

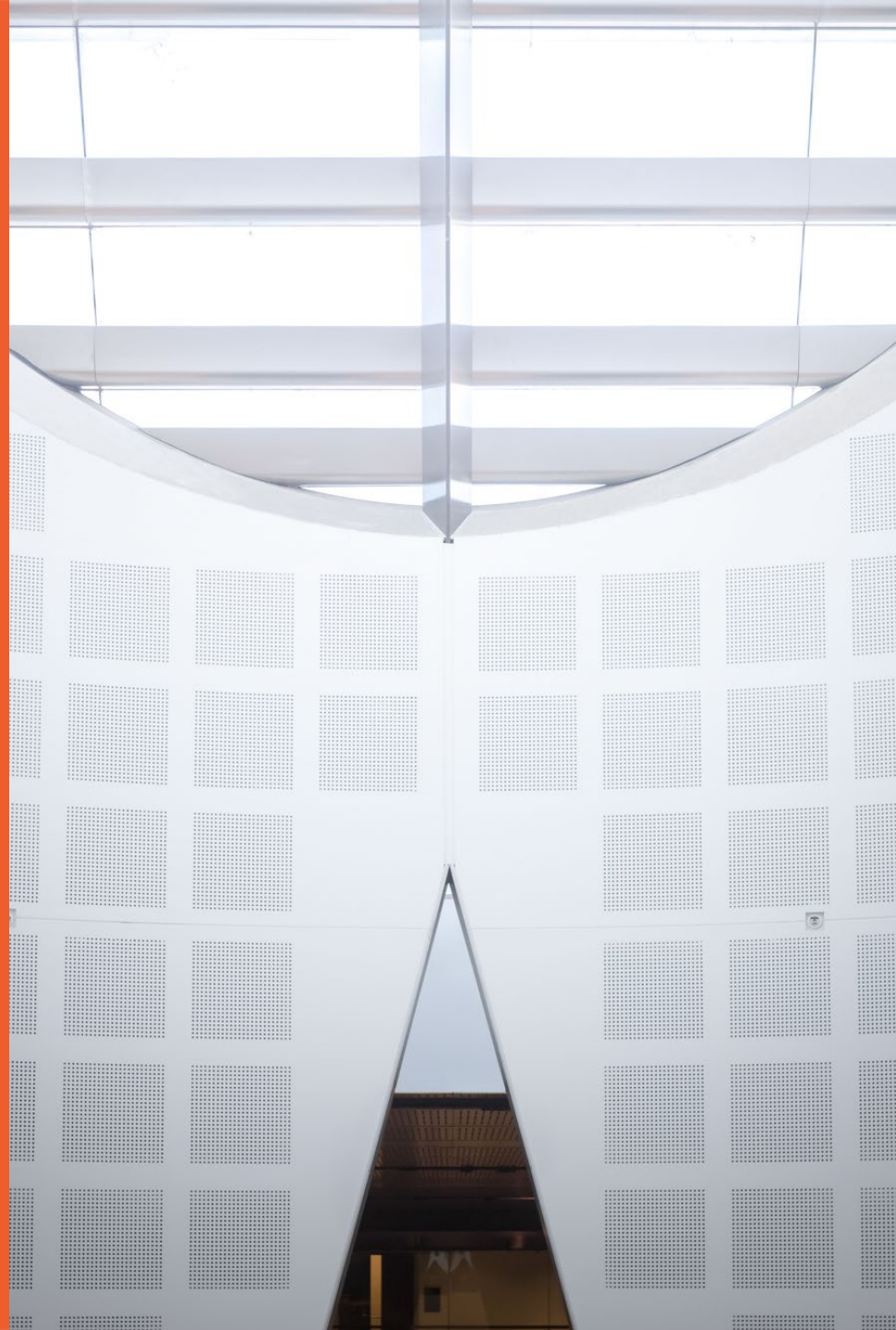
Indigenous Evaluation: Better than Nothing is Not Good Enough

Professor Deborah Cobb-Clark
University of Sydney

Applied Experimental/Behavioral
Economics
UTS Centre for Policy and Market Design
(CPMD)
23 November 2018



