London School of Economics and Political Science

Beyond birth outcomes: the impacts of perinatal conditions on child development

Mary-Alice Doyle

A thesis submitted to the Department of Social Policy at the London School of Economics for the degree of Doctor of Philosophy

14 February 2025

Declaration of authorship

I certify that the thesis I have presented to the Department of Social Policy of the London School of Economics and Political Science for examination for the PhD degree in Social Policy is my own work other than where I have indicated it is the work of others (in which cases the extent of any work caried out jointly by me and any other person is identified).

I declare that my thesis consists of 59,710 words (excl appendices).

Chapter 2 of my thesis has been published as:

> Doyle, MA, 'Seasonal patterns in newborns' health: Quantifying the roles of climate, communicable disease, economic and social factors', *Economics & Human Biology*, Volume 51, 2023. https://doi.org/10.1016/j.ehb.2023.101287.

Statement of joint work

Chapters 3, 4 and 5 were coauthored. Below are details of the extent of my contribution and that of my coauthors:

- Chapter 3: M-A.D. designed the study, wrote the plan for the analysis, conducted the data analysis, drafted the manuscript, and revised the manuscript following consultation with Bernard Leckning (B.L). B.L. managed the administrative data used in this analysis, as well as ethical review and consultation processes. This statement confirms that M-A.D. contributed 90% of this work.
- Chapter 4: M-A.D. designed the study, wrote the plan for the analysis, conducted the data analysis, drafted the manuscript, and revised the manuscript following consultation with Steven Guthridge (S.G.). S.G. managed the administrative data used in this analysis, as well as ethical review and consultation processes. This statement confirms that M-A.D. contributed 90% of this work.
- Chapter 5: M-A.D. designed the study, in consultation with Stefanie Schurer (S.S.). M-A.D. wrote the plan for the analysis, conducted the data analysis, drafted the manuscript, and revised the manuscript following consultation with S.S. and S.G. S.G. managed the administrative data used in this analysis, as well as ethical review and consultation processes. This statement confirms that M-A.D. contributed 80% of this work.

Data and ethics statement

The research data used in this study was provided by the Child and Youth Development Research Partnership (CYDRP) between the Menzies School of Health Research and the Northern Territory Government. I thank the South Australia-Northern Territory Datalink data integrating authority for facilitating linkage of the multiple datasets made available through CYDRP. The CYDRP First Nations Advisory Group have independently reviewed and endorsed the research to ensure it is respectful of Aboriginal perspectives. Approval for the research has been provided by the Northern Territory Department of Health and Menzies School of Health Research Human Research Ethics Committee (Ref: 18-3261).

Acknowledgements

I would like to thank my supervisors, Stephen Jenkins and Berkay Özcan, for their support and advice throughout my PhD studies. They have pushed me to think about my research from different perspectives and helped me to rein vague ideas into tangible research projects. I really appreciate the time they have both dedicated to helping to improve the work in this thesis. I am also grateful to have been part of the broader Social Policy community at LSE – I have learned so much from being in a multidisciplinary department, especially from my fellow students and friends. Our online co-writing sessions kept me motivated and sane while we were all stuck at home during the lockdowns.

This research was also made possible by many others. First, Stefanie Schurer at the University of Sydney. I had carried the idea of studying the impact of income management since my undergraduate thesis in 2011, and was told by many potential supervisors at the time that it was too hard or impossible. When I brought the idea to Stefanie several years later, she made it possible. Beyond our collaboration on this research, she has been an incredible advocate and mentor to me since we met 9 years ago. I'd also like to thank Steven Guthridge and Bernard Leckning at the Menzies School, who manage the data I use in this thesis. This research would not be possible without the heavy lifting of Steven, Bernard and their team, in building, managing and maintaining the linked data and the secure research environment. They have both helped shepherd each paper through due process, provided input on data and context.

I must also thank my family, and above all, my husband Dilhan, who moved across the world with me so that I could study at LSE. He has not only provided practical and emotional support throughout the PhD journey, but has also effectively become my third supervisor, advising on all things from coding to statistics to interpretation. I also thank my son Ravi, in whose eyes riding the tube to the PhD offices in Holborn is a great adventure, and his enthusiasm is infectious. I'd also like to thank Verity, Geoff, Ellie and Eddie, without whom London would not feel like home. And of course my parents, Elizabeth and Bruce, who are invariably supportive and willing to discuss the minutiae of my research.

Abstract

This thesis investigates the impacts of perinatal conditions on children's health and development, focusing on two exposures in utero: ambient heat exposure and a temporary reduction in means-tested benefits. Existing research demonstrates that perinatal conditions can influence health at birth, as well as health, education, and employment in adulthood. However, less is known about how these processes manifest, if at all, in childhood.

My thesis analyses the impacts of prenatal exposures on outcomes at birth and in childhood, through four empirical chapters. I use linked administrative data from the Northern Territory of Australia, which allows for longitudinal analysis.

Chapter 2 investigates what drives seasonality in birth outcomes, identifying heat exposure as the primary driver. Chapter 3 examines which specific aspects of heat exposure matter most for health at birth. I find that prenatal exposure to both moderate and extreme heat affects newborns' health. In Chapter 4, I find that prenatal exposure to extreme heat increases the risk of hospital admissions in early childhood and lowers test scores at ages 8, 10 and 12. In contrast to my findings in Chapter 3 on birth outcomes, I see little impact in childhood of moderate heat exposure.

Chapter 5 analyses the impact of a reduction in means-tested benefits in pregnancy, finding that this led to increased hospital admissions in childhood. In contrast to the more general increase in admissions from heat exposure, the impact of the reduction in family income was driven by admissions for infections.

Together, my analysis reveals that while the effects of various prenatal exposures may look similar at birth, their impacts in childhood are different. Furthermore, the effects I measure on health at birth do very little to explain or predict effects in childhood.

Contents

Chapter 1: Introduction

The perinatal period – that is, the time from the beginning of pregnancy to soon after birth – is a key period in human development. As Almond et al. (2018) outline in their review, a large body of empirical research has shown that the conditions that children are exposed to in this period can have long-lasting impacts, leading to changes decades later in those individuals' health status, education attainment, and earnings. This evidence has an important practical implication: it suggests that policies and programs that shield parents and babies from harmful conditions during this contained period can have outsized returns in terms of population health, wellbeing and productivity.

