Journal of Health Economics 84 (2022) 102618

Contents lists available at ScienceDirect

Journal of Health Economics

journal homepage: www.elsevier.com/locate/jhealeco

Unintended consequences of welfare reform: Evidence from birthweight of Aboriginal children in Australia☆

Mary-Alice Doyle[a,*], Stefanie Schurer[b], Sven Silburn[c]

[a] London School of Economics and Political Science UK
[b] University of Sydney, Australia and Institute for the Study of Labour (IZA), Germany
[c] Menzies School of Health Research, Australia

ARTICLE INFO

JEL:
D04
I14
I38

Keywords:
Welfare reform
Aboriginal children
Birthweight
Income management
Unintended consequences

ABSTRACT

In 2007, Australia introduced its most radical welfare reform in recent history, targeting Aboriginal communities with the aim of protecting children from harm. The 'income management' policy forced Aboriginal welfare recipients to spend at least half of their government transfers on essentials (e.g. food, housing), and less on non-essentials (e.g. alcohol, tobacco). By exploiting its staggered rollout, we estimate the impact of in utero exposure to the policy rollout on birthweight. We find that exposure to the income management policy reduced average birthweight robustly by 85 g and increased the risk of low birth weight by 3 percentage points. This finding is not explained by behavioral change (fertility, maternal risk behavior, access to care), or survival probabilities of at-risk fetuses. More likely, a lack of policy implementation planning and infrastructure led to acute income insecurity and stress during the rollout period, exacerbating the existing health inequalities it sought to address.

☆ 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, 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), 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).

* Corresponding author.
E-mail address: m.s.doyle@lse.ac.uk (M.-A. Doyle).

https://doi.org/10.1016/j.jhealeco.2022.102618
Received 10 May 2021; Received in revised form 26 March 2022; Accepted 2 April 2022
Available online 7 April 2022
0167-6296/

1. Introduction

In 2007, the Australian Government embarked on its most radical and contentious welfare reform in recent history. The new policy, referred to as 'income management', restricted the way Aboriginal and Torres Strait Islander benefit recipients in the Northern Territory[1] could spend their entitlements. Most benefit recipients in Australia receive a cash transfer. Under income management, people living in Aboriginal communities who received government benefits – including unemployment, disability support, parenting payments and the pension – had half of those payments quarantined into a separate account, which was designated to be spent on priority needs and could not be withdrawn as cash. The policy's official objective was to improve the welfare of Aboriginal children living in remote communities. Gaps in social, economic and physical health between Aboriginal and non-Aboriginal people in Australia have persisted over decades, a phenomenon that many rich nations with Indigenous populations are struggling with (Mitrou et al., 2014; Cooke et al., 2007).

This policy was introduced with the goal of addressing what officials saw as the behavioral causes of persistent disadvantage in Aboriginal communities.[2] The Australian Prime Minister at the time emphasized "the parental responsibility that accompanies [parents'] right to welfare support" (Howard, 2007, 73), reflecting the policy's aim to reduce entrenched disadvantage by changing individual behavior. However, whether the policy change was successful in improving outcomes for children has not been rigorously tested.

In this paper, we are the first to evaluate the effectiveness of the initial rollout of income management on children's health outcomes. Specifically, we study birthweight, a particularly useful outcome when considering the effects of a policy that targeted budget priorities. Birthweight is the most accessible marker of newborns' health, making it an ideal outcome for comparing our findings with the international literature. It is also predictive of children's survival in the first year of life (Wilcox, 2001), a key health policy target for the Northern Territory's Aboriginal communities, where twice as many children are born with low birthweight than the Territory's average. Although birthweight markers have their limitations, they are suitable to pick up environmental and nutritional shocks that affect the mother during pregnancy (Almond et al., 2005). Birthweight variations at the lower end of the distribution also tell a good story about the likely long-term consequences. Low birthweight is associated in a meaningful way with adult health (Risnes et al., 2011), schooling attainment (Bharadwaj et al., 2018), and later-life income (Currie and Moretti, 2007; Bharadwaj et al., 2018).

Income management affected all benefit recipients within the targeted communities.[3] The communities that were targeted by the initial rollout of income management are not only remote, but exceptionally poor in economic terms. Most experience high levels of food insecurity, and nutrition is mainly based on energy-dense and nutrient-poor foods (Brimblecombe et al., 2010). Over half of the population experiences cash poverty (Markham and Biddle, 2018). For most Aboriginal women living in such remote areas, government transfers are their main source of income (Venn et al., 2020).

We analyze the impact of in utero exposure to income management. In the same way as Almond et al. (2011), we define a baby as 'treated' by income management if the policy was introduced in their community before the beginning of the third trimester of pregnancy: that means that we include babies for whom the policy was introduced before conception, or in the first or second trimesters of pregnancy as 'treated'. The comparison group is babies who are born before income management came into their community – and therefore the policy could not have affected their birthweight – or for whom it came in at the very end of pregnancy, meaning a very short period of in utero exposure. Our event study estimates break down the effects of these different periods of exposure.

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 et al., 2011). Conversely, if the spending restriction was not binding because households already spent more than half of their income on essentials, we would expect no change in consumption patterns or in birthweight. Data on households' spending and consumption at the time are not available. However, earlier survey data for the region suggest that most low-income households were already spending over half of their income on essentials. If this is also true for the targeted communities, we may expect the policy to have no discernible impact on birthweight.

Our analysis draws on administrative records on the universe of all births to women living in affected communities, from the Northern Territory Data Linkage Study (Silburn et al., 2018). To identify the impact of the policy, we exploit its gradual rollout that occurred over 13 months with 32 different implementation dates across clusters of communities. We apply a standard generalized

[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] This means that all families with newborn or children were affected by the policy, because of their entitlement to conditional and unconditional family benefits including an unconditional lump-sum 'baby bonus' and the means-tested single and partnered parenting payments. Furthermore, around one-third of adults in the affected communities were employed, suggesting that the remaining two-thirds were eligible to receive unemployment-related government transfer payments (AIHW, 2010). In contrast, in other countries which use welfare restriction policies, restrictions apply to a smaller 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).

difference-in-difference model, as well as a newly developed estimator which addresses potential biases because of multiple rollout dates (Borusyak et al., 2021). To deal with potential violations in the common trend assumption, we use methods suggested by Roth and Freyaldenhoven et al. (2019).

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 pre-existing variations in birthweight or community characteristics. Contrary to expectations, we find that the policy introduction reduced average birthweights of children who were exposed to income management in utero by 85 g. The adverse effects are present throughout the birthweight distribution and are larger at the lower end. Not surprisingly, the policy increased the risk of low birthweight by 3 percentage points. The policy was most detrimental the earlier the in-utero exposure.

These findings raise the question of what elements of the income management policy may have caused harm to babies in utero. One hypothesis is that the income management policy itself changed behavior in unintended ways. Another hypothesis is that the policy was not correctly implemented, causing disruption of benefit payments. To disentangle these mechanisms, we evaluate a range of explanations. We show that the negative treatment effect cannot be explained by recipients' behavioral changes in response to the policy, such as changes in fertility, maternal risky health behaviors, or access to perinatal care. There is also no evidence that income management improved chances of survival for at-risk fetuses, which may reduce average birthweight, but would nevertheless be a beneficial outcome.

This leaves us with a more qualitative assessment of why income management reduced birthweight. We find that during the implementation period, recipients who did not immediately comply with new rules for accessing their funds had their payments suspended – this affected up to one-third of all recipients. Furthermore, those who could access their funds had their spending restricted in ways that were not intended by the policy. The most acute of these issues were resolved within the first year of the policy's operation, but during the rollout period, it led to substantial income insecurity. Qualitative evidence from an early Government report supports this conclusion. Because of the lack of consultation with Aboriginal communities, recipients were confused about how to access existing benefits, and in some cases, this led to short-run food insecurity (Yu et al., 2008). Cobb-Clark et al. (2021), who evaluated the policy's impact on children's school attendance, suggested that the policy led to family discord and a temporary disruption to school attendance.

Much of the discussion around income management has focused on whether or not it 'works', with many commentators questioning the underlying assumption that households receiving government transfers were spending more than half of their income on alcohol, tobacco and gambling (Altman, 2016; Klein, 2016; Bielefeld, 2018). While this is an important question, our findings show that it is not the only determinant of the policy's impact, at least in the short-term. Even if the policy's theory of change was sound, its ability to meet its goals was hampered by the way it was implemented. In itself, any significant change to a long-standing program may cause disruption when it is first introduced, as recipients must adjust to a new way of accessing and managing their money. This disruption was compounded by the fact that much of the financial and logistical infrastructure required to administer the program did not exist when the policy rollout began. Furthermore, the top-down and racially-targeted imposition of the policy has made many recipients less receptive to it (Yu et al., 2008; Dalley, 2020).

Many countries use paternalistic approaches to welfare policy (Currie and Gahvari, 2008). Australia's experience with income management demonstrates that the way that such policies are implemented – as distinct from the details of the policies themselves – can drive impact, leading to unintended consequences. Indeed, implementation issues are particularly likely with paternalistic policies, because they are more complex and costly to administer (Margolies and Hoddinott, 2015), creating more opportunity for the policy 'as implemented' to deviate from the policy 'as designed'.

Our findings therefore contribute to an international literature that attributes the quality of implementation of a policy as the defining factor in whether it achieves its intended goal (Durlak and DuPre, 2008; Cerna, 2013). Examples in the literature are abound. For instance, the Learnfare initiative that linked families' welfare payments to the school attendance of their teenage children, failed in part due to challenges administering it (Ethridge and Percy, 1993). Program outcomes can vary with how financial incentives (or sanctions) are structured (Dee, 2011), the coherence of the underlying statutes (Meier and McFarlane, 1995), and the way that parents are engaged with a program (Gennetian et al., 2016).