THE UNIVERSITY OF
SYDNEY

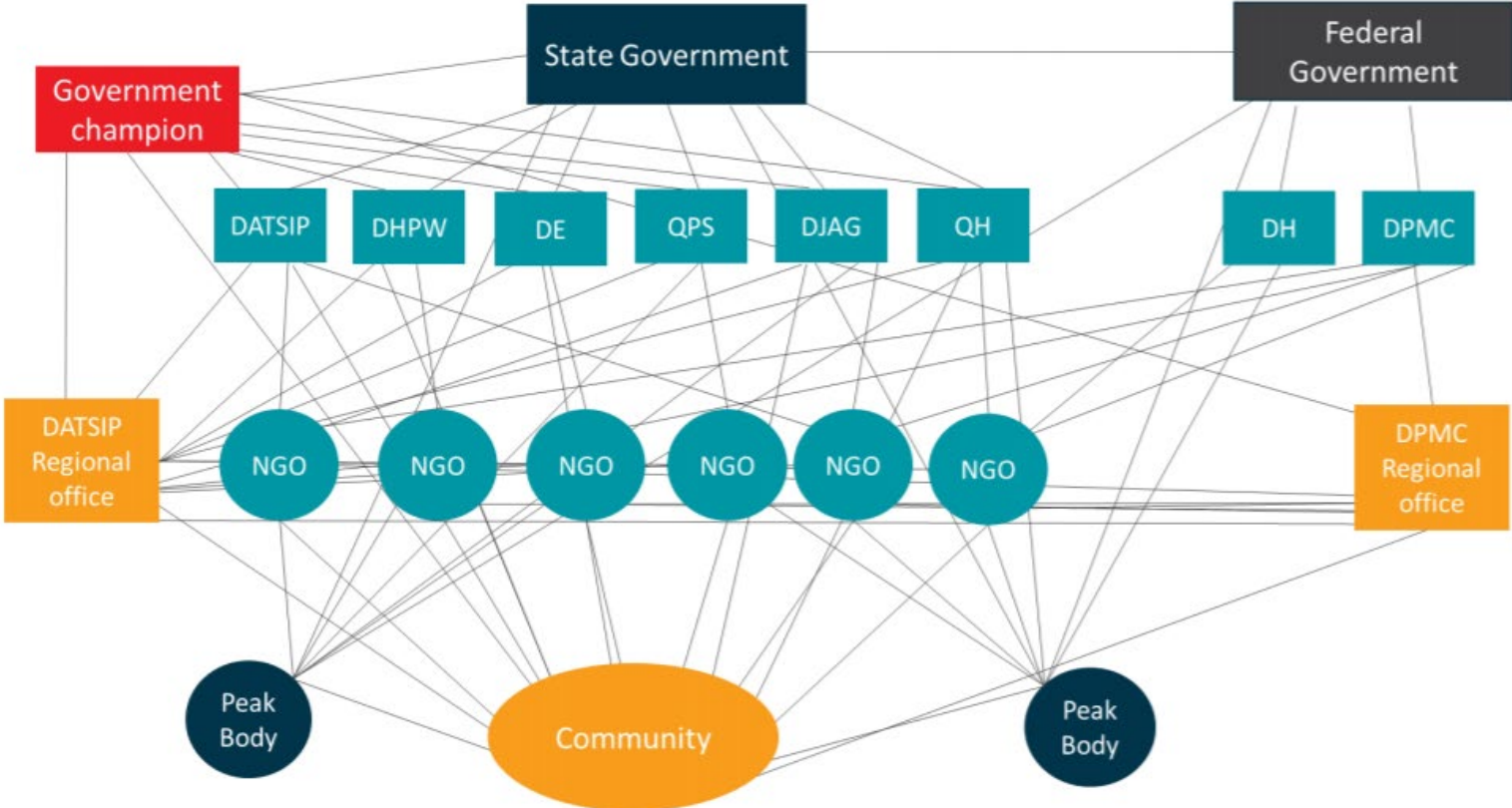


Indigenous Program Evaluation

- Small population size (i) makes many data sets totally unsuitable and (ii) limits statistical power ... Only large impacts can be detected;
- Uniqueness of Indigenous communities means that it is difficult to define a meaningful control group;
- Indigenous policy is targeted at communities which are characterised by fluid, extended family structures and cultural norms for resource sharing ... We can really only estimate community-level impact of the intention to treat;
- Indigenous communities are NOT randomly selected for treatment ... Effects of the selection process itself will be part of the estimated treatment impact;
- Multiple interventions affect the Indigenous at any point in time and “control group” is also treated ... Estimates then capture the marginal effect of one set of interventions and another set, many of which overlap. Is this interesting?

Service Delivery in Remote and Discrete Aboriginal & Torres Strait Islander Communities (QLD Productivity Commission 2017)

Figure 4 The bureaucratic maze



Note: To simplify, the map shows only a subset of the departments, authorities and NGOs involved in service delivery, design and coordination.

Has a lack of Indigenous-specific evaluation limited our ability to learn from past policies?

I have no direct evidence, ...but:

- ❖ Very hard to believe that this is not the case;
- ❖ If nothing else, the continuing gap in outcomes in the face of very substantial resources commitments clearly shows that we must do better.

Program evaluation that is well-done, methodologically sound, and corresponds to accepted scientific principles is critical to achieving our goal to eliminate the gap in outcomes.

Lessons from Health Care Evaluation

Reviews in the 1990s cast doubt on the scientific reliability of published evaluations in health care ... increased transparency and wider dissemination of results was seen as fundamental to raising the evaluation quality.