Many different types of conditions and experiences in pregnancy and early childhood affect both health at birth, and long-term outcomes. For instance, we know that social programs and policies such as cash transfers (Garfinkel et al. 2022; Bailey et al. 2020), external influences on maternal stress (Black, Devereux, and Salvanes 2016; Aizer, Stroud, and Buka 2016), transmission of disease (Almond 2006; Schwandt 2019), and environmental conditions like heat and pollution (Brink et al. 2024; Martin-Bassols et al. 2024) have been shown to impact children's health at birth, as well as their health and economic outcomes decades later.

In addition to the range of early life conditions that have been studied, the outcomes are also wide-ranging. They include height and stunting (Hu and Li 2019), education attainment (Pacheco and Wagner 2023), earnings (Bailey et al. 2020), and health in adulthood (Martin-Bassols et al. 2024).

Despite these wide-ranging exposures and outcomes studied, this field of research is united in its reliance on the so-called 'fetal origins hypothesis' – that is, the idea that prenatal conditions and experiences affect fetal development in a way that can have long-term impacts. Almond and Currie (2011) set this out in their comprehensive review. As Almond et al. (2018) highlight in their updated review, this body of research has also expanded to consider the postnatal period – and particularly early childhood (ages 0-5) – in addition to pregnancy, as potentially key periods within which both major and subtle changes to a child's environment can have outsized impacts. Hence, we may more accurately refer to this field of research as the 'early origins hypothesis'. Indeed, in public

health research, this field is referred to as 'developmental origins of health and disease'. As Gluckman et al. (2010) explain in their review and conceptual framework, there is no reason why such developmental origins must be limited to the prenatal period.

In their review of previous research, Almond et al. (2018) highlight the 'missing middle' in early origins research. They point out that we now have a large body of research demonstrating the impacts of many types of prenatal conditions or exposures on birth outcomes (e.g. birthweight and preterm birth). We also have a growing body of research demonstrating impacts of prenatal and early life conditions on adult outcomes. But we know little about what happens in between. Only a handful of studies have demonstrated the impacts of specific perinatal exposures on outcomes in childhood, and even fewer are able to analyse outcomes in infancy, early childhood and middle childhood together. Almond et al. (2018) also note that the effects of some prenatal conditions appear to 'fade out' over time. But, equally, in some cases even when researchers detect no impact of a given exposure on health at birth, effects can become evident in adulthood (Schwandt 2019).

If we are to make use of our knowledge of the link between perinatal conditions and long-term health and wellbeing, we need to know more about the developmental trajectories that link the two. Is it that early conditions set off a chain of events and, if we could break the chain, we can limit the longer-term effects? Or do some exposures lead to permanent changes in a child's gene expression which cannot be reversed, but whose effects could be mitigated by targeted intervention? Understanding this would help us identify how to help affected children to catch up. Separately, we also want to know which exposures pose the greatest and most persistent risks, and whether we can reliably detect early effects of those exposures in childhood. Knowing this would make it possible to target remedial care to those who would benefit most from it.

There are a few potential causal pathways that could, in theory, link perinatal conditions to adult outcomes. One possibility is that a given perinatal condition affects adult outcomes because it affects health at birth, and health at birth affects health in early childhood, which limits a child's opportunities to engage in developmentally optimal activities, and so on. Masten and Cinccetti (2010) refer to this dynamic as 'developmental cascades', and explain that this overall model of development sits behind much of the

child development and child psychology literature. If developmental cascades are a causal pathway linking perinatal and adult outcomes, it may be possible to identify early indicators of such effects at work and provide targeted interventions to break the links in this chain.

Another possibility is that a given perinatal condition has biological effects on the child. It could cause congenital conditions or other health conditions, which themselves affect the child's development.

A third possibility is that the effects may come through 'developmental programming': that is, a given exposure 'switches on' certain genes that lead to poorer health or impaired cognitive development. This is the mechanism first suggested by Barker (1998), who proposed that poor nutrition in utero may 'switch on' genes that cause chronic diseases to develop later in life. If this is the case, it may not be possible to reverse long-term impacts after the exposure has occurred. But, as Billing and Schnepel (2018) found in the case of lead poisoning, remedial care could help affected children to catch up to their peers in other ways.

Distinguishing between these developmental pathways has important implications for how we think about the large body of research analysing health at birth as an outcome. Is health at birth a mediator in the chain of events linking a given prenatal exposure to longer-term outcomes? Is it an early symptom of effects to come? Or is it incidental to the long-term effects of a given exposure?

It is likely that both developmental cascades and developmental programming operate simultaneously, for all children – there is strong empirical evidence for both pathways. But different types of conditions and experiences may affect children more through one pathway than the other. Early origins research considers a wide range of causes. While most have been shown to affect the same outcomes at birth (birthweight and preterm birth), it is likely that their effects on children's health and development are not the same. Therefore, while there is value in considering the early origins literature as a whole, to make progress, we must also study whether and how specific exposures affect child development differently from other exposures, and the developmental pathways through which these effects occur.

In this thesis, I analyse two distinct prenatal conditions and their impacts on both health at birth and longer-term outcomes: heat exposure during pregnancy, and a temporary reduction in means-tested benefit income. These are seemingly unrelated – what ties them together is that both exposures have been shown in earlier research to have long-term impacts on a person's health, education and employment outcomes: broadly speaking, their human capital (see Brink et al. (2024) for a review on the impacts of prenatal heat exposure, and Garfinkel et al. (2022) on the effects of income). Both also have effects on children's health at birth (González and Trommlerová 2022; Reader 2023; Chersich et al. 2020). Analysing both conditions in this thesis allows me to study whether and how different prenatal exposures affect children's development differently.