Independent of the reasons for which the income management policy harmed exposed babies in utero, our findings suggest that at least one birth cohort of babies were born with lower birthweights and higher risks of long-term health and developmental problems. Our findings on a poorly implemented policy therefore also contribute to a body of research which is concerned with the adverse effects of in utero exposure to inadequate income, risky health behaviors, stress and environmental pollution (see Aizer and Currie, 2014). Although previous studies have shown that restricted-use transfers can prevent or reduce the health effects of early-life shocks (Currie and Rossin-Slater, 2015), we demonstrate the potential for changes in transfer policies – particularly when they are implemented with a very short transition period – to have unintended consequences on early-life health outcomes. While we are unable to say whether these effects persisted for future birth cohorts, it is likely that there were long-term impacts on human capital for the cohort of children who born during or shortly after the policy introduction. It is well understood that poor early-life health outcomes can have long-term consequences (Almond et al., 2018). 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.

2. How might the income management policy affect children?

2.1. Policy background

The Northern Territory (NT) is a large and sparsely populated geographic area, covering approximately one-sixth of the Australian continent. Although nation-wide Aboriginal peoples make up only 3% 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 peoples – especially those who live in remote communities – experience substantial health disparities relative to non-Aboriginal people. 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 low birthweight, 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. (2021) 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 health and safety of children in remote Aboriginal communities.[4] The central policy was income management. The NTER also included a range of other policies such as alcohol and pornography bans, the withdrawal of an employment program in remote communities (the Community Development Employment Projects), 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.[5] 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.[6] Since 2010, income management has been rolled out universally in the Northern Territory, which allowed the *Racial Discrimination Act* to be reinstated. Income management's successor, the 'Cashless Debit Card', has been introduced in additional locations around Australia. Yet, little empirical evidence exists on whether these policies have been effective in their goals of protecting children from harm or improving their welfare.

2.2. Institutional set up of the income management policy

Income management imposed restrictions on what benefit recipients could do with their payments. Before income management, recipients had 100% of their payments deposited into their bank account to spend however they choose. Under income management, the total value of payments did not change, but recipients were paid only half of their transfer into their bank account as usual. The other half of each regular payment was set aside into an income management account held with the government welfare agency. It could only be spent on priority needs such as food, housing, bills and clothing.[7] It could not be withdrawn as cash and could not be spent on alcohol, cigarettes, pornography or gambling.

Participants 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 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 community residents. Statistics from a 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% being income managed at a point in time after the rollout was complete.[8] Women were more likely to be income managed than men

---

[4] 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.

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

[6] 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).

[7] The full list of 'priority needs' are: food; non-alcoholic beverages, clothing, footwear, basic personal hygiene items, basic household items, housing, household utilities, rates and land tax, health, child care and development, education and training, items required for employment, funerals, public transport and private vehicles. These items are listed in the *Social Security and Other Legislation Amendment (Welfare Payment Reform) Act 2007*.

[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.

Table 1
Summary statistics in year before income management (1 July 2006–30 June 2007).

	NTER communities			Rest of NT
	Communities in first half of rollout	Communities in second half of rollout	Diff.	
Outcome variables				
Birthweight (grams)	3023	3037	13.88	3334
	(33)	(32)	(46.07)	(12)
Low birthweight (%)	16.05	15.22	-0.53	7.53
	(1.67)	(1.63)	(2.33)	(0.5)
Obstetric complications				
Premature (%)	17.08	16.53	-0.55	8.64
	(1.71)	(1.67)	(2.39)	(0.53)
Intrauterine growth restriction (%)	4.94	3.23	-1.71	1.37
	(0.98)	(0.79)	(1.3)	(0.22)
Anemia (%)	9.26	9.48	0.217	1.98
	(1.32)	(1.32)	(1.86)	(0.26)
Gestational diabetes (%)	7.00	8.67	1.67	6.81
	(1.16)	(1.26)	(1.72)	(0.48)
Any obstetric complication (%)	43.00	45.36	2.4	24.56
	(2.25)	(2.24)	(3.17)	(0.82)
Other characteristics				
Average age of mother	23.80	23.92	-0.13	28.57
	(0.27)	(0.28)	(0.4)	(0.12)
Average Apgar 5 score	8.67	8.77	0.10	8.92
	(0.08)	(0.07)	(0.11)	(0.02)
Community characteristics				
Community population	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. The third column tests for differences between the first and second half averages. There are 80 communities (including town camps and outstations) with any births in the year before the policy was introduced, and 86 communities with births in the full analysis sample period. Apgar 5 stands for Appearance, Pulse, Grimace, Activity and Respiration measured 5 min after birth. Each of the five categories is scored 0, 1 or 2, for a maximum total score of 10. Community characteristics come from the 2006 Census and are available for 54 communities; characteristics for 'Rest of NT' represent the NT average. $^{*}p < 0.10$. $^{**} p < 0.05$. $^{***} p < 0.01$.

(60 versus 40% of participants). Families with children were all affected by the policy because they were entitled to unconditional cash transfers tied to the birth of a child, as well as means-tested parenting payments.

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 (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 649 out of 21,763 clients who were ever income managed during the rollout period (3%) were granted exemptions (AIHW, 2010).

The institutional set up of income management made it a costly policy. Recent estimates suggest that to administer the policy, it cost around A$9000 per participant per year. This amount excludes the value of the payment itself (Department of Social Services, 2017).

2.3. How might the policy affect consumption?

The policy intended to affect consumption choices by limiting how much money households could spend on excluded goods using their welfare entitlements. To have the intended impact, it assumed that households had previously spent over half of their benefits on excluded goods. For households already spending most of their benefits on essentials, the policy would not have been expected to affect consumption (Southworth, 1945) or change measurable outcomes (see Appendix Fig. A.1).

Based on this reasoning, for the policy to have its intended impact, it requires that a non-negligible share of households do not already spend most of their benefits on essentials; that is, that there are a large number of ‘extramarginal’ households.

Given the lack of data from the affected communities, we do not know how many households were extramarginal pre-policy. However, earlier survey data covering low-income households in the NT shows that a minority would have been extramarginal, and thus may have had their spending on non-essentials curtailed by the policy as intended (see discussion in Appendix B).

Conversely, the policy could have affected consumption in other ways, if it led to changes in income and purchasing power (see Cobb-Clark et al., 2021 for a more detailed discussion). We revisit the scope for such potential unintended effects in Section 7.

2.4. *Will changes in consumption affect birthweight?*

If the policy increased food consumption of participants and in particular pregnant women, then it is likely that the policy also affected birthweight of the children who were in utero at the time of the policy. There is ample empirical evidence that increased food consumption increases birthweight of babies by boosting intrauterine growth (see Gresham et al., 2014).[9] Previous research has also shown that transfer programs can improve nutrition in pregnancy (Barber and Gertler, 2008; Almond et al., 2011; Hoynes et al., 2011).

Based on this evidence, we analyze in utero exposure to income management. We expect that if income management increased maternal 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, 65% of quarantined funds were spent on food during the rollout period (AIHW, 2010).

2.5. *Empirical evidence on the effect of the income management policy*

There is little existing evidence on the effectiveness of income management in changing consumption patterns. Brimblecombe et al. (2010) study the effect of the policy on consumption in ten local stores in the North-East of the NT. They find a significant drop in the sales of soft drinks and a non-significant drop in sales of tobacco in the first half year following the introduction of income management, but after one year, sales in both categories returned to pre-policy levels. The authors conclude that the policy had no beneficial effect on priority good sales. Lamb and Young (2011) study the amounts spent on gambling in two towns located in the center of the NT (Katherine and Alice Springs) after the introduction of income management. They find a drop in electronic poker machine revenues in one gambling venue in each town, suggesting that the policy may have had a small effect in reducing formal gambling expenditure.

Both studies relate to specific locations in the NT, so their findings may not be representative of the effect of the policy across the whole of the NT.[10] Therefore, it is unclear whether these same conclusions – of no change in spending on groceries but a reduction in gambling – would apply more generally.

Apart from these two academic studies, there is one government report that evaluated the introduction of income management (AIHW, 2010). The report concludes that the policy led to increased spending on essentials. Although informative, these conclusions rely mainly on surveys or focus groups with small, non-random samples of community residents and staff involved in administering the program, making it non-representative of the target population. No baseline data were collected that would have allowed for a before-and-after comparison.

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 impact on children’s health outcomes – a key policy objective.

3. Data and definitions

We use 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 all 74,425 children born in the Northern Territory between 1994 and 2013. For this study, we have linked in daily, location-specific rainfall and temperature data from NASA and 2006 Census community profile data from the Australian Bureau of Statistics.

[9] The meta-analysis by Gresham et al. (2014) shows that randomized trials that provide food or fortified food during pregnancy increase birthweights by 125 g, on average across the studies included.

[10] 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.

3.1. Definition of treatment

Income management was rolled out in stages in all 73 remote Aboriginal communities (and associated satellite communities – known as '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, 86 of which had at least one birth in the sample period.[11] 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.[12] 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 AIHW, 2010).

We define a child as being treated if income management was introduced in their community before the start of the third trimester in utero – that is, up to 28 weeks after the estimated date of conception. This means that if income management was introduced in the mother's community before conception, in the first trimester or in the second trimester, the baby is defined as 'treated'. If income management was introduced during the third trimester or after birth, the baby is assigned to the comparison group. Almond et al. (2011) use this same treatment definition in analyzing the effect of in utero exposure to food stamps in the US, finding an increase in birthweight if the policy was in place for the full third trimester or longer. To test the appropriateness of this definition, we show an event study with treatment effects by length of in-utero exposure. Additional robustness checks present estimation results in which we exclude babies exposed during the third trimester from the analysis.

3.2. 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 take the cohort of babies born from the first day of the rollout period (17 September 2007) until 90 days after the final community received income management (25 January 2009). This gives a sample during the rollout period of 1,187 babies.[13] However, to aid reliable estimation of community fixed effects and time trends, we include five and a half years of pre-rollout period data (born 1 January 2001–16 September 2007), and a short post-rollout period (born 26 January 2009 to 31 December 2009).

3.3. Outcome variables

The outcome variables of interest are birthweight and the probability of low birthweight. Low birthweight is common in NTER communities – with around 14% of infants born with low birthweight in the year before income management was introduced – compared with 7% in other parts of the NT.

We focus on birthweight for two reasons. First, as described in Section 2.4, improved maternal nutrition during pregnancy can increase birthweight. This means that 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 et al., 2018 for an overview).