Incomplete, non-systematic reporting of results undermines the RCT methodology:

- Results of many trials are never published and those that are published are systematically more likely to show that the intervention is effective;
- This sort of positive publication bias makes it impossible to form valid judgements about an intervention's true effectiveness from the published literature;
- There may be strong financial incentives to withhold negative results and suppress data;

Lessons from Health Care Evaluation

We need periodic methodological assessments of economic evaluations using adequate sampling frames. The assessments should be ongoing and publically accessible. **Unless swift action is taken, low methodological quality risks bringing the practice of economic evaluation into disrepute—an outcome that is in no one's interest.**

Jefferson & Demicheli (2002) *British Medical Journal*

Since then:

- Key drug failures (e.g. Paxil) have focused the collective mind on the value of sound evaluations in a way that is unlikely to happen in other areas.
- In 2004, ICMJE made trial registration a necessary condition for the publication of any manuscript reporting trial results;
- In 2005, Australian New Zealand Clinical Trials Registry (ANZCTR) was established at U. Syd as part of (WHO) Registry Network;

Economic and Social Policy Evaluation

The arguments in favour of an arrangement like a clinical trials registry are as compelling in economic and social policy as they are in health care:

- Initiatives in Indigenous, education, or income-support policy have as profound an effect on individuals' lives as those in health care ... Must be as important to evaluate them;
- Positive publication bias must surely skew the published results. Here, “positive” refers not to the efficacy of the particular drug or treatment, but to the desirability of the program from bureaucrats' or politicians' perspectives.
- The tension between private (manufacturer) and public interest in publicizing the results of medical trials is the same as the tension that arises when government departments or non-governmental organizations have a private (political) incentive to withhold information about the impact of particular programs or policy initiatives;

Economic and Social Policy Evaluation

Increased transparency and wider dissemination of results is essential to improving the quality impact evaluations of Indigenous policy, education initiatives, as well as income-support, disability, and job training programs, etc.

In particular, greater transparency would

- put pressure on evaluators to lift their game;
- allow evaluations themselves to be evaluated against sound scientific principles ... we can then judge which to weight more heavily and which to ignore;
- provide an opportunity for truly informed public debate about the issues facing us; and
- substantially enhance our chances for sound decision making.

Better than nothing is not good enough

We are currently spending billions on major social initiatives in Indigenous policy, educational reform, and supporting the disabled.

Sound, independent program evaluation will be critical, yet current system generally produces poor-quality evaluations that do not tell us very much:

- We are almost never doing a true impact evaluation.
- Often evaluations are conducted within the very government agencies responsible for meeting program objectives.
- When external evaluators are used, it is common for the government to insist that the results not be published.

“We need to be aware that not all evaluations are equally compelling. There can be a temptation for a government department to conduct tokenistic, low-quality evaluations that tick-the-box for a program being evaluated ... Such evaluations are usually inconclusive, which has the added benefit of not risking embarrassment to the minister championing the program.” Siminski (2016)

Better than nothing is not good enough

What to do?

- Publically register economic and social program evaluations;
- Embed and fund impact evaluation plan in the design of the program from the start;
- Use capacity constraints which require programs to be “rolled out”;
- Bring academics (and academic publications) back into the picture;
- Establish separate, independent agency to commission all policy evaluations on behalf of the government;
- Make all components of any program evaluation--including the unit-record data on which it rests--widely and publically available, so that results can be replicated and confirmed;
- Make widespread publication of evaluation results the norm.

The Effect of Quarantining Welfare on School Attendance in Indigenous Communities

**Deborah Cobb-Clark
Nathan Kettlewell
Stefanie Schurer**

University of Sydney



Income Management in the Northern Territory

In 2007 the Australian Government radically changed the way welfare payments were delivered.

- For those living in prescribed (remote) Aboriginal communities, 50% of welfare benefits were quarantined.
- Quarantined benefits could not be spent on prohibited products (i.e. alcohol, tobacco, gambling and pornography).
- Quarantined income was stored in an income management account. Initially there were three access methods:
 - Make purchases directly from licensed stores;
 - Obtain store cards (gift cards) from Centrelink;
 - Third party deductions (e.g. to landlord, utility company)
- In 2008 the Basics Card was introduced, which significantly improved access to welfare benefits (AIHW, 2010).

Overview

Goals: Evaluate the causal effect of income management on school attendance (2006 – 2009).