My analysis uses data from the Northern Territory of Australia. My motivation to analyse the impacts of these two exposures in the Northern Territory stems from an observation: the income reduction I study (in Chapter 5) was a landmark policy change which, due to poor implementation (Yu, Duncan, and Gray 2008), led to a sharp but temporary fall in means-tested benefits (Australian Institute of Health and Welfare 2010). Previous research found its impact on health at birth was large, reducing average birthweight by 85 grams (M.-A. Doyle, Schurer, and Silburn 2022). This is a large effect relative to other research on policy changes impacting birthweight. But this was a one-off policy change, with effects aggravated by poor implementation: it is unlikely that later birth cohorts experienced such large effects, even though their parents were subject to the same policy (I discuss this further in Chapter 5). While a single cohort of children were impacted by this policy change, another set of risks occur each year, affecting all children. As I show in Chapter 2, the season a baby is born can have large effects on their health at birth in the context I study: average birthweight is 55 grams lower for a baby born in January than for a baby born in October. This is two-thirds of the magnitude of the previously estimated impact of the policy change.

My investigation into the causes of these seasonal risks was originally motivated by the need to account for them, to facilitate the analysis in Chapter 5. But I began to question: if both an income shock and the season of birth appear to have impacts on birthweight of comparable magnitudes, does this mean that their impacts on children's health and development are also similar?

This is an interesting question academically, filling the 'missing middle' gap I have outlined above in the early origins research. It is also of practical importance. As I show in Chapter 2, a major driver of seasonality in birth outcomes is heat exposure. Exposure to extreme heat at some point during the 9 months of pregnancy is unavoidable in many parts of the world. With the projected impacts of climate change, this is very likely to increase. Regions that are already hot will experience more extreme heat, and currently-mild climates will experience more frequent extreme heat conditions as well (Calvin et al. 2023). In addition, heat exposure will accentuate other seasonal risks, like disease transmission (Mora et al. 2022). Hence, my research is relevant now to the Northern Territory and the many other locations with similar climates. It is also relevant for future cohorts in locations where extreme heat is currently rare but will become more frequent in coming decades. If we are to prevent the adverse effects of climate change on population level health and human capital, we must understand its impacts better.

In Chapter 2, I investigate what drives seasonality in health at birth. Much research on seasonality in health outcomes implicitly attributes it to the weather, or examines other seasonal risks in isolation. But there could be several explanations: weather conditions, seasonal disease, seasonal economic conditions, and seasonality in conception timing – which may differ across socioeconomic and demographic groups. I analyse these factors together, to assess which risks contribute most to the seasonal patterns we observe. I find that by far the most important driver of seasonality in birth outcomes in the Northern Territory is heat exposure, though seasonal disease and economic conditions also contribute.

In Chapter 2, I take a simple approach to measurement of heat exposure, in line with prior literature: I analyse average temperatures in pregnancy, complemented with indicators of extreme conditions. However, it is important to go beyond the conventions of past research in measurement of heat exposure. Although the topic of measurement may seem esoteric, defining measures carefully is an important practical question: it can help us to understand what types of heat exposure pose the greatest risk. It is also an important step to ensuring measures we use capture the full effects of heat exposure.

In Chapter 3, I investigate what it is about heat exposure that affects fetal development, and hence how to measure heat exposure to capture its full effects. Despite the large body

of research showing that heat exposure leads to poorer health at birth, the literature lacks consistent measurement. We do not know, for instance, whether it is exposure to extreme temperatures, to moderate heat, or the confluence of heat and humidity that drives these effects. This means we cannot extract practical recommendations around which heat conditions pose the greatest risks, and hence should be avoided during pregnancy. It also means we cannot predict the implications of climate change on neonatal health and healthcare needs at a population level. Using data from the Northern Territory, I compare the explanatory power of heat exposure metrics that are commonly used in research, alongside additional metrics supported by theory. I find that my conclusions around whether and how much heat exposure affects health at birth would vary dramatically depending on the metric used. Metrics based on 'bands' of exposure and incorporating daily minimum as well as maximum measures provide the best fit; this is consistent with the theoretical understanding that the risks of heat exposure do not increase linearly with higher temperatures, and instead, both moderate and extreme heat affect fetal development in different ways. I find that some measures commonly used in previous research do not have theoretical or empirical basis. In the Northern Territory, using my preferred heat metric, I estimate the impact of prenatal exposure is orders of magnitude larger than would be implied by some metrics commonly used in research. This informs my modelling approach in Chapter 4.

In Chapter 4, I analyse the impacts of prenatal heat exposure on health and education outcomes, from birth to age 12. I find that exposure to unusually hot temperatures in pregnancy reduces school test scores and increases hospital admissions in early childhood. The impacts on hospital admissions fade after age 2, but the impacts on test scores are persistent, evident at ages 8, 10 and 12. Importantly, although heat exposure also affects health at birth, the effects on admissions and test scores cannot be predicted by the initial impacts on health at birth.

In Chapter 5, I return to analysis of the policy-induced reduction in means-tested benefits, against which I compare the impacts of prenatal heat exposure. I analyse the gradual rollout of a policy called 'income management' in the Northern Territory in 2007 and 2008, which represented a change in the way that recipients of government benefits could access their money. Due to implementation challenges, the policy change led to a sharp but temporary reduction in income for some households. We know from past research

that this reduced birthweight for babies in utero during this period (M.-A. Doyle, Schurer, and Silburn 2022); in Chapter 5 I find that these negative effects persist into childhood. The cohort of children who were in utero or newborns at the time of the policy change had more hospital admissions than their peers, spending an average of 4.7 additional days in hospital through to age 8. Most of this impact is manifested in hospitalisations for infections. As with the impacts of heat exposure in Chapter 4, these effects cannot be predicted by initial impacts on health at birth.

In Chapters 4 and 5, I find that the effects of the two prenatal shocks I study are not mediated through birth outcomes. Hence, a prediction of the impacts of either condition on health in childhood, based on the impacts of health at birth, would severely under-estimate the scale of the effect. But, importantly, the effects of these two different exposures on health differ: the income shock led to increases in hospital admissions specifically for infections, and in asthma diagnoses – both of which indicate weaker immunity. This is consistent with the 'immune programming hypothesis' – i.e. the idea that perinatal stress and poor nutrition can lead to permanent changes in the way a child's immune system develops. In contrast, prenatal heat exposure led to an increase in hospital admissions but not in any specific health condition or diagnosis. I interpret this to mean heat exposure affects health in a way that does not generally manifest in acute conditions requiring hospitalisation. Instead, it may increase the risks of complications for conditions that could otherwise be treated at home, making hospitalisation more likely.