To reduce the potential effect of outliers, we top- and bottom-code values of birthweight above the 97.5th and below the 2.5th percentiles, respectively.

In our analysis, we also show a range of secondary outcomes, including preterm birth, small for gestational age, Apgar scores, various obstetric complications, and admission to special care nursery. These outcomes provide a fuller picture of the policy's impact, to assess whether income management affected birthweight alone, or whether it impacted infant health in other ways.

4. Empirical framework

To identify the causal impact of the introduction of income management, we exploit its staggered rollout. Our approach is to compare birthweight for babies that were not yet born when income management was introduced in their communities to babies who were already born by the time income management came to their communities. Our analysis is limited to babies born in communities where income management was eventually introduced.

As shown in Fig. 1, the policy was introduced over a period of 13 months with 32 distinct rollout dates. 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 exogenous, and that before the rollout, there were parallel trends in – and levels of – birthweight between the earlier- and later-adopting communities.

[11] 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.

[12] Some communities are known by many different names. 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.

[13] As a robustness test, we alternatively define our sample based on date of conception instead of date of birth to avoid the 'fixed cohort bias' (Strand et al., 2011).

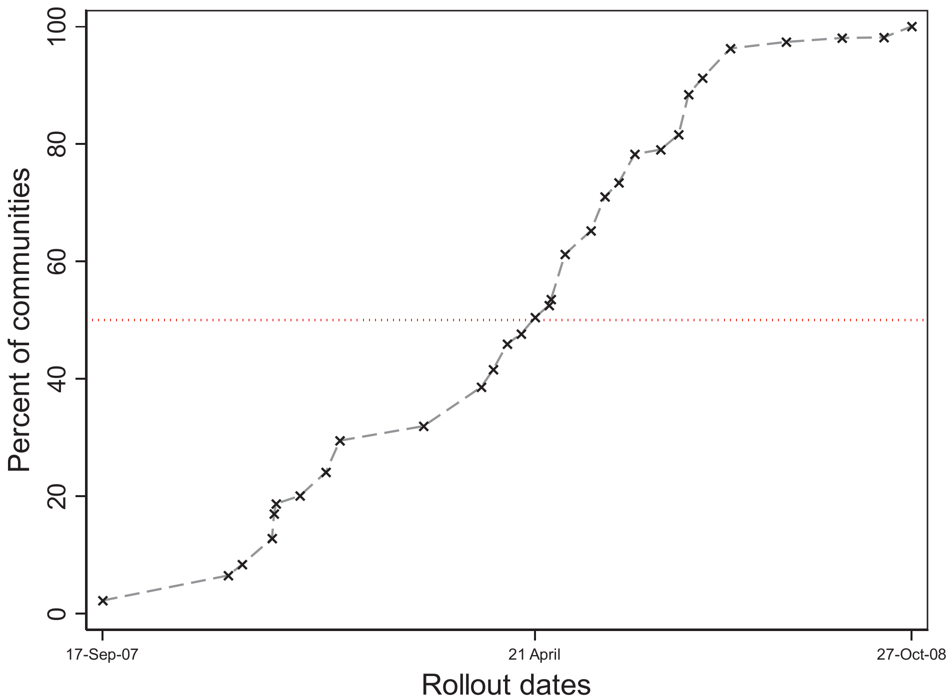


Fig. 1. Rollout of the Income Management Policy. *Notes*: The graph shows the cumulative share of NTER communities that were covered by income management on each of the implementation dates (indicated by crosses). Data are weighted by number of births in each community. For full details on the rollout schedule, see AIHW (2010).

4.1. Exogenous timing of the rollout

Income management was rolled out on a pre-defined timeline.[14] However, the rollout did not follow any clear geographic pattern (Fig. 2). 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. The geographic pattern of the rollout makes clear that the timeline was not following any systematic logistical pattern (e.g. targeting communities closest to urban areas first).

We also see no pre-policy differences in birthweight, birth complications, or community characteristics – this suggests that the rollout schedule did not, for example, target the most deprived or the largest communities first. Table 1 shows this by reporting means for early- and late-adopting communities in the year prior to rollout, as well as testing for differences between those means. It uses 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.[15]

Most birth outcome measures, including obstetric complications, characteristics of the mother and Apgar scores 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 households 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.

As an empirical test for pre-existing differences between communities, we run a regression to predict policy implementation timing from community characteristics and previous birthweight. It confirms that these variables are not predictive; birthweight explains less than 0.1% of the variation in rollout timing, while community characteristics explain up to 14% (Appendix Table B.2).[16]

---

[14] 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, but a robustness test (Appendix Tabel D.5) controls for consultation periods pre-rollout, accounting for the fact that these were extended due to delayed implementation in some communities.

[15] Table B1 tests 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.

[16] 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

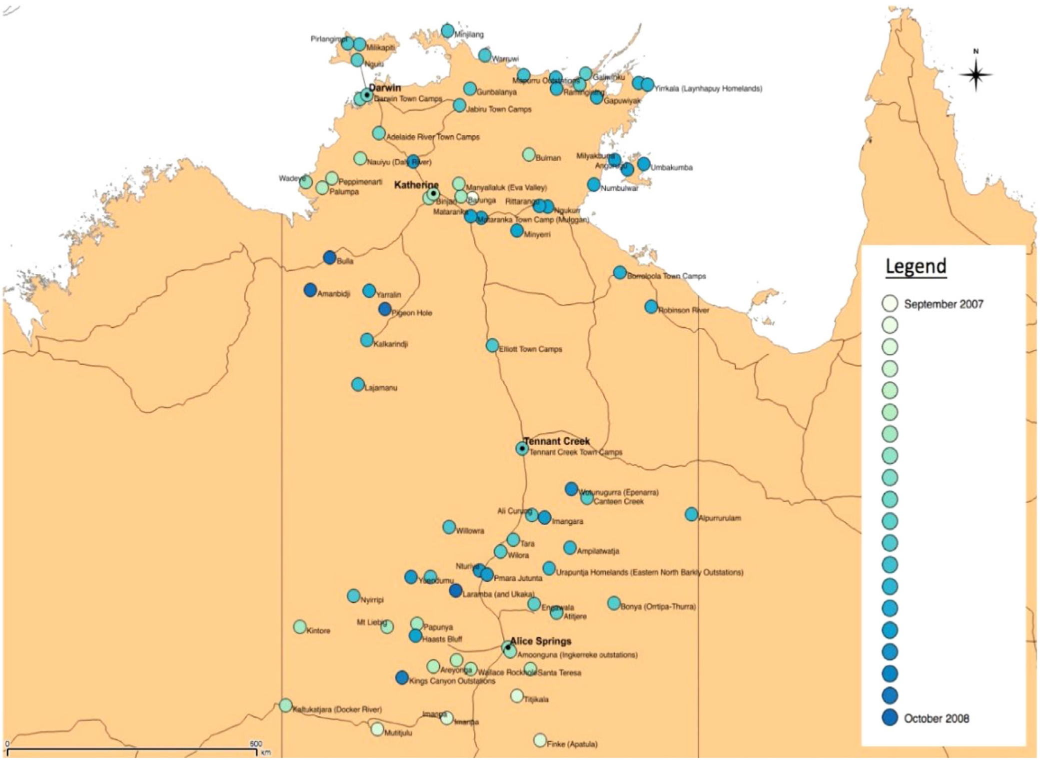


Fig. 2. Geographic pattern of the income management rollout. *Notes:* Color coding reflects the date income management started in the relevant community, as indicated in the Legend. Major cities or towns in the Northern Territory are displayed for reference. People living in the municipal parts of those towns were not subject to IM, but people living in associated town camps were. Lines connecting communities represent major roads.

Finally, although income management was one of a range of policies implemented as part of the NTER, we are able to separately identify its effects because other NTER policies were each rolled out on different timelines and with communities receiving each policy in a different order (see Cobb-Clark et al., 2021 for a detailed discussion). As shown in Appendix Table A.1, some NTER policies were introduced in all communities at the same time and before income management commenced (e.g. alcohol and pornography bans, and community and housing clean-up and repairs). Other policies were introduced on a much slower timeline (e.g. increasing police presence, and store licensing, which included ensuring community stores sold nutritious food).

To address the possibility of confounding, we conduct robustness checks to our main estimation model, where we include control variables for the second-largest component of the NTER – the cancellation of the Community Development Employment Projects (CDEP).[17] Our estimates are robust to this inclusion (see Appendix Table D.5).

4.2. Parallel trends before income management

In the year before the rollout, both the level and trend in birthweight in communities that received income management early were no different from those that received it later. Constructed from our administrative data, Fig. 3 suggests no apparent trend in birthweight in either group prior to the rollout.[18] We will show more formally in an event study that the common trend assumption is valid in our preferred model (see Section 5.2).

4.3. Crossover between early- and late-adopting communities

Some residents in early-adopting communities may have wished to move to late-adopting communities to delay participating in income management. If this occurred, it may bias our estimates. But the scope for this was very limited, as eligibility was deter-

control variables in this regression explain up to 14 percent 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.

[17] The CDEP was seen to be important in providing employment opportunities in remote communities (Hunter, 2009). Its cancellation was controversial, and thus the cancellation was halted in December 2007, after 23 communities were affected.

[18] 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.