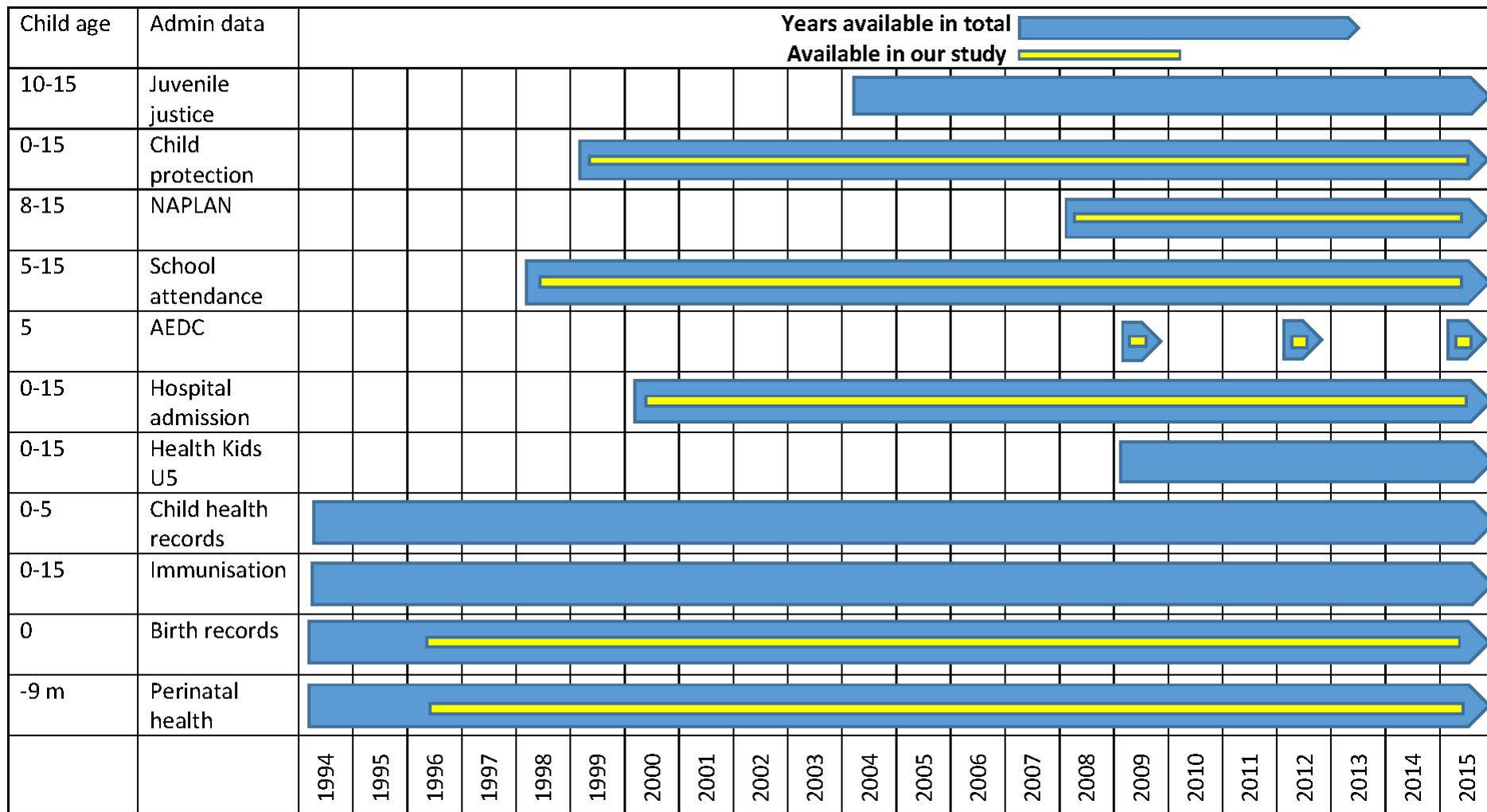
Method: Our identification strategy exploits the staggered rollout of income management across the Northern Territory.

Data: Administrative data covering the enrolment/attendance records of all public sector schools in the Northern Territory.

Results: We find no evidence that income management improved school attendance. In fact, attendance decreased in the short run. We provide evidence that this was due to implementation issues.

This is the first paper to estimate the causal effect of income management on social outcomes.

NT Data Linkage Project (Joint with Menzies)



Funding is through a Partnership Project (2014-2017) between the National Health and Medical Research Council (NHMRC) and the NT Government.

Data – NT Early Childhood Data Linkage Project

- We use daily attendance and enrollment records from the NT Department of Education.
- The sample is restricted to students (born after 1993) enrolled in grades 1 – 12 the period 2006-2009 (inclusive).
 - This covers approximately ± 1.5 years from the rollout start/end.
- We have 130 schools in 78 separate communities.
- Observation unit is student-day.
 - We end up with an unbalanced panel of 9,162 students and ≈ 3.5 m student-day observations.

Community Characteristics

	Aus Pop.	Sample Mean	Sample Min	Sample Max
Population size		428	83	1904
Percent male	49.4%	48.6%	40.9%	56.5%
Median age (years)	37	22.1	18	27
English only language spoken at home	78.5%	17.23%	0%	94.4%
Labour force part. rate	64.6%	37.8%	6.90%	83.5%
Employment rate	94.8%	86%	9.2%	100%
Median Personal Income	\$466	\$209.82	\$148	\$466
Average people per household	2.6	6.08	3.3	9.6

Note: Data are from the 2006 Australian Census. For the sample characteristics, N=64 in the case of population and percentage males. N=55 for all other variables. Community data are for the Indigenous Local Area for that community. For the missing observations, a suitably granular spatial unit could not be identified in the Census data.

Identification

We require that the rollout of income management is unrelated to temporal changes in school attendance patterns.

- We explicitly consider whether our results are confounded by other NTER initiatives. (This doesn't seem important).
- There is no clear spatial pattern to the rollout (see Figure 2).
- We also show that the rollout is largely unrelated to pre-existing community characteristics; $R^2 < .1$ (see Table 1).
- An event study analysis shows that there was no systematic trends preceding income management and that the policy effect occurs precisely around the program start date.
- No evidence of a “pseudo policy effect”;

Overall, we have strong evidence that our identification assumption holds.