There is much to learn from my analysis that is specific to the exposures and outcomes studied, and these are discussed within the relevant chapters. I return to these findings in the Conclusions (Chapter 6) and reflect on what they mean for our understanding of how prenatal exposures affect children. I also draw out some general policy recommendations around administration and implementation of means-tested benefits, and around anticipation and adaptation to the likely effects of climate change. While my research makes an initial comparison of two outcomes, there is still much more we do not know, both about the specific prenatal exposures I study and their mechanisms, and about how these and other prenatal conditions compare with each other. I propose steps for further research.

In the remainder of this chapter, I set out some concepts that inform my analysis. Specifically, I explain in more detail the concepts and literature behind the 'developmental cascades', and 'developmental programming' theories as potential causal pathways linking perinatal exposures with health at birth and children's development. I then turn to the specifics of my research. I give some background information on the Northern Territory of Australia and its population, which is the setting for all four empirical chapters, and I describe the linked administrative data I use. While each empirical chapter explains its methods in detail, there are commonalities which I set out in the introduction. Based on this, I outline the contributions this thesis makes to our knowledge base.

1.1 From perinatal exposures to childhood: potential pathways to long-term impact

The causal mechanisms linking perinatal exposures to longer-term human capital outcomes are complex and interrelated. As a result, theoretical explanations span multiple dimensions. Here I give a high-level summary of theories around the persistence or fade-out of early life health inequalities and relatedly, explanations for why the perinatal period is so important. I then set out three broad theoretical pathways linking perinatal risk factors with a child's development.

1.1.1 Fade-out vs self-reinforcing health stocks

In the field of economics, Grossman (1972) pioneered the consideration of health as a form of human capital. Although his model does not consider health in childhood, if we were to apply it to prenatal conditions, we may consider health at birth as a 'stock' of health with which individuals are endowed. Individuals (or, more likely in this case, their parents) can make investments to maintain good health, though their health stock depreciates over time. A range of factors, including education, can make investments more productive.

As Almond and Currie (2011) show, viewing health simply as a depreciating 'stock' would imply that the impacts of health at birth should fade over time, and would be

negligible by mid-adulthood. This fade-out hypothesis appears true in some cases. For instance, Cozzani et al. (2021) find that British children with low birthweight have statistically significantly lower cognitive scores than other children at ages 3, 5 and 7, but these differences are almost entirely attenuated by age 14 after accounting for family characteristics; Boardman et al. (2002) find the same pattern in the USA.

However, there are also many empirical studies that find persistent effects of perinatal conditions on adult outcomes. Heckman's (2007) alternative theoretical framework may better explain these findings, as his concept of 'self-productivity' allows for early life health setbacks to be self-reinforcing. Applied to analysis of the prenatal environment, the idea of 'self-productivity' implies that a higher stock of health at birth can mean future investments are more productive. In such a framework, the effects of prenatal and perinatal exposures do not necessarily fade out – they may lead to greater inequalities between those exposed and unexposed over time.

1.1.2 Sensitive and critical periods

Closely related to the question of whether early life health is self-reinforcing or fades out is the idea of sensitive and critical periods. A sensitive period is one within which investments may have particularly high returns, but low investment can be compensated for later. In contrast, a critical period is a specific period during development, within which if a child experiences some harmful shock or does not receive all that is needed (e.g. adequate nutrition), there may be irreversible effects. The possibility that there are critical periods in child development fits with the idea of 'developmental programming', which I explain in more detail below. While the prenatal period is generally considered to be critical, it is not the only such period. For instance, in their review of the impacts of prenatal and early life nutrition, Marques et al. (2013) present evidence that conditions and exposures in the newborn stage can have similar 'programming' effects to prenatal exposures.

Whether a given shock is persistent or fades over time, or even the degree to which it contributes to 'self-productivity', may depend on the nature of the shock. But it may also depend on the timing: if it occurs during a critical period, its effects are more likely to be permanent than if it occurs at another point in time – this is the source of variation that I use to identify the effects of an income shock in Chapter 5.

1.1.3 Prenatal and perinatal environments

Throughout this thesis I use the terms 'prenatal' and 'perinatal'. By perinatal, I mean the three trimesters of pregnancy and the period shortly after birth – sometimes referred to as the 'fourth trimester'. By prenatal, I mean the three trimesters of pregnancy.

My main focus is on the prenatal environment. This is by necessity in Chapters 2 and 3, as I analyse the impact of the prenatal environment on birth outcomes, and hence exposures after birth are irrelevant.

However, in Chapter 5, I analyse the impact of a perinatal shock – that is, one occurring in pregnancy or as a newborn. I discuss the reasons for this in more detail in that chapter, and present analysis that shows impacts of prenatal and newborn exposure separately. In short, the inclusion of the 'fourth trimester' in that analysis relates to the discussion of critical and sensitive periods. For the nutrition and maternal stress mechanisms that I expect to be at play in explaining the impact of an income shock on children's health, these exposures, while qualitatively different after birth, can continue to affect a baby's development. There is good reason to expect that the critical period during which these specific risks can affect a child's development extends into the newborn period.

The concepts of self-productivity and critical periods are instructive but are somewhat abstract. At a more granular level, it is helpful to understand how a given prenatal shock may affect children's early health and hence their longer-term development. I now turn towards a consideration of possible causal pathways, which draw on these theories, alongside others.

1.2 Causal mechanisms

In this thesis I analyse three sets of outcomes: measures of health at birth (e.g. birthweight, preterm birth, and others explained in the relevant chapters), health in childhood (measured as hospital admissions and primary care use), and cognitive development (measured as school test scores).

There could be several potential mechanisms linking prenatal and perinatal conditions to these outcomes. Here I set out these potential mechanisms, drawing on concepts from the fields of economics, public health, and developmental psychology.

A first point to make is an obvious one on composition: when we are observing population-level exposures and changes in average outcomes, we can think of two ends of the spectrum, in terms of how effects are distributed. Let us consider an exposure that is shown to reduce average test scores by 1 percent in a population of 100 children. It may be that all children are affected similarly by the exposure, and it leads to a 1 percent reduction for each child. Conversely, it may be that the exposure has no effect on most children but reduces one child's test scores by 100%. Both would lead to the same average change in test scores, but they have very different implications for our understanding of how that exposure affects children, and what could be done about it. The impacts of most prenatal exposures likely sit somewhere in the middle of these two extremes. But it is worth considering whether average impacts we observe are the result of acute impacts felt by a minority. In my research, such extreme effects would likely be the result of severe health conditions or developmental delays. This is a particular focus in public health research, where researchers investigate links between in utero exposures and the risk of specific congenital conditions. For instance, several studies draw a link between prenatal heat exposure and congenital conditions such as heart defects (Auger et al. 2017) or neural tube defects (Bruckner et al. 2024).

