.023
	(52.41)	(0.022)
Control for month of conception	-108.03	0.038
	(52.02)	(0.022)

Notes: Community-clustered standard errors in parentheses. Including community fixed effects and controls for: year of birth, rainfall, conceived in wet season, sex of baby, whether mother's first pregnancy, and mother's age.

Finally, as described in the main text, our estimation results are also not sensitive to inclusion of additional controls for maternal characteristics, maternal smoking and drinking behaviour (Table C.5). We do not include these controls in the main model as data are missing for some observations. Similarly, our results are not sensitive to inclusion of health care at birth (Table C.6).

TABLE C.5 – TREATMENT EFFECT ON BIRTHWEIGHT, CONTROLLING FOR MOTHER'S BEHAVIOUR AND CHARACTERISTICS

	Birthweight (grams)			Low birthweight		
	(1)	(2)	(3)	(4)	(5)	(6)
Income management	-93.12	-65.59	-99.64	0.029	0.015	0.033
	(49.78)	(52.32)	(49.49)	(0.021)	(0.026)	(0.021)
Aboriginal or Torres Strait Islander	-188.72	-141.11	-146.25	0.078	0.055	0.063
mother	(51.19)	(55.79)	(0.030)	(0.030)	(47.56)	(0.023)
Mother has pre-existing medical	20.51	32.89	25.75	-0.002	0.001	-0.004
conditions	(24.80)	(27.27)	(0.015)	(0.015)	(25.73)	(0.014)
<u>Controls</u>						
Smoking at first antenatal visit, drinking at first antenatal visit, smoking at 36 weeks, drinking at 36 weeks	No	Yes	Yes	No	Yes	Yes
Missing smoking and drinking data	No	No	Yes	No	No	Yes
Community & year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Standard controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,982	1,378	1,982	1,983	1,379	1,983
R-squared	0.37	0.36	0.38	0.394	0.361	0.398

Notes: Community-clustered standard errors in parentheses. Models are estimated using ordinary least squares and linear probability models.

TABLE C.6 – TREATMENT EFFECT ON BIRTHWEIGHT, CONTROLLING FOR QUALITY OF CARE

	(1)	(2)	(3)	(4)	(5)	(6)
		Birthweight (grams)			Low birthweight	
Income management	-95.96	-95.60	-96.15	0.030	0.031	0.031
	(49.42)	(49.55)	(49.71)	(0.021)	(0.021)	(0.021)
Main hospital	98.47	92.94	77.39	-0.034	-0.044	-0.044
	(42.68)	(42.66)	(42.66)	(0.020)	(0.020)	(0.020)
Emergency delivery		30.50	-29.78		0.051	0.051
		(34.40)	(47.46)		(0.016)	(0.020)
Spontaneous delivery			-75.57			-0.000
			(32.75)			(0.016)
Community & year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,982	1,982	1,982	1,983	1,983	1,983
R-squared	0.37	0.37	0.37	0.393	0.396	0.396

Notes: Models are estimated using ordinary least squares and linear probability models. Community-clustered standard errors in parentheses.