Identification

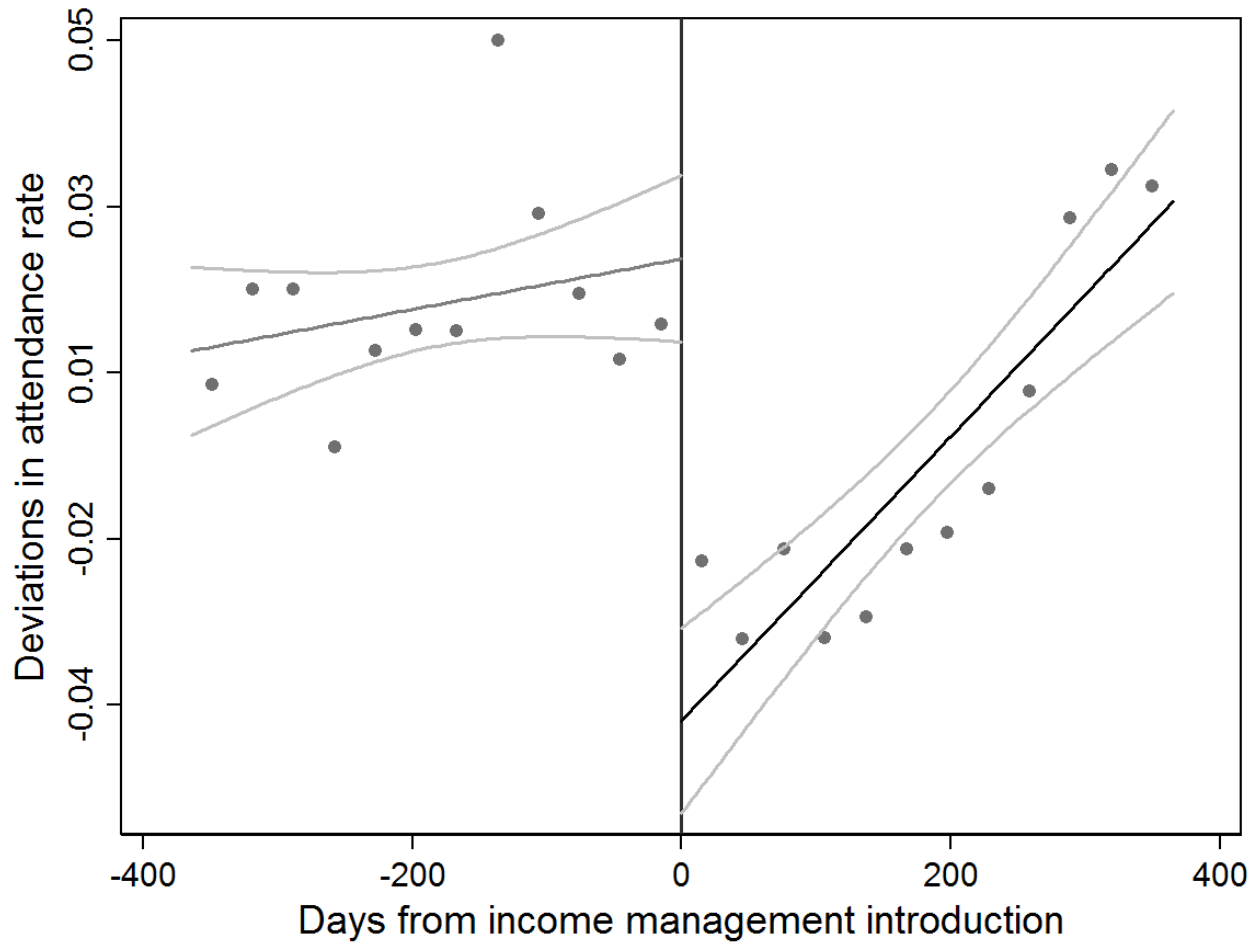
Event Study Analysis

- Before turning to our main results, we provide evidence on the validity of our identification strategy using an event study model.
- This approach is also used to test for systematic time trends wrt the timing of any policy effect.
- We estimate the following model:

$$Y_{ist} = \alpha + \sum_{d=-365}^{365} \pi_i \mathbf{1}(\widetilde{\tau}_{st} = d) + \gamma_s + \epsilon_{ist}$$

- Event-date coefficients map the relationship between onset of income management and attendance probability.

Event Study Results



Model

Our baseline model is a difference-in-differences estimator:

$$Y_{isldt} = \alpha + \beta IM_{isldt} + \gamma_s + \tau_t + \lambda_l + \delta_d + \epsilon_{isldt} \quad (1)$$