A second point, as alluded to above, is whether the effects of a given exposure are incremental, with a chain of events mediating the links between the original exposure and the final outcomes in adulthood. The concept of 'developmental cascades' (as set out by Masten and Cincchetti (2010)), is a helpful framework in considering such a chain of events. It may be the case, for instance, that a specific exposure in utero leads to a higher risk of preterm birth, preterm birth leads to higher risk of recurrent infections in infancy, more time spent recovering from infections as an infant leads to less exposure to developmentally beneficial interactions and experiences, and so on, with ultimate impacts on a child's cognitive skills.

In this framework, it could well be that two different prenatal conditions, if they both affect health at birth in the same way, have similar effects on longer-term outcomes.

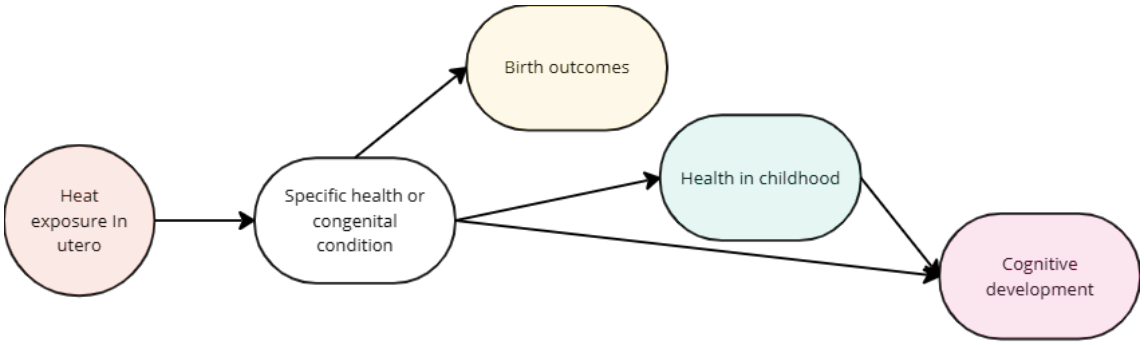
While not always explicitly stated, much policy discussion implicitly adopts this framework. For example, the UK House of Lords' Preterm Birth Committee (2024, 17) considers that preterm birth itself is a "leading cause" of child mortality, disability and cognitive impairments.

An alternative possibility is more deterministic: that the prenatal environment has direct effects on later outcomes, for which certain earlier developmental measures may be symptoms, but do not mediate the effect. This is the idea of 'developmental programming (also referred to as 'fetal programming' in economics) or, as it is referred to in the public health literature, 'developmental origins of health and disease'.

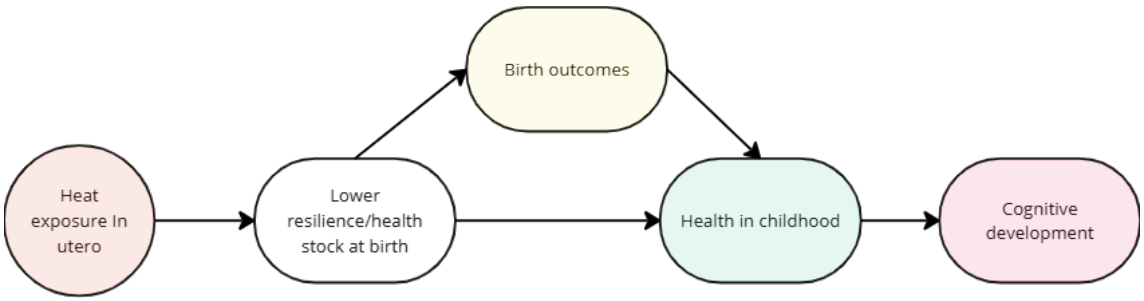
In Figure 1, I illustrate these possible causal pathways and their implications for the links between prenatal conditions and the childhood outcomes I analyse, in relation to heat exposure. Figure 1 is not intended as a set of directed acyclic graphs ('DAGs') or nested hypotheses for which it would be possible to empirically estimate a full model and 'test down' to settle on a single pathway. Rather, these are schematic diagrams to illustrate potential links (or lack thereof) between the three sets of outcomes that I analyse in this thesis.

Pathway 1: congenital or specific health conditions

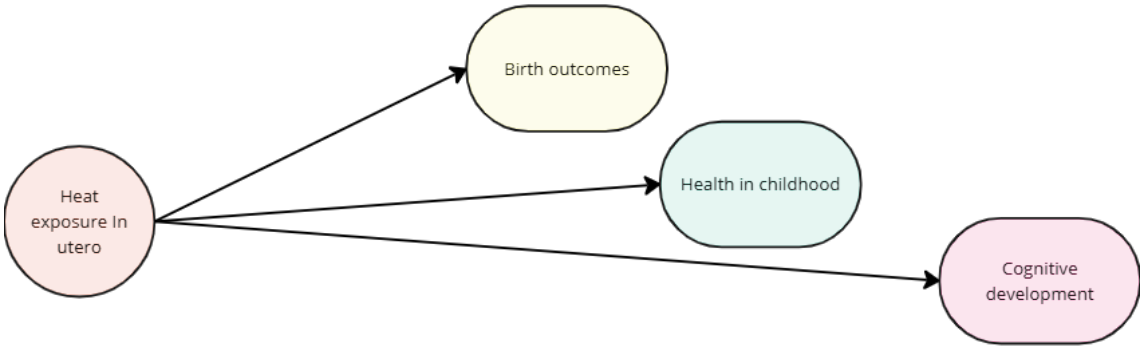
As described above, prenatal conditions may lead to certain congenital or other specific health conditions. Children with these conditions are more likely to be born with poor health at birth, though this is a symptom of the underlying condition, and not itself a cause of any impacts on health and cognitive development. These same conditions can also lead to poorer health, and hence increased interactions with the healthcare system, in childhood. Certain congenital conditions may themselves affect cognitive development directly (Bonthrone et al. 2021), though as shown in Figure 1(a), there may be some element of 'self-productivity', with poorer health limiting children's ability to engage in their education.