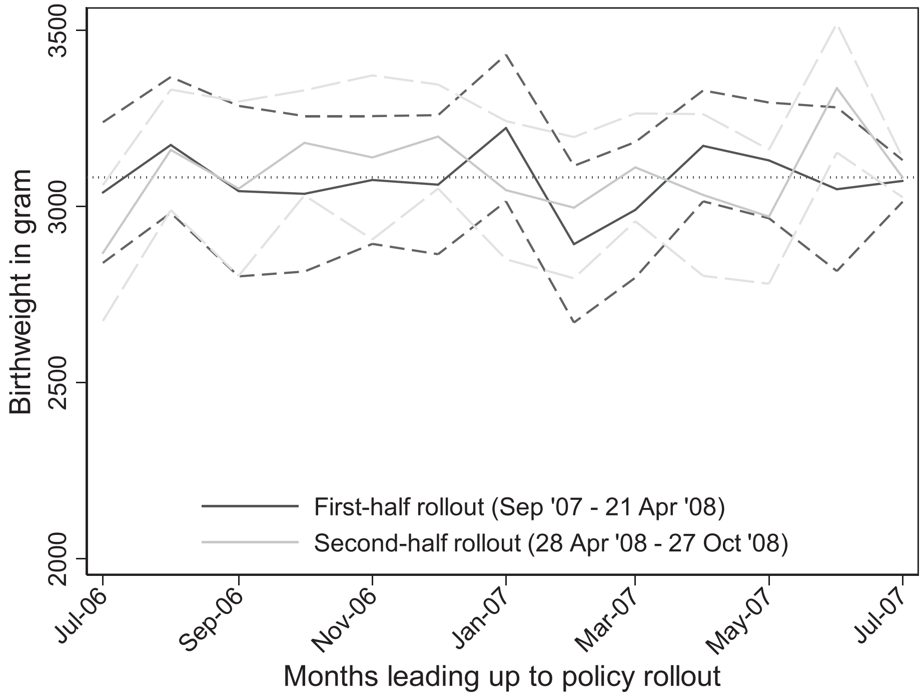


Fig. 3. Birthweight trends by timing of the rollout. *Notes*: The graphs display the unadjusted mean birthweight in NTER communities in each month in the year before income management was introduced. Trends are shown, separately for communities that received IM in the first half of the rollout (weighted by the number of births, those that received it from the beginning of the rollout in September 2007 to 21 April 2008) and the second half of the rollout (those that received it after 28 April). Dashed lines indicated 95% confidence intervals. No communities received IM between 21 and 28 April 2008.

mined based on place of residence one week after the policy was announced. Cobb-Clark et al. (2021) show empirically that income management did not impact short-term mobility.

4.4. Estimation model

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

$$Y_{itc} = \alpha + \tau P_t + \delta IM_{itc} + \pi P_t \times IM_{itc} + \gamma X_{ic} + \beta\theta_t + \rho S_{tr} + \eta_c + \varepsilon_{itc,} \tag{1}$$

where η_c denotes community fixed effects, θ_t represents a linear time trend measured in quarters and S_{tr} captures controls for region-specific seasonal patterns (see Appendix Fig. B1).[19] There are strong seasonal patterns in the NT, with birth outcomes being worse for babies born in the summer/wet season. This is partly related to extreme heat exposure in the summer/wet season and seasonal disease prevalence, and partly to road flooding, which limits access in and out of some remote communities, sometimes for months at a time. We capture these seasonal patterns by including controls for rainfall in the three months to birth and its square, and for the number of days during pregnancy with maximum temperatures above 30 °C. X_{ic} is a set of individual-level control variables: the sex of the baby, whether the baby is Aboriginal or Torres Strait Islander, the mother's number of previous pregnancies carried to 20 weeks or more, and the mother's age (categorized into 5-year age groups). Standard errors are clustered at the community level.

P_t is the rollout period indicator, which equals 0 during the rollout period, and 1 for the pre-rollout period (where no births are treated), and 2 for the post-rollout period (where all births are treated). 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. We interact the rollout period and the reform indicator $P_t \times IM_{itc}$. By defining the rollout period as $P_t = 0$, the coefficient δ on IM_{itc} gives the treatment effect of main interest. This is what we report in our estimation results as the treatment effect.

[19] 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, and the relatively short policy rollout period. Instead, we include seasonal controls. The advantage of our method of controlling for seasonality is that it allows for the timing of these seasonal effects to vary from year to year. In our robustness tests, we show that results are robust to different approaches to controlling for time and seasonal variation (e.g. using year-quarter fixed effects, controlling for month of conception instead of season of conception, and inclusion of additional seasonal control variables).

Table 2
Main estimation results.

	(1)	(2)	(3)	(4)	(5)
Panel A: Birthweight (pre-treatment mean: 3061g)					
Treatment effect	-109.544*** (39.053)	-101.808*** (35.439)	-114.727*** (35.274)	-118.631*** (35.070)	-115.406*** (34.566)
Observations	8604	8604	8604	8604	8604
Panel B: Probability of low birthweight (pre-treatment mean: 0.148)					
Treatment effect	0.047** (0.019)	0.044** (0.018)	0.044** (0.018)	0.045** (0.019)	0.044** (0.018)
Observations	8604	8604	8604	8604	8604
Time period controls	√	√	√	√	√
Basic controls	✗	√	√	√	√
Community FE	✗	✗	√	√	√
Weather controls	✗	✗	✗	√	√
Linear time trend	✗	✗	✗	✗	√

Notes: Difference-in-difference model, where the outcome variable is birthweight in grams (Panel A) or the probability of low birthweight (Panel B) as specified in Eq. (1). Treatment is identified as income management being introduced in the mother's community of residence before the start of the third trimester of the pregnancy. The treatment indicator is interacted with dummy variables that indicate pre-rollout (before 17 September 2007) and post-rollout (after 25 January 2009) period; the treatment effect shown is the impact of the policy during the rollout period. Basic controls are: newborn is a boy, baby is identified as Aboriginal and Torres Strait Islander, age categories of mother (<20 versus 20–24, 25–29, 30–34, 35+), dummy variables for number of previous pregnancies (0 versus 1, 2, …, 5+). Weather controls are number of days during pregnancy with maximum temperatures above 30 °C, and a quadratic polynomial of total rainfall (in ml) in the three months to birth in the mother's place of residence. Fixed effects are for the 86 NTER communities (including town camps and outstations. Time trends: Linear time trend proxied by a continuous measure of quarter by year of birth. Cluster robust standard errors by NTER community in parentheses. The model is estimated on all births that occurred between 2001 and 2009 inclusive. $^{*}p < 0.10$. ** $p < 0.05$. *** $p < 0.01$.

Our treatment estimate of δ captures intention-to-treat (ITT) effects. We do not observe which mothers received government transfers (and thus were directly affected by income management). But with the intervention affecting three-quarters of the population (AIHW, 2010) and all families with a newborn child, 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. However, our findings are robust to applying more recently developed methods of estimating two-way fixed effects models that address potential biases caused by multiple rollout dates (see Section 5.3).

5. Estimation results

5.1. Did the income management policy affect birthweight?

Table 2 reports our main estimation results. In a model without control variables (Panel A, column 1), we find that average birthweight is around 110 (S.E. 39 g) grams lower in the treatment group than in the untreated group, and the probability of low birthweight is 4.7 percentage points (S.E. 1.9) higher (Panel B, column 1).

The treatment effects remain relatively stable, with a reduction of 102–119 g, and an increase in the probability of low birthweight of around $4\frac{1}{2}$ percentage points when we add basic control variables (column 2), control for community fixed effects (column 3), and controls for weather conditions during pregnancy (column 4). Our preferred model – which also controls for a linear time trend – suggests similar effects. It shows that income management reduced birthweight by 115 g (S.E. 35) and increased probability of low birthweight by 4.4 percentage points (S.E. 1.8). These estimates are statistically significant at the 1% and 5% levels, respectively.

A reduction in birthweight can be a symptom either of reduced nutrition (leading to slower intrauterine growth), or of a worsening in infant health for other reasons that may lead to shorter gestational age or preterm birth. Table 3 presents the impact of income management on a range of additional measures of infant health. It suggests that both gestational age and intrauterine growth were reduced.

Panel A shows a statistically significant reduction in gestational age by around half of one week, and an increase in preterm birth by 3 percentage points (though this is not statistically significant). There are no observable changes in other measures (e.g. Apgar 5 scores or obstetric complications). The probability of admission to a special care nursery increased by 5 percentage points, which is consistent with the increased probability of low birthweight, given that low birthweight is a criteria for admission.

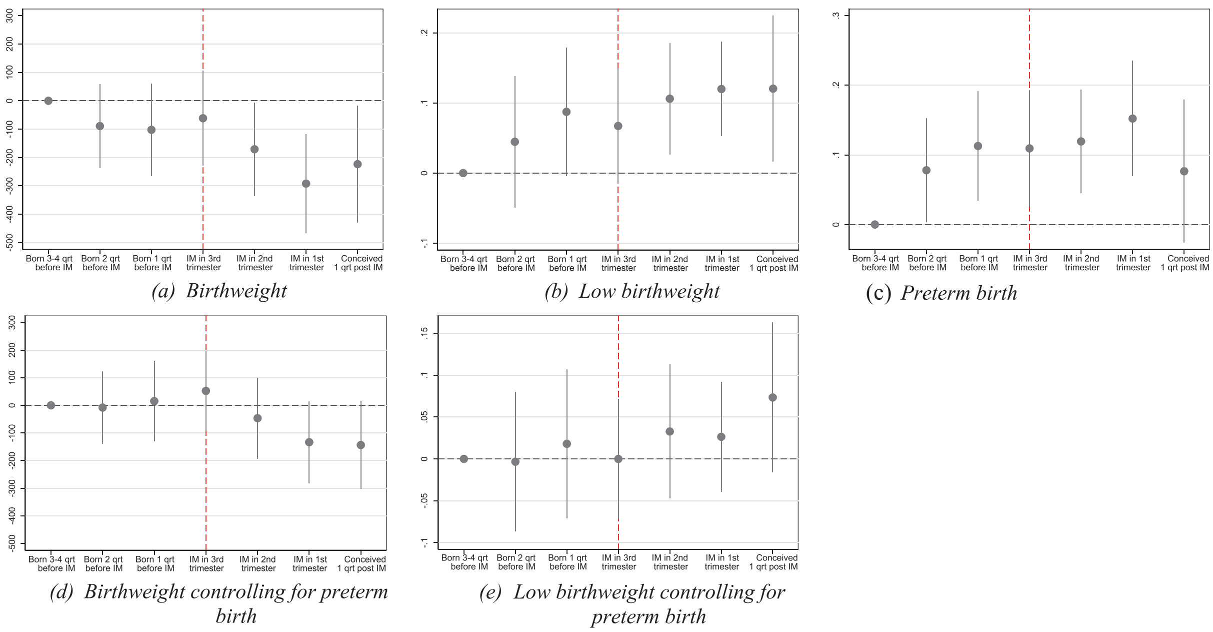
5.2. Event study approach and model refinement

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. Fig. 4 shows five panels; Fig. 4(a) and (b) show the model described above, with our

Table 3
Impact of income management on other indicators of infant health.