$Y = 1$ if attended school for the whole day

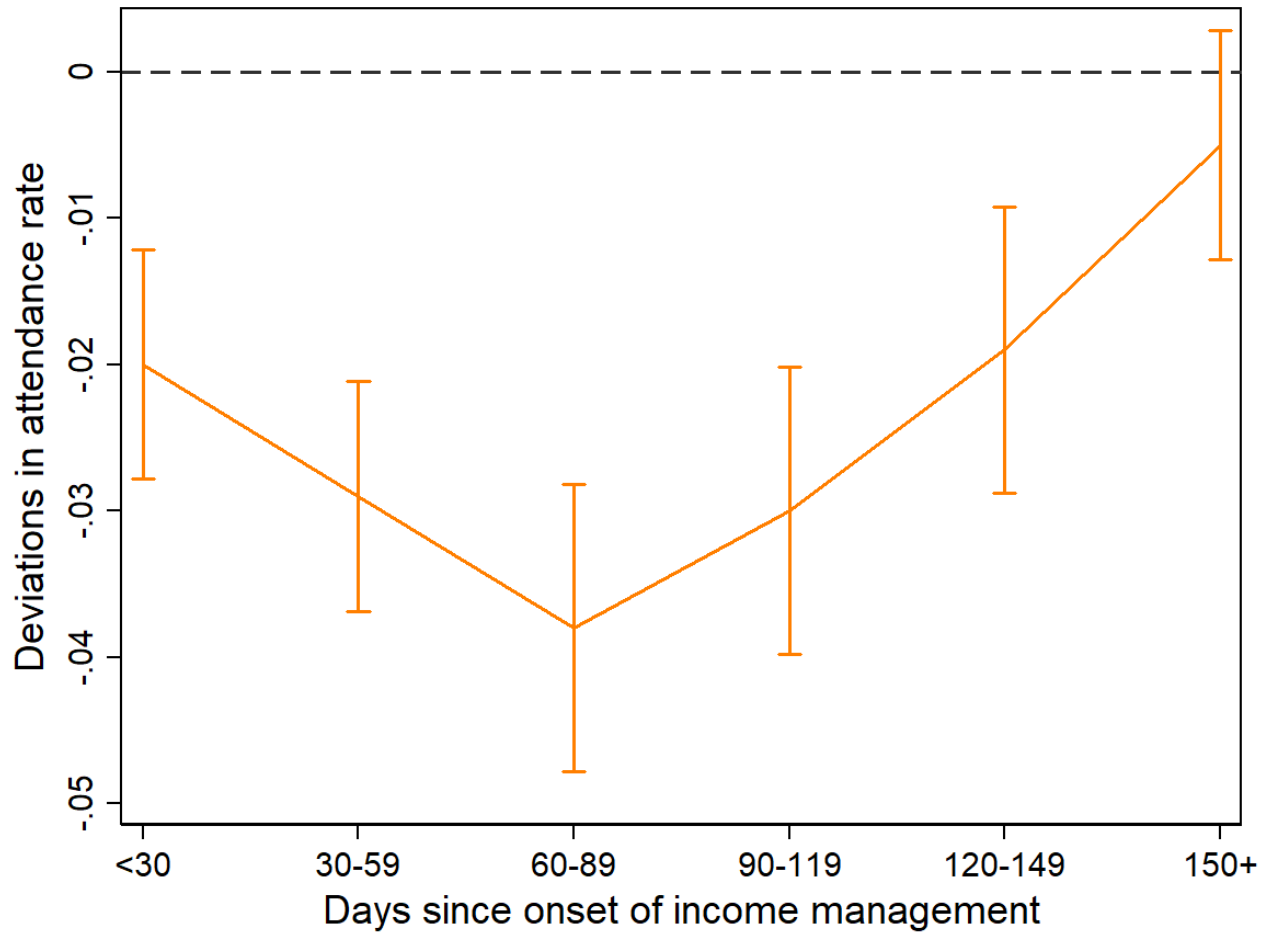
IM = dummy for attending school in income managed community

- We control for school (s), time (t), grade level (l) and day of the week (d) fixed effects.
- We also estimate a version of (1) allowing for school-specific linear time trends with school-specific school-term shifters.