(a) Pathway 1: Congenital or specific health conditions



(b) Pathway 2: Developmental cascade



(c) Pathway 3: Developmental programming

Figure 1: illustration of three potential pathways explaining the link between heat exposure (as an example of a prenatal condition) and outcomes in childhood
Source: Authors' analysis

Pathway 2: developmental cascades

There is much research on the impacts of a wide range of prenatal conditions and exposures on health at birth. Many conditions have been shown to cause children to be born smaller or earlier than they otherwise would be. It could be that the birth outcome itself (e.g. preterm birth) causes poorer health. It could also be caused by some latent underlying health status at birth, for which preterm birth is a symptom. Compared with pathway 1 above, pathway 2 brings forward the point at which 'self productivity' might begin to have effects. In this scenario, better health at birth means that other 'investments' are more productive. It may be, for instance, that children are more receptive to experiences that enhance their cognitive development if they are in good health.

We may also think about this in terms of 'developmental cascades', which could explain interrelations between different domains of child development.[1] For instance, children developing otitis media (a middle ear infection) at a young age – which is particularly common in the Northern Territory – are at greater risk of speech and language problems. As Williams and Jacobs (2009) outline, impaired hearing and the resulting delays in speech and language can lead to poorer cognitive and education outcomes. This is because, as Peng et al. (2024) describe, delayed language development in early childhood causes conduct problems, and when children start school, conduct problems impede their ability to learn. Therefore, as shown Figure 1(b), while the in-utero environment may have no direct impact on health in childhood and cognitive abilities, it may indirectly affect these outcomes through its impact on resilience or health stock at birth. Though self-productivity and developmental cascades represent different theories with their own nuances, they fit within the broad prediction of path-dependence, with long-term effects mediated by early life health.

Pathway 3: developmental programming

Finally, it may be that prenatal conditions affect health at birth, health in childhood and cognitive development, but that these are not directly related to one another, and do not necessarily occur in the same children. It may be that for some children, a given prenatal condition affects their health in early life, and for other children it may have epigenetic

[1] As a theory, Masten and Cincchetti (2010) argue that developmental cascades may incorporate developmental programming as an additional level of effect which itself may set of a chain reactions with behavioural consequences. Here I make the distinction between 'cascade' effects and epigenetic effects.

effects inhibiting their cognitive development. These differences may reflect different mechanisms underlying the effect itself. As an example of this, I explain in Chapter 3 how the impacts of heat exposure on the fetus may be different depending on whether the exposure leads to heat strain, heat stress, or changes in maternal behaviour. As shown in Figure 1(c), these effects need not be self-reinforcing: unlike for the other pathways, health does not necessarily mediate the impact on cognitive outcomes.

1.2.1 Why does identifying the causal mechanism matter?

Distinguishing which of these three pathways best reflects the impacts of a given prenatal condition or exposure is important because the appropriate response differs. If impacts come through specific health conditions, attention could be focused on prevention and early treatment of those conditions. If the impacts are due to cascade effects set off by poor health at birth, intensive early healthcare interventions could help. If the effects come through developmental programming, this may mean both that they are irreversible, and that the effect of the exposure on one outcome (e.g. cognitive development) cannot be predicted by early indicators on another (e.g. health). However, this does not mean that nothing can be done: interventions could help children to work around these 'programmed' effects – in a similar way, for instance, to successful treatments for children with neurological damage due to lead exposure (Billings and Schnepel 2018).

Of course, these three pathways are not necessarily independent of each other. They could be 'stacked' together and interact: it may be that specific health conditions lead to developmental cascades. Perhaps these are accentuated by developmental programming. And as Almond, Curie and Duque (2018) point out, there may be heterogeneous effects of the same exposure, as parents may respond differently to the same conditions, based on their knowledge, their budget constraints, and their own 'parenting technology'.

That said, which of these pathways is activated by a given prenatal exposure may vary based on the nature of the exposure. If this is the case, it would be very useful to know whether, for instance, a specific shock leads to a condition that can be observed and treated early in life to prevent cascade effects, or whether, through its impact on gene expression, it leads to disadvantages – like impaired immunity – with potential lifelong effects.

1.3 Context and data

1.3.1 The Northern Territory

All four empirical chapters in this thesis analyse data from the Northern Territory of Australia. The Northern Territory is one of Australia's eight states and territories. It is located in the central part of northern Australia (see Figure 2). The Northern Territory is sparsely populated, with around 233,000 residents, approximately 60 percent of whom live in or around the capital city, Darwin.

Aboriginal people comprise approximately one-third of the territory's population. Eighty percent of Aboriginal residents live outside Darwin, primarily in communities based on established kinship networks. Many of these communities are very remote, and residents have limited access to healthcare, quality housing, and economic opportunities. Median income for Aboriginal people in the Northern Territory is one-quarter of that of the non-Aboriginal population, with approximately 45 percent of Aboriginal households below the national poverty line (Altman 2017). It is because of these stark differences that in Chapters 2, 3 and 4, I conduct heterogeneity analysis to investigate whether the relationships I identify differ between Aboriginal and non-Aboriginal children. Chapter 5 focuses solely on children in Aboriginal communities.

Figure 2: Map of Australia with Northern Territory highlighted, and indicating Darwin (the capital city of the Northern Territory)
Source: simplemaps.com

I study the Northern Territory because this is where income management, the policy which is the focus of Chapter 5, was first introduced. Income management has since been introduced in specific localities in other states of Australia, but most people affected by this policy are Aboriginal people in the Northern Territory.

While this research focuses on a specific sub-population, it is an instructive example more generally. Income management is similar to major programs in other countries, such as the Supplemental Nutritional Assistance Program (SNAP, formerly 'food stamps') in the United States. The challenges surrounding the rollout of income management – which is the focus of my analysis – reflect more general implementation challenges in payment administration, relevant to any agency administering means-tested income support. As I explain in Chapter 5, my analysis from the Northern Territory provides a case study in how such implementation challenges can affect the most disadvantaged communities.