Panel A: Other birth outcomes	Preterm birth	Gestational age	Small for gestational age	Apgar 5	Admitted to special care nursery
Treatment effect	0.03	-0.52***	0.03	-0.10	0.05**
	(0.02)	(0.17)	(0.02)	(0.09)	(0.02)
Pre-treatment mean	0.16	38.11	0.20	8.74	0.24
	8604	8604	8604	8584	8526
Panel B: Indicators of obstetric complications	Intrauterine growth restriction	Gestational diabetes	Pre-eclampsia	Anemia	
Treatment effect	0.01	-0.03*	0.00	0.01	
	(0.01)	(0.02)	(0.01)	(0.02)	
Pre-treatment mean	0.05	0.06	0.06	0.07	
N	8604	8604	8604	8604	

Notes: Difference-in-differences model as specified in Eq. (1), where the outcomes is one of many other birth outcomes (Panel A) or one of many obstetric complications (Panel B). Preterm birth is defined as birth before 37 complete weeks of pregnancy. Gestational age is defined as the number of weeks from the estimated date of conception to birth. Small for gestational age is defined as birthweight below the 10th percentile of the national distribution (Dobbins et al., 2012), given the baby's gestational age and sex. Apgar 5 is a summary measure of the baby's appearance, heart rate, reflexes, muscle tone and respiration, taken by the birth attendant 5 min after birth. Clustered Standard errors by NTER community in parentheses. The model is estimated on all births that occurred between 2001 and 2009 inclusive. *$p < 0.10$. ** $p < 0.05$. *** $p < 0.01$.



(a) *Birthweight*

(b) *Low birthweight*

(c) *Preterm birth*

(d) *Birthweight controlling for preterm birth*

(e) *Low birthweight controlling for preterm birth*

Fig. 4. Event study models. *Notes:* These figures show 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/trimesters. The base category 'Born 3-4 qtrs before IM' includes babies born 180–360 days before income management was introduced in their community. We follow Borusyak et al. (2021) to include two periods in the base category to avoid under-identification. We use the same controls as the benchmark model, and standard errors are clustered by community.

two main outcomes. In Fig. 4(a), we see no pre-treatment trends in birthweight, and that newborns who were exposed to income management for longer (i.e. during the first trimester or pre-conception) experienced the largest reductions in birthweight.

However, Fig. 4(b) suggests that there was a pre-treatment trend in low birthweight. The probability of low birthweight appears to increase for babies that were born before income management was introduced. This introduces a puzzle – why was there a pre-intervention increase in low birthweight?

Some contextual details provide a possible explanation. First, as described in Section 4.4, there is substantial seasonal variation in newborns' health in the NT, with birth outcomes being worse for babies born in the summer/wet season. It is possible that some of the same factors that produce these poorer birth outcomes (e.g. heavy rainfall leading to road flooding, restricting access to routine medical care and fresh food) may also have been considerations in the timing of the policy introduction. If so, the rollout in those communities for which road access would likely be blocked by flooding may have been scheduled to begin after the wet season. Following this logic, birth outcomes would have been worse than average in those communities in the quarters before the policy

Table 4
Final model specification and tests for parallel trends.

	Birthweight	Birthweight	Low birthweight	Low birthweight
Panel A: OLS regressions				
Treatment effect	-115.41***	-85.06**	0.04**	0.03*
	(34.57)	(33.66)	(0.02)	(0.01)
Observations	8604	8604	8604	8604
Control for preterm birth	✗	√	✗	√
Panel B: Bousyak et al. (2021) estimator				
Treatment effect	-118.81***	-88.82***	0.04***	0.03***
	(19.30)	(17.36)	(0.01)	(0.01)
Observations	7714	7714	7714	7714
Control for preterm birth	✗	√	✗	√
Panel C: Placebo test				
Placebo treatment effect	-46.28	-11.65	0.02	0.00
	(35.57)	(32.97)	(0.02)	(0.02)
Observations	7646	7646	7646	7646
Control for preterm birth	✗	√	✗	√
Panel D: Pre-trend test				
F-statistic	0.64	0.36	1.46	0.14
P-value	0.59	0.79	0.23	0.93
Control for preterm birth	✗	√	✗	√

Notes: Difference-in-differences model as specified in Eq. (1), shown with and without including controls for preterm birth (Pane A). Panel B uses Borusyak, Jaravel and Spiess' imputation estimator (Stata command: did_imputation). Panel C shows a one-year placebo treatment effect (i.e. with the placebo-policy rollout beginning in September 2006 and ending in October 2007, instead of the actual dates of September 2007–Occtober 2008). Panel D shows results from an *F*-test of joint significance of the coefficients on the event study estimates for income management being introduced during the third trimester, or in the four quarters after birth. The models are estimated on all births that occurred between 2001 and 2009 inclusive. * $p < 0.10$. ** $p < 0.05$. *** $p < 0.01$.

introduction. We have no way of confirming whether this was a consideration in the timing of the policy rollout, but we judge it likely given that the same roads tend to flood each year.

Recent econometric literature suggests that pre-trends can be addressed using such context-specific information. Roth (2021) suggests using relevant parametric controls to address pre-trends, and Freyaldenhoven et al. (2019) suggest including covariates that may confound the outcome measure. In our case, preterm birth is such a candidate, given the high rates of preterm birth in the wet season/summer (see Appendix Fig. B.1), and the fact that we see a similar pre-trend in both low birthweight (Fig. 4(b)) and preterm birth (Fig. 4(c)). Preterm birth appears to drive the pattern we see in low birthweight, with an increase for babies born in the 6 months before the policy introduction, but no further increase for those born after the policy introduction.

Preterm birth also works in this case as a proxy for unexplained seasonality in birth outcomes. In the NT, seasonal patterns vary substantially from year to year, as the timing of the seasons can vary. For example, in 2005 rainfall in Katherine (one of the major towns in the NT) peaked in December, whereas in 2007, rainfall peaked in March. While our inclusion of covariates for rainfall and heat exposure during pregnancy work reasonably well for apparent seasonal variation in birthweight, they do not control for apparent seasonal variation in preterm birth (Fig. 4(c)). We therefore control for preterm birth directly in Fig. 4(d) and (e).

We know that low birthweight can be caused either by earlier delivery or by restricted intrauterine growth. In this analysis, we set out to test the impact of income management on birthweight, based on the hypothesis that income management would impact food consumption through the intrauterine growth channel. Therefore, we can include preterm birth as a covariate in our model, given that it appears to drive the pre-trend in low birthweight, and its inclusion as a control helps us to isolate the policy's hypothesized impact on intrauterine growth.

We find that when we include preterm birth as a covariate, the pre-intervention trend in low birthweight disappears (Fig. 4(e)). This confirms our suspicion that there were likely other factors unrelated to income management – such as seasonality in preterm births – contributing to the pre-trend in low birthweight. After controlling for preterm birth, we see a possible increase in the probability of low birthweight for babies exposed to income management for at least one full trimester of pregnancy, however, none of the individual treatment effects are statistically significant at the 5% level. Meanwhile, we continue to see a clear negative treatment effect on birthweight (Fig. 4(d)), but it is somewhat but smaller in magnitude than that shown in Fig. 4(a).

In sum, our event study analysis suggests that our main model should control for preterm birth, to avoid pre-existing confounding trends in that variable. When we do that, we see a clear reduction in average birthweight, which is largest for babies exposed to income management throughout all, or almost all, of their pregnancy. There is also evidence of an increase in the probability of low birthweight, although we are less confident in the magnitude of this increase given that it is less precisely estimated.

Table 4 sets out our main model estimates with and without controlling for preterm birth (Panel A). As is evident in Fig. 4, the estimated treatment effect is smaller for both birthweight (a reduction of 85 g, from 115 g), and for low birthweight (an increase of

3 percentage points, from 4.4 percentage points). The estimates are statistically significant at the 5% and 10% levels respectively. These are our preferred estimates.

5.3. Robustness checks

Recent advances in the econometric literature have highlighted shortcomings of the standard generalized difference-in-difference model using fixed effects. They find that in cases with heterogeneous treatment effects, models estimated using OLS may be biased, and may lead to incorrect conclusions due to negative weights being placed on some individual treatment effects. We test the robustness of our estimates to use of the Borusyak et al. (2021) estimator. This is an imputation method, similar to other proposed estimators (Gardner, 2021; Wooldridge, 2021). While there are multiple other estimators available (e.g. Callaway and Sant'Anna, 2020; de Chaisemartin and D'Haultfœuille, 2020), we find this to be the most suitable in our case, as it is flexible, allowing us to use individual-level, repeated cross-section data, with treatment defined based on exact birthdate and date of policy implementation, but to run models with unit fixed effects at the community level. Furthermore, because of the many implementation dates within our application, we find alternative estimators to be overly computationally intensive.

Table 4 (Panels B) shows that our results using this estimation method are very close to our benchmark results. They show a reduction in birthweight of 89 g, and an increase in the probability of low birthweight of 3 percentage points.[20]

As further empirical tests of the validity of our identification method, we run a placebo test, pretending that the policy rollout occurred one year before it actually happened. If our identification method is valid, we would expect to see no placebo-treatment effect. Conversely, if our method were picking up seasonal or other patterns that occur at the same time each year, we would find a placebo effect similar to our actual treatment effect. Panel C shows the results: we find small and statistically insignificant placebo effects, and these estimates become smaller after controlling for preterm birth.

Finally, as outlined by Borusyak et al. (2021), a placebo test may be a weak test of pre-trends; they suggest instead conducting an *F*-test on the event study coefficients from before the policy introduction. There are no significant pre-trends for either outcome, with small *F*-statistics and large *p*-values (Panel D). The F-statistics are even smaller in the models that control for preterm birth.

Together, the estimates in Panel B, C and D give us confidence that our main treatment estimate captures the effect of income management, is not biased by negative weights, and not simply capturing unobserved trends in birthweight that were present before income management.