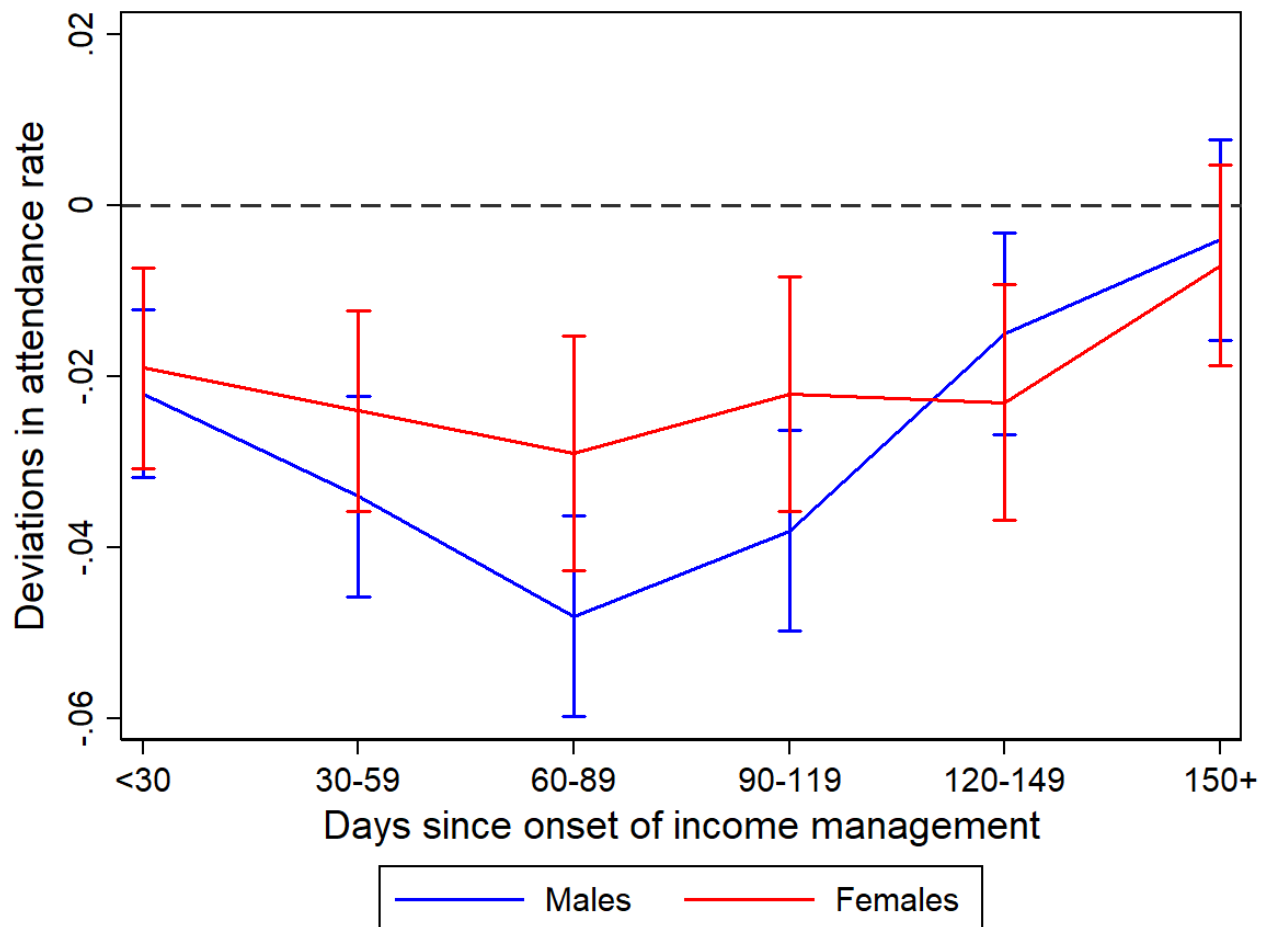
Table 4: Main Regression Results: Aggregate Treatment Effect

	(1)	(2)	(3)	(4)
Treatment	-0.015*** (0.003)	-0.021*** (0.004)	-0.018*** (0.003)	-0.018*** (0.003)
School FE		Y	Y	Y
Time FE		Y		
Time trend			Y	Y
School-term FE			Y	Y
School X Term			Y	Y
School X Trend				Y
Trend X Term				Y
School X Term X Trend				Y
Grade FE		Y	Y	Y
Day of week FE		Y	Y	Y