The Northern Territory is also a particularly interesting setting in which to study the impacts of seasonal risks and heat exposure on health outcomes. This is because it spans both tropical and arid climates zones. The climate is not in itself unique: half of the world's population lives in tropical or arid climates (United Nations Environment Management Group 2011; State of the Tropics 2020). What is unique is the opportunity to analyse the impacts of seasonal risks and heat exposure in these climates using high quality, longitudinal data covering the full population.

Most research on heat exposure is about countries with milder climates. As I explain in Chapter 2, some of the seasonal risks that residents in tropical and arid climates face differ from milder, temperate climates: monsoonal rain and flooding, frequent exposure to extreme heat, and certain seasonal diseases. Furthermore, as I show in Chapter 3, the relationships between alternative measures of heat exposure differ in tropical climates, making it all the more important to carefully consider which measures we use.

While my findings are most likely to translate to contexts with similar climates, the fact that I can study the impacts of extreme heat conditions is also instructive more broadly. The extreme heat conditions currently experienced in arid and tropical climates are likely to become more common even in currently mild climates due to climate change. It will be important to be able to predict their effects on population levels of health, healthcare demand, and education.

1.3.2 The data

In all four empirical chapters, I use linked administrative data, covering every child born in the Northern Territory and their interactions with the public health and education systems.

Most of my research focuses on health outcomes, proxied by hospital records. Australia has a public healthcare system. Care is provided free of charge at the point of service, funded through taxes. In this thesis I analyse hospital admissions, as well as community clinic use in Aboriginal communities. These cover the majority of healthcare encounters in the populations I study. [2]

A limitation of this dataset is that because it comes from the Territory government, it does not include information on benefit receipt or income (in Australia, taxes and benefits are administered by the federal government). This means I have limited socio-demographic information, and cannot, for instance, analyse outcomes separately by socioeconomic status, benefit receipt or income. I also cannot track interstate migration: this is most relevant to Chapter 4, and I discuss potential attrition in detail there.

1.3.3 Methods

My analytical methods and specifications vary, but across all empirical chapters, I analyse data using fixed-effects regression estimators – that is, regressions that control for small area location and time. This approach adjusts for time-invariant and location-invariant unobserved characteristics. Using location fixed effects means that I analyse changes within towns or communities, holding characteristics like climate, demographic characteristics, and public infrastructure fixed. Most of my research uses data spanning a decade or more – using time fixed effects means I can abstract from trends over time, which may be caused by changes in the policy environment or in healthcare practices, for example.

[2] Australia also has a parallel private healthcare system. But there is just one private hospital in the Northern Territory which does not provide comprehensive care, and is co-located with a public hospital. The number of children admitted to this hospital is very small. The remote communities for which I use clinic data do not have private clinics.

The fixed effects regression approach is particularly helpful when studying a heterogeneous region like the Northern Territory, where standard regression techniques that do not control for small area location and year fixed effects might conflate underlying differences and changes related to climate zone, Indigenous status, and public infrastructure with the effects of specific interventions or exposures.

1.4 Contributions of this thesis

Each empirical chapter in this thesis speaks to a slightly different literature, with specific contributions discussed in each chapter. But as detailed above, there is a common thread across all four chapters: I am interested in what types of exposures affect fetal development, how this manifests in health at birth, and whether and how these effects persist in childhood.

My thesis also speaks to two other broad themes: the implications of climate change on population health and human capital, and benefits policy and implementation, seen through an 'early childhood lens'. I set out these themes here, which I will return to in my conclusions (Chapter 6).

1.4.1 Early origins of human capital

Across my four empirical chapters, there are three specific questions within the literature on the early origins of human capital to which I aim to contribute.

First, as described above, I aim to test whether two different changes to the prenatal environment (heat exposure and an income shock) – which have similar effects on health at birth – have different effects in childhood.

Second, a recurring theme is on whether and to what extent 'birth outcomes', such as birthweight and preterm birth, should be considered outcomes at all. As Romero et al. (2014) put it, preterm birth represents 'one syndrome, many causes'. Conti et al. (2020) make a similar point in relation to birthweight: health in utero and at birth is multidimensional, and no single measure can proxy all of these dimensions. One solution is to consider multiple measures. As I argue in Chapter 2, some additional measures like

an Apgar score (itself based on multiple dimensions), can help to provide a more comprehensive picture of newborn health.

We know from prior research that many different types of risks cause poorer health at birth. As I discuss in Chapter 3, even the same risk ('heat exposure') can affect fetal development through many mechanisms. We also know that poorer health at birth is linked to poorer health and human capital outcomes in adulthood. But it does not necessarily follow that every risk that affects health at birth has long-term effects. Relatedly, it may be that certain prenatal conditions do not consistently lead to lower birthweight or to preterm birth, but can still have long-term consequences. We see this, for instance, in Schwandt's (2019) analysis of prenatal exposure to influenza – he finds that babies exposed in the second trimester of pregnancy had no observable differences in health at birth, but had lower average earnings and were more likely to claim means-tested benefits in adulthood. This suggests that influenza exposure affected fetal development in a way that did not affect birth outcomes.

In Chapters 4 and 5, I provide new evidence on these dynamics. I test whether controlling for both standard measures of health at birth (preterm birth, birthweight) and a broader set of measures attenuates the impact of the two risks I study on outcomes in childhood.

Third, as described in detail above, I discuss evidence for the three pathways through which the perinatal environment may have impacts in childhood. Of course, all three – as well as other processes – are likely at play simultaneously. In Chapters 4 and 5, I discuss how my findings fit with theories of developmental programming, compared with alternative causal pathways.

1.4.2 Climate change

In Chapters 3 and 4, I focus on the impacts that prenatal heat exposure had on the cohort of children born from 2000 onwards, who are now young adults. The Northern Territory has always had a hot climate but, since this cohort were in utero, average temperatures in the Northern Territory have risen and are projected to rise further (NESP Earth Systems and Climate Change Hub 2020). Even if global efforts to limit the increase in temperatures to 1.5 degrees are successful, this average will not be spread evenly. We are likely to see more extreme temperatures and weather events globally, but locations with

hot climates, like northern Australia, are likely to see average temperatures rise by more than 1.5 degrees (CSIRO and Bureau of Meteorology 2022).