5.4. Is lower birthweight a worse outcome?

A decline in average birthweight could represent better infant health if concentrated among heavier babies. However, quantile regressions estimates (Fig. 5), show the effect was present throughout the distribution, and larger at the lower end.

5.5. Other robustness tests

Our findings are robust to changes in research design (matching methods), sample selection criteria and restrictions, and alternative approaches to controlling for time and seasonal trends. These robustness tests are summarized in Fig. D.1, with details of each model in Appendix D. The treatment effects vary between -65 and -120 g.

6. Why did income management reduce birthweight?

We explore three channels that may explain the treatment effect: changes in fertility and maternal characteristics, maternal risky health behaviors and better access to quality care. We summarize the results of this mechanism exploration in Table 5, which shows, where applicable, the impact of income management directly on the mechanism (column 1), and the estimated treatment effect on birthweight after controlling for the mechanism (column 2). More details on each test can be found in Appendix E.

6.1. 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. But we find no evidence for this hypothesis. The policy had no impact on community-level fertility rates, and we see no significant difference in the medical history of women who gave birth after income management was introduced (Table 5, Panel A). Therefore, not surprisingly, controlling for the mother's medical history does not change the treatment effect.

[20] Note that the sample size in Panel B is smaller. This is because this model drops time periods for which all units are treated; this means that babies born in the post-rollout period (January–December 2009) are dropped. Indeed, further analysis following Jakiela (2021), shown in Appendix C, confirms that those later observations were most likely to have treatment effects estimated with negative weights.

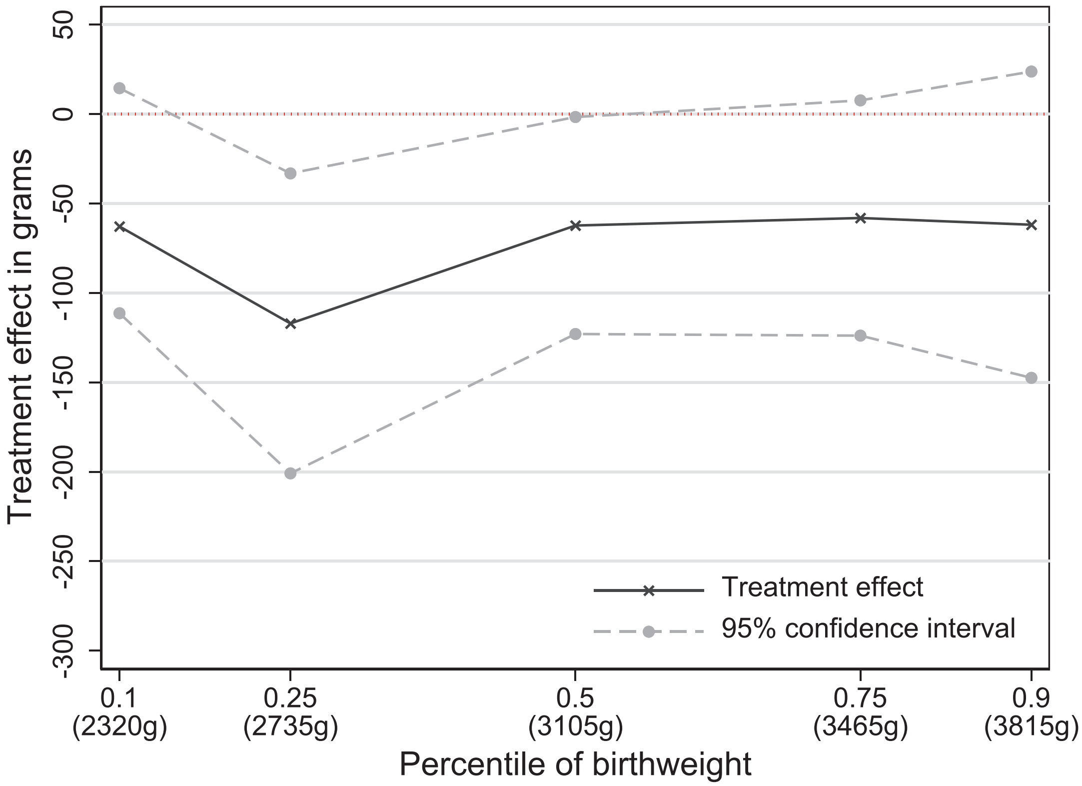


Fig. 5. Treatment effect by birthweight quantile. *Notes:* Estimated coefficients obtained from quantile regression models, where conditional treatment effect is estimated as the benchmark model with a control for preterm birth. Figures in parentheses on the horizontal axis represent birthweight at that percentile.

6.2. Maternal health behaviors

The negative treatment effect may be the result of a change in maternal health behaviors. 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. Alternatively, it may have led some households to stockpile restricted goods before the policy was introduced, leading to higher consumption of those goods in the following weeks and months. We can test this possibility because our data include self-reported information on whether the mother was drinking or smoking at the time of the first antenatal visit.[21]

We find that this is not the case. In fact, if income management was introduced before the first antenatal visit, women were 5 percentage points less likely to report smoking at that time, and 2 percentage points less likely to report drinking alcohol (Table 5, Panel B). But in both cases, the S.E. is so large that we have no certainty that this effect is not due to random variation. Furthermore, controlling for smoking and drinking behavior directly in the regression model does not change the treatment effect. We thus conclude that there is insufficient evidence to say whether the policy affected smoking and drinking during pregnancy, but if it did occur, this does not to explain the treatment effect.

6.3. Access to quality care and survival

A reduction in birthweight, even at the lower end of the distribution, does not necessarily indicate a worsening in infant health if the alternative is that the baby would not survive. 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, earlier detection of serious complications, and thus referral to emergency C-sections or delivery in a larger, better-resourced hospital. This could have increased survival probabilities of at-risk babies. Despite lower birthweight, it would have been a preferred health outcome. However, we find no effect of income management either on the location of birth or delivery by emergency C-section, and our treatment effect is little changed if we included these variables as controls (Table 5, Panel C).

[21] 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 included in the rollout period is therefore different from our main analysis sample.

Table 5
Mechanism tests.

	Impact of income management on mechanism (1)	Impact of income management on birthweight after controlling for mechanism (2)
Panel A: Selection into pregnancy		
Fertility (births per 1000 women per year, community-level analysis)	0.95	na
	(1.57)	
Any complications in medical history	0.00293	-85.03**
	(0.0380)	(33.69)
Panel B: Maternal health behaviors		
Smoking at first antenatal visit	-0.0504	-83.08**
	(0.0400)	(33.19)
Drinking alcohol at first antenatal visit	-0.0235	-84.30**
	(0.0186)	(33.64)
Panel C: Access to care		
Emergency C-section	0.00	-86.10**
	(0.02)	(33.63)
Born in main hospital	0.01	-85.91**
	0.01	(33.73)
Panel D: Survival		
Stillbirth	0.00606	-81.16**
	(0.00742)	(32.80)
Boy	0.01	na
	(0.03)	

Notes: Column 1 shows the impact of income management on the potential mechanism, after including the controls set out in Eq. (1), and controlling for preterm birth, where relevant. Community-clustered standard errors in parentheses. Column 2 shows the main estimation model for the impact of income management on birthweight, including all controls and a control for preterm birth, after including the mechanism variable as a control. In Panel B column 1, 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. In panel D, column 2 is not shown for the 'boy' variable as this is included as a control in the main model. Community-clustered standard errors in parentheses. The models are estimated on all births that occurred between 2001 and 2009 inclusive. $^{*}p < 0.10$. $^{**}p < 0.05$. $^{***}p < 0.01$.

We can also test the policy's impact on survival directly, by analyzing stillbirths. While we do not have data on miscarriage (defined in the NT as fetal loss before week 20 of pregnancy), we can use the baby's sex as a proxy, given that male fetuses are less likely to survive than female fetuses in adverse conditions (Catalano and Bruckner, 2006). A change in the sex balance may therefore indicate a changing probability of miscarriage. Again, we do not find evidence for these hypotheses. Treated and untreated babies did not differ in the probability of stillbirth, and the sex ratio did not change (Panel D). Controlling for these variables does not affect the estimated treatment effect on birthweight.

7. Are policy implementation challenges the cause of the negative effect?

Having ruled out medical and behavioral channels as cause of the reduction in birthweight, we are left with a more qualitative assessment of the policy's impact. The effect we find is large. For comparison, our effect size is one-third of the estimated effect of in-utero exposure to the Dutch famine (Stein and Susser, 1975) and 40% of the estimated effect of maternal fasting during pregnancy (Savitri et al., 2014). This is remarkable for an intervention that was not intended to change recipients' income, and in a population where most recipients were likely inframarginal. This comparison leads us to the hypothesis that the policy may have reduced consumption.

Returning to the policy logic, we can consider two ways through which the introduction of income management may have affected consumption. First, if the policy had its intended impact, it would affect consumption choices for some households by increasing spending on essentials. Our finding of a reduction in birthweight, evident throughout the birthweight distribution, is not consistent with an increase in consumption of essentials. Second, the policy may have affected consumption in unintended ways, potentially reducing household income and purchasing power.

This explanation is most consistent with our findings. The most likely channel would be through a temporary income shock that was caused by the new procedures for accessing benefits, leading to large values of unallocated quarantined funds and some payment suspensions.

7.1. Fund allocations and auto-income management

While we do not have data on income at the individual level, aggregate figures suggest very acute reductions in recipients' take-home income during the rollout period. This relates to the requirement for recipients to meet with a case officer to allocate their funds; until recipients did so, their quarantined funds could not be accessed. At the beginning of the rollout, up to half of quarantined funds were unallocated, or 'auto-income managed'. Two months into the rollout, about 50% of recipients could not access their quarantined benefits (Australian Institute of Health and Welfare, 2010, 30). This share decreased throughout the rollout period. But for most of the implementation period, about one in five dollars earmarked for the consumption of essentials was not available. By March 2009, over 18 months after the rollout began, only 3% of quarantined funds remained unallocated, though this is still substantial, representing an average of over A$200 per recipient (Australian Institute of Health and Welfare, 2010, 31).