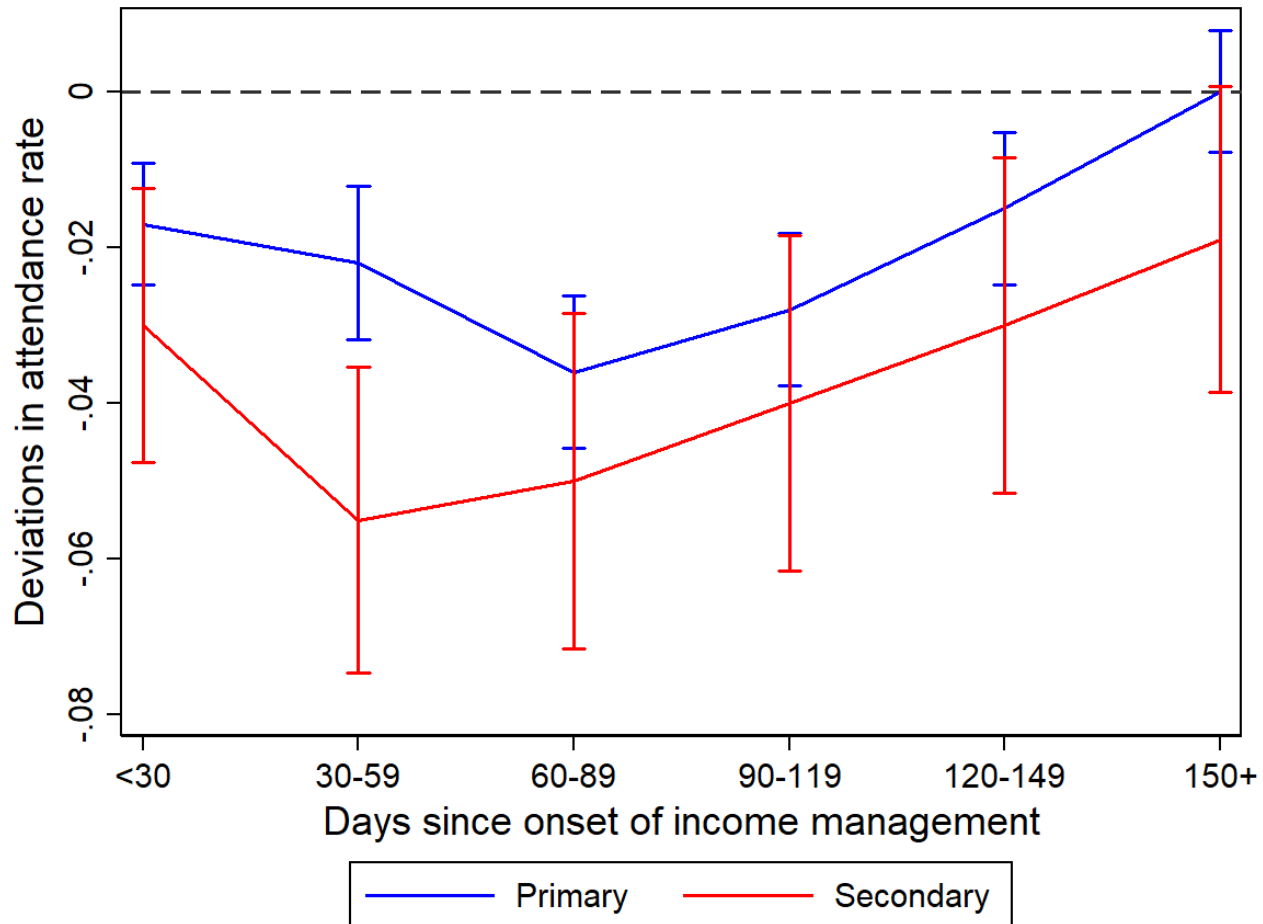
Treatment Effect by Days Since Onset



Treatment Effects by Gender



Treatment Effects by School Sector



Mechanisms

Why did income management lower attendance?

Four possibilities:

1. Confoundedness with the NTER.
2. Changes in school enrolment.
3. Changes in geographic mobility.
4. Implementation issues.

We find little evidence for 1-3 whereas there is evidence for 4.

Implementation issues

- Including a dummy for the Basics Card changes our estimate for 150+ days from **-0.005** to **-0.019***** in our most flexible specification.
- The coefficient on the Basics Card dummy is **0.016*****.

In the medium term, the positive effect of the Basics Card almost fully offsets the negative effect of income management generally.

This indicates that implementation issues are one reason why we observe a decrease in attendance following income management.

Wrap Up

We provide the first causal evidence linking income management to a key policy target – school attendance.

- Income management was associated with a short-term decrease in attendance and return to trend – no positive effect.

Implementation matters. Poorly implemented policy may not only fail to meet objectives, but also have harmful effects.

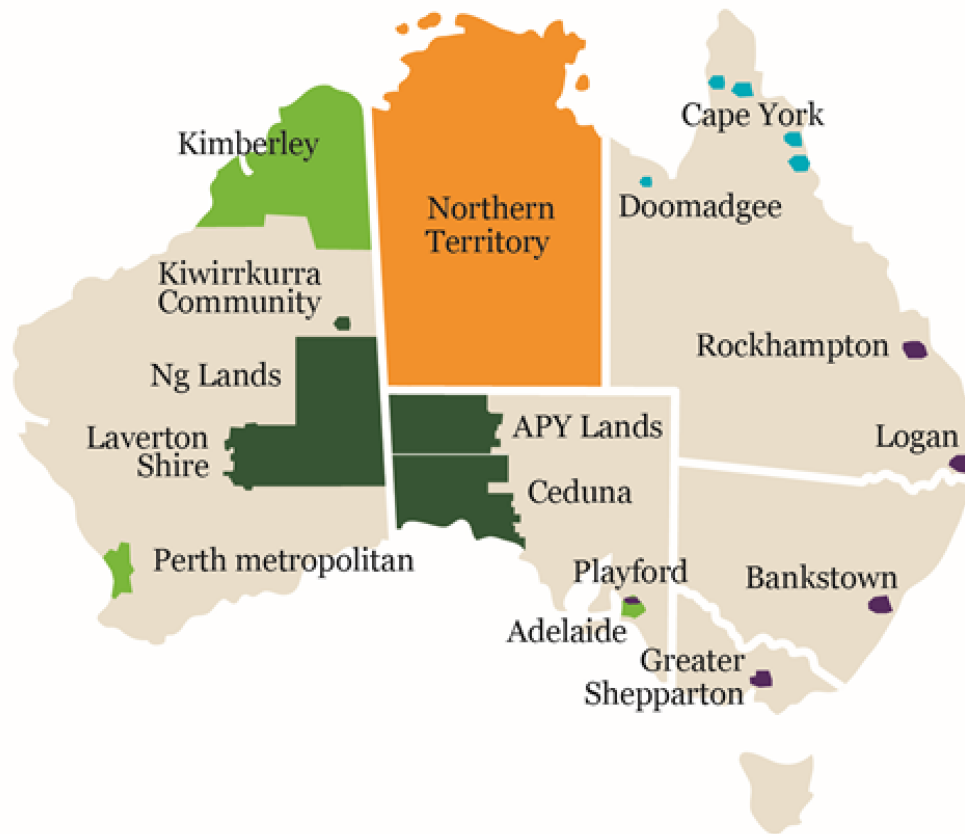
No evidence that in-kind transfers are superior to cash.
However, we only look at one outcome.

Important to evaluate other outcomes of income management, particularly in light of **significant administrative costs.**

- Cost of administering income management was \$451 million (AUD) between the 2007-08 and 2009-10 financial years, or approximately \$20,700 per income managed person.

Wrap Up

Following its use in remote NT Aboriginal communities (2007), income management has become an important policy instrument for the Australian Government:



Thank You