7.2. Payment suspensions

A second source of disruption was payment suspension, which led to even larger income shocks for a smaller number of recipients. If a recipient's account was suspended, then all benefit payments (not just the quarantined amount) were withheld. 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 suspensions was large (AIHW p. 26-27). About one-third of all clients had at least one payment suspended. Assuming that one payment was equivalent to AU$210 per week (AU$105 quarantined),[22] then in total AU$2,067,660 were missing. Auto-income management was the most common reason for these suspensions (AIHW, 2010).

During the rollout period, 1833 clients experienced a suspension because they failed to contact the government welfare agency within 13 weeks of being auto-income managed. For these clients, they would have had access to only half of their benefit payments for 13 weeks in a row. 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.

Independent of whether the suspended payment was restored in week 14, those clients would have been short by at least AU$1,365 (that is, $105 quarantined per week over 13 weeks). These shortfalls represent a severe temporary income shock, which is likely to cause both a reduction in consumption and an increase in stress.[23]

7.3. Stress

Stress experienced by pregnant women can affect development in utero, as cortisol is passed on to the fetus (Aizeret al., 2016). The income shortfalls described above are likely to have caused stress, potentially exacerbating the effect of the income shock itself. In addition, 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" (Yu et al., 2008, 20).

7.4. Additional logistical and implementation challenges

Aside from the initial income shock, more recent evidence shows that income management can reduce purchasing power and change behavior, even for participants receiving the full nominal value of their payments. This is because the policy limits recipients' participation in the cash economy, and their ability to purchase second-hand goods informally. It also restricts participants to shop at certain retailers, which may not have the lowest prices (Marston et al., 2020). This was acknowledged in a 2020 report from the government department that administers the policy, stating that income management "is a largely incoherent policy that has a limited ability to create change within communities", citing technological limitations which restrict participants' spending choices in ways that are not intended by the policy design (Department of Social Services, 2020, 33).

Marston et al. (2020) describe additional challenges based on interviewee and survey respondents' experiences with income management in other parts of Australia. They find that when setting up initial spending allocations, some recipients are not open with case workers about their expenses, for fear of having their payments cut off entirely. This fits with survey evidence from the NT, finding that participants do not trust the payment agency and would not feel safe to speak with case workers when they have a problem (Equality Rights Alliance, 2011). Therefore, recipients may end up with inappropriate allocations, leading to ongoing challenges in meeting their expenses.

Finally, income management may affect behavior and spending in more subtle ways. For instance, some participants have reported feelings of stigma and shame when using their quarantined funds at retailers (Chaloner and Kaelah, 2021; Mendes et al., 2020; Watt, 2020; Equality Rights Alliance, 2011; Deloitte Access Economics, 2015), and this may lead to a change in shopping patterns to avoid those experiences.

While these additional channels may have reduced recipients' purchasing power on an ongoing basis, it is unlikely that these factors alone would be sufficient to lead to the large reduction in birthweight that we find. Given the magnitude of the effect we find, we conclude that it was largely driven by large income reductions relating to auto-income management and exacerbated by the stress accompanying these income reductions.

[22] 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).

[23] 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). We may assume suspension rates for these other reasons would have been similar before income management, but we do not have any data to confirm this.

8. Discussion and conclusion

Since 2007, the Australian government has continued to expand income management to additional sites across Australia. Although the welfare of children has taken center stage in the policy debate, there is little evidence on how the introduction of this new policy affected children. This study is one of the first attempts to quantify the impact on children's welfare by estimating the effect of in utero exposure to the rollout of income management on birthweight.

We conclude that the introduction of income management did not increase birthweight. Instead, it had a negative effect. Our benchmark model suggests a reduction in birthweight of 85 g. These estimates are robust. We also find evidence that this increase was largest at the lower end, with an increase in the risk of low birthweight by 3 percentage points. With a pre-treatment probability of 14.8 of being born with low birthweight, this implies an increase of 20%.

These effects are large relative to the literature on the effects of transfer programs (Almond et al., 2011; Hoynes et al., 2011). We find that the impact of income management is closer in magnitude to the effect of fasting or famine than to other transfer programs. For instance, Savitri et al. (2014) find that mothers who fasted during pregnancy gave birth to babies that were on average 200 g lighter. The effect was largest for babies exposed to fasting during the first trimester – a finding consistent with our event study estimates.

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. This conclusion stands in contrast to findings from the main government evaluation report (AIHW, 2010), which relied on focus groups and a small, non-randomly sampled survey. This contrast underscores the importance of appropriate methods when evaluating the 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 largely to a reduction in consumption because of initial implementation challenges. In particular, when benefit recipients begin income management, 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 we estimate reduced short-term spending across all communities by a total of A$2 million.

With our data, we cannot disentangle the direct contribution of this income shortfall, the accompanying increase in stress, and other longer-term logistical issues, to the reduction in birthweight. 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 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 (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. Furthermore, logistical issues that reduce participants' purchasing power relative to cash have continued, over a decade after the policy was initially introduced (Marston et al., 2020), showing that with an administratively complex policy like income management, discrepancies between policy 'as designed' and policy 'as implemented' are not necessarily resolved during the initial rollout.

While our findings raise additional questions, they also convey a clear message. As Gracey and King (2009) describe, there are demonstrated and persistent inequalities between Indigenous and non-Indigenous peoples worldwide. Income management represents one attempt to reduce these disparities, but we find evidence it exacerbated them. The unexpected negative effect of the introduction of income management highlights the importance of careful design, testing and consultation for major social policy changes.

Our identification mechanism cannot provide evidence on whether these negative effects have persisted for children born after the policy rollout period. However, Bray (2020) analyses aggregate data up to 10 years after income management was introduced, to test for evidence of longer-term improvements. He finds that rates of low birthweight and infant mortality among Aboriginal children in the NT have worsened since the policy introduction, both relative to non-Aboriginal children in the NT, and to Aboriginal children in other parts of Australia. This is not necessarily evidence that income management caused these divergences. However, it does suggest that income management has not achieved its goal of reducing early life disadvantages faced by Aboriginal children in the NT, even over the longer-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 employment prospects (Almond et al., 2018).

Author statement

Professor Dr Stefanie Schurer (University of Sydney) has no competing interests.
Ms Mary-Alice Doyle (LSE) has no competing interests.
Professor Sven Silburn (Menzies School of Health Research) has no competing interests.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jhealeco.2022.102618.

References

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

Aizer, A., Currie, J., 2014. The Intergenerational transmission of inequality: maternal disadvantage and health at birth. Science 344 (6186), 856–861. doi:10.1126/science.1251872.

Aizer, A., Stroud, L., Buka, S., 2016. Maternal stress and child outcomes: evidence from siblings. J. Hum. Resour. 51 (3), 523–555. doi:10.3368/jhr.51.3.0914-6664R.

Almond, D., Chay, K.Y., Lee, D.S., 2005. The costs of low birth weight. Q. J. Econ. 120 (3), 1031–1083.

Almond, D., Currie, J., Duque, V., 2018. Childhood circumstances and adult outcomes: act II. J. Econ. Lit. 56 (4), 1360–1446. doi:10.1257/jel.20171164.

Almond, D., Hoynes, H.W., Schanzenbach, D.W., 2011. Inside the war on poverty: the impact of food stamps on birth outcomes. Rev. Econ. Stat. 93 (2), 387–403. doi:10.1162/REST_a_00089.

Altman, J., 2016. Blind-Sided by basics: three perspectives on income management in an aboriginal community in the northern territory. Aust. J. Soc. Issues 51 (4), 487–502. doi:10.1002/j.1839-4655.2016.tb01245.x.

Australian Human Rights Commission. 2011. "The suspension and reinstatement of the RDA and special measures in the NTER." November 2, 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, S.L., Gertler, P.J., 2008. The impact of Mexico's conditional cash transfer programme, oportunidades, on birthweight. Trop. Med. Int. Health 13 (11), 1405–1414. doi:10.1111/j.1365-3156.2008.02157.x.

Bharadwaj, P., Eberhard, J.P., Neilson, C.A., 2018. Health at birth, parental investments, and academic outcomes. J. Labor Econ. 36 (2), 349–394. doi:10.1086/695616.

Bharadwaj, P., Lundborg, P., Rooth, D.O., 2018. Birth weight in the long run. J. Hum. Resour. 53 (1), 189–231. doi:10.3368/jhr.53.1.0715-7235R.

Bielefeld, S., 2018. Government mythology on income management, alcohol, addiction and indigenous communities. Crit. Soc. Policy 38 (4), 749–770. doi:10.1177/0261018317752735.

Borusyak, K., Jaravel X., and Spiess J.. 2021. "Revisiting event study designs: robust and efficient estimation," 48.

Bray, J.R., 2020. Measuring the Social Impact of Income Management in the Northern Territory: An Updated Analysis, 136. Centre for Aboriginal Economic Policy Research, Canberra, Australia CAEPR Working Paper.

Brimblecombe, J.K., McDonnell, J., Barnes, A., Dhurrkay, J.G., Thomas, D.P., Bailie, R.S., 2010. Impact of income management on store sales in the northern territory. Med. J. Aust. 192 (10), 549–554. doi:10.5694/j.1326-5377.2010.tb03632.x.

Brough, M. 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.

Callaway, B., Sant'Anna, P.H.C., 2020. Difference-in-differences with multiple time periods. J. Econom. doi:10.1016/j.jeconom.2020.12.001, December.

Catalano, R., Bruckner, T., 2006. Secondary sex ratios and male lifespan: damaged or culled cohorts. Proc. Natl. Acad. Sci. 103 (5), 1639–1643. doi:10.1073/pnas.0510567103.

Cerna, L., 2013. The Nature of Policy Change and Implementation: A Review of Different Theoretical Approaches. OECD, Paris.

Chaisemartin, C., D'Haultfœuille, X., 2020. Two-way fixed effects estimators with heterogeneous treatment effects. Am. Econ. Rev. 110 (9), 2964–2996. doi:10.1257/aer.20181169.

Chaloner, C., and Kaelah. 2021. "The Cashless welfare card makes life hard – you can't do normal things | Craig Chaloner and Kaelah." The Guardian, January 3, 2021, sec. Opinion. http://www.theguardian.com/commentisfree/2021/jan/04/the-cashless-welfare-card-makes-life-hard-you-cant-do-normal-things.

Cobb-Clark, D.A., Kettlewell, N., Schurer, S., Silburn, S., 2021. The effect of quarantining welfare on school attendance in indigenous communities. J. Hum. Resour. 1218. doi:10.3368/jhr.1218-9909R2, September.

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

Cooke, M., Mitrou, F., Lawrence, D., Guimond, E., Beavon, D., 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 Int. Health Hum. Rights 7 (1), 9. doi:10.1186/1472-698X-7-9.

Currie, J., Gahvari, F., 2008. Transfers in cash and in-kind: theory meets the data. J. Econ. Lit. 46 (2), 333–383.

Currie, J., Moretti, E., 2007. Biology as destiny? Short- and long-run determinants of intergenerational transmission of birth weight. J. Labor Econ. 25 (2), 231–264. doi:10.1086/511377.

Currie, J., Rossin-Slater, M., 2015. Early-life origins of life-cycle well-being: research and policy implications. J. Policy Anal. Manag. 34 (1), 208–242. doi:10.1002/pam.21805.

Dalley, C., 2020. The 'white card' is grey: surveillance, endurance and the cashless debit card. Aust. J. Soc. Issues 55 (1). doi:10.1002/ajs4.100.

Dawson, J., Augoustinos, M., Sjoberg, D., Canuto, K., Glover, K., Rumbold, A., 2021. Closing the gap: examining how the problem of aboriginal and torres strait islander disadvantage is represented in policy. Aust. J. Soc. Issues 56 (4), 522–538. doi:10.1002/ajs4.125.

Dee, T.S., 2011. Conditional cash penalties in education: evidence from the learnfare experiment. Econ. Educ. Rev. Spec. Issue Educ. Health 30 (5), 924–937. doi:10.1016/j.econedurev.2011.05.013.

Deloitte Access Economics. 2015. "Place based income management - medium term outcomes evaluation report." https://www.dss.gov.au/sites/default/files/documents/10_2015/pbim_medium_term_outcome_report_final_26062015.pdf.

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.

Department of Social Services. 2020. "Cashless debit card final assessment regulation impact statement." Office of Best Practice Regulation 24100. Canberra: Australian Government. https://ris.pmc.gov.au/sites/default/files/posts/2020/12/final_ris_-_cashless_welfare_-_26.11.20.pdf.

Dobbins, T.A., Sullivan, E.A., Roberts, C.L., Simpson, JM., 2012. Australian national birthweight percentiles by sex and gestational age, 1998–2007. Med. J. Aust. 197 (5), 291–294. doi:10.5694/mja11.11331.

Durlak, J.A., DuPre, E.P., 2008. Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. Am. J. Commun. Psychol. 41 (3–4), 327–350. doi:10.1007/s10464-008-9165-0.

Equality Rights Alliance. 2011. "Documenting women's' experience of income management in the northern territory." Canberra, Australia. https://www.alrc.gov.au/wp-content/uploads/2019/08/cfv_143_equality_rights_alliance_-_womens_voices_for_gender_equality_.pdf.

Ethridge, M.E., Percy, S.L., 1993. A new kind of public policy encounters disappointing results: implementing learnfare in Wisconsin. Public Adm. Rev. 53 (4), 340–347. doi:10.2307/977146.

Freyaldenhoven, S., Hansen, C., Shapiro, J.M., 2019. Pre-event trends in the panel event-study design. Am. Econ. Rev. 109 (9), 3307–3338. doi:10.1257/aer.20180609.

Gardner, J., 2021. Two-Stage Differences in Differences. Department of Economics, University of Mississippi.

Gennetian, L., Darling, M., Aber, J., 2016. Behavioral economics and developmental science: a new framework to support early childhood interventions. J. Appl. Res. Child. Inf. Policy Child. Risk 7 (2), 1–35. https://digitalcommons.library.tmc.edu/childrenatrisk/vol7/iss2/2.

Gracey, M., King, M., 2009. Indigenous health part 1: determinants and disease patterns. Lancet N. Am. Ed. 374 (9683), 65–75. doi:10.1016/S0140-6736(09)60914-4.

Gresham, E., Byles, J.E., Bisquera, A., Hure, A.J., 2014. Effects of dietary interventions on neonatal and infant outcomes: a systematic review and meta-analysis. Am. J. Clin. Nutr. 100 (5), 1298–1321. doi:10.3945/ajcn.113.080655.

Howard, J., 2007. To stabilise and protect: little children are sacred. Sydney Pap. 19 (3), 68.

Hoynes, H.W., Page, M., Stevens, A.H., 2011. Can targeted transfers improve birth outcomes? Evidence from the introduction of the WIC program. J. Public Econ. 95 (7–8), 813–827. doi:10.1016/j.jpubeco.2010.12.006.

Hoynes, H.W., Schanzenbach, D.W., 2009. Consumption responses to in-kind transfers: evidence from the introduction of the food stamp program. Am. Econ. J. Appl. Econom. 1 (4), 109–139. doi:10.1257/app.1.4.109.

Hunter, B., 2009. A half-hearted defence of the CDEP scheme. Fam. Matters (81) 43–54.

Jakiela, P.. 2021. "Simple diagnostics for two-way fixed effects." ArXiv:2103.13229 [Econ, q-Fin], March. http://arxiv.org/abs/2103.13229.

Klein, E., 2016. Neoliberal subjectivities and the behavioural focus on income management. Aust. J. Soc. Issues 51 (4), 503–523. doi:10.1002/j.1839-4655.2016.tb01246.x.

Lamb, D., Young, M., 2011. 'Pushing buttons': an evaluation of the effect of aboriginal income management on commercial gambling expenditure. Aust. J. Soc. Issues 46 (2), 119–140.

Margolies, A., Hoddinott, J., 2015. Costing alternative transfer modalities. J. Dev. Eff. 7 (1), 1–16. doi:10.1080/19439342.2014.984745.

Marston, G., Mendes P., Bielefeld S., Peterie M., Staines Z., and Roche S.. 2020. "Hidden costs: an independent study into income management in Australia."

Markham, F., Biddle, N., 2018. Income, Poverty and Inequality. Centre for Aboriginal Economic Policy Research Census Paper (2). https://caepr.cass.anu.edu.au/research/publications/income-poverty-and-inequality.

Meier, K.J., McFarlane, D.R., 1995. Statutory coherence and policy implementation: the case of family planning. J. Public Policy 15 (3), 281–298.

Mendes, P., Roche, S., Marston, G., Peterie, M., Staines, Z., Humpage, L., 2020. The social harms outweigh the benefits: a study of compulsory income management in greater shepparton and playford. Aust. Soc. Work 0 (0), 1–15. doi:10.1080/0312407X.2020.1820536.

Mitrou, F., Cooke, M., Lawrence, D., Povah, D., Mobilia, E., Guimond, E., Zubrick, S.R., 2014. Gaps in indigenous disadvantage not closing: a census cohort study of social determinants of health in Australia, Canada, and New Zealand from 1981–2006. BMC Public Health 14 (1), 201. doi:10.1186/1471-2458-14-201.

Risnes, K.R., Vatten, L.J., Baker, J.L., Jameson, K., Sovio, U., Kajantie, E., Osler, M., et al., 2011. Birthweight and mortality in adulthood: a systematic review and meta-analysis. Int. J. Epidemiol. 40 (3), 647–661. doi:10.1093/ije/dyq267.

Roth, J., 2021 forthcoming. Pre-test with caution: event-study estimates after testing for parallel trends. AER Insights. https://jonathandroth.github.io/assets/files/roth_pretrends_testing.pdf.

Savitri, A.I., Yadegari, N., Bakker, J., van Ewijk, R.J., Grobbee, D.E., Painter, R.C., Uiterwaal, C.S., Roseboom, T.J., 2014. Ramadan fasting and newborn's birth weight in pregnant Muslim women in The Netherlands. Br. J. Nutr. 112 (9), 1503–1509. doi:10.1017/S0007114514002219.

Silburn, S., Guthridge, S., McKenzie, J., Su, J.Y., He, V., Haste, S., 2018. Early Pathways to School Learning: Lessons from the NT Data Linkage Study. Menzies School of Health Research, Darwin.

Southworth, H.M., 1945. The economics of public measures to subsidize food consumption. Am. J. Agric. Econ. 27 (1), 38–66. doi:10.2307/1232262.

Stein, Z., Susser, M., 1975. The Dutch famine, 1944–1945, and the reproductive process. I. Effects on six indices at birth. Pediatr. Res. 9, 70–76.

Strand, L.B., Barnett, A.G., Tong, S., 2011. Methodological challenges when estimating the effects of season and seasonal exposures on birth outcomes. BMC Med. Res. Methodol. 11 (1), 49. doi:10.1186/1471-2288-11-49.

Venn, D., Biddle N., and Sanders W.. 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.

Watt, E., 2020. Is the BasicsCard 'Shaming' Aboriginal people? Exploring the differing responses to welfare quarantining in Cape York. Aust. J. Soc. Issues 55 (1), 40–50. doi:10.1002/ajs4.94.

Wilcox, A.J., 2001. On the importance—and the unimportance— of birthweight. Int. J. Epidemiol. 30 (6), 1233–1241. doi:10.1093/ije/30.6.1233.

Wild, R., Anderson, P., 2007. Ampe Akelyernemane Meke Mekarle 'little children are sacred. Northern Territory Board of Inquiry into the Protection of Aboriginal Children from Sexual Abuse. Northern Territory Government, Darwin.

Wooldridge, J.M., 2021. Two-Way Fixed Effects, the Two-Way Mundlak Regression, and Difference-in-Differences Estimators. Social Science Research Network, Rochester, NY SSRN Scholarly Paper ID 3906345 doi:10.2139/ssrn.3906345.

Yu, P., Duncan, M.E., Gray, B., 2008. Northern Territory Emergency Response: Report of the Review Board. Commonwealth of Australia, Canberra.