We already know that heat exposure adversely affects human health, and the impact of heat exposure during pregnancy on newborns' health is particularly well-studied. But what we do not know is what it is about heat exposure that affects fetal development. Every review on this topic, including Chersisch et al. (2020), Brink et al. (2024) and Baharav et al. (2023), has lamented the inconsistency in the ways researchers measure heat exposure – making it very difficult to draw firm conclusions and estimate the precise effects of different aspects heat exposure.

Given the projected increase in prevalence of extreme weather conditions, it will be important to understand what it is about heat exposure that affects fetal development. Is it prolonged exposure to moderate heat, or extreme temperatures, or something else? Knowing this could help healthcare professionals provide more targeted advice to pregnant people. If the effects of moderate and extreme temperatures are different, then prior research based on average temperatures may be less informative.

This is why it is important that we know what it is about heat exposure that affects fetal development most. Chapter 3 addresses this. While I cannot provide a definitive answer from the single context I study, I present an approach that can be applied in other contexts in future research, and some general recommendations.

In Chapter 4, I analyse the impacts of heat exposure in utero on health in childhood and school test scores. This helps to fill in the 'missing middle' (Almond, Currie, and Duque 2018) in evidence, between impacts on health at birth and adult measures of human capital. As I outline in that paper, the 'missing middle' is particularly acute in research of the impacts of prenatal heat exposure. My analysis of comprehensive linked data, which allows for longitudinal analysis, in a location with a hot climate where the effects of extreme heat can be measured, is a unique contribution.

1.4.3 Benefits policy and early childhood lens

Many commentators conflate policies to boost early childhood development with early childhood education. This is for good reason – there is strong evidence, for instance, that

quality early education and care can have lifelong effects on children (Heckman 2007). But there is growing appreciation that a much broader set of factors, including environmental policy, economic conditions, the tax and transfer system, and the built environment have important impacts on children's early development.

The USA's National Scientific Council for the Developing Child (2023) argues that all policy should be considered through an 'early childhood lens'. That is, the likely effects of a policy change on children's development should be specifically considered, even if the policy is not 'about' children. We know, for instance, that tax and transfer policies affecting parents' incomes have the potential to improve children's human capital – with studies showing that over the long-term, this type of public spending may pay for itself. This is because such programs can improve children's health, boost their education attainment and earnings potential, and even reduce the risk of criminal activity (Garfinkel et al. 2022; Bailey et al. 2020).

Research on these effects focuses on the impacts of landmark policies affecting families' financial resources, such as the introduction of public health insurance, earned income tax credits, and new benefit entitlements. But it is important to also consider more subtle policy changes. In particular, we have seen an increase in conditionality of means-tested benefits across OECD countries in recent decades (Knotz 2018). Such conditionality seeks to impose rules either on how recipients spend their money, and how they spend their time. We know little about how such policies affect the children of benefit recipients.

I study one such policy change in Australia – the introduction of income management. Challenges in policy implementation meant that recipients had difficulties adapting to the new way that their benefits were paid. They were also affected by delays in receiving payments, and experienced stress caused by poor official communication. This evaluation is, of course, specific to an Australian policy, and furthermore, specific to the way it was rolled out. We can only speculate on the longer-term impacts of the policy once implemented. But such implementation challenges are far from unique to the policy I study. Conditionality is inherently more expensive and complex to administer than unconditional cash transfers (Margolies and Hoddinott 2015). When the US SNAP program moved from physical food stamps to its current electronic benefits card, this led to a temporary spike in recipients' food insecurity (Lovett and Xue 2017). And when the

UK transitioned its Healthy Start food voucher scheme from paper vouchers to a debit card, many recipients were temporarily unable to access this support (Defeyter et al. 2022). We therefore already know that these types of 'conditional welfare' policy changes can be disruptive due to implementation challenges. What we do not know is how this type of disruption affects children's development – in Chapter 5 I address this question. When we see these disruptions through an early childhood lens, even if implementation challenges are resolved, their impacts are not necessarily temporary.

1.5 Outline of thesis

This thesis consists of four empirical chapters. In Chapters 2 and 3, I focus on understanding the link between season of birth and health at birth. In Chapter 2 I investigate the drivers of seasonal variations in birth outcomes, examining weather conditions, seasonal diseases, and economic factors. In Chapter 3, I seek to understand what it is about prenatal heat exposure that affects health at birth, and hence how researchers should measure it. In Chapter 4, I extend this analysis by examining the longer-term health and education impacts of prenatal heat exposure from birth to age 12. In Chapter 5, I shift focus to analyse the effects of a policy-induced perinatal income shock, investigating its implications for child health through age 8. Each chapter is followed by a set of appendices that provide additional methodological details, robustness checks, and supplementary analyses. Chapter 6 contains my conclusions, in which I revisit the contributions outlined above in light of my findings, and suggest productive avenues for future research.

Chapter 2: Seasonal patterns in newborns' health: quantifying the roles of climate, communicable disease, economic and social factors

Abstract

Poor health at birth can have long-term consequences for children's development. This chapter analyses an important factor associated with health at birth: the time of year that the baby is born, and hence seasonal risks they were exposed to in utero. There are multiple potential explanations for seasonality in newborns' health. Most previous research has examined these in isolation. We therefore do not know which explanations are most important – and hence which policy interventions would most effectively reduce the resulting early-life inequalities. In this chapter, I use administrative data to estimate and compare the magnitudes of several seasonal risks, seeking to identify the most important drivers of seasonality in the Northern Territory of Australia, a large territory spanning tropical and arid climates and where newborn health varies dramatically with the seasons. I find that the most important explanations are heat exposure and disease prevalence. Seasonality in food prices, employment conditions, and road accessibility have smaller effects on some outcomes. Seasonal fertility patterns, rainfall and humidity do not have statistically significant effects. I conclude that interventions that protect pregnant women from seasonal disease and heat exposure would likely improve newborn health in the Northern Territory, with potential long-term benefits for child development. It is likely that similar impacts would apply in other locations with tropical and arid climates, and that, without action, climate change will accentuate these risks.

2.1 Introduction

There is much evidence showing the long-term consequences of poor health at birth. Children with poor health at birth are less likely to do well at school (Bharadwaj, Eberhard, and Neilson 2018), less successful in the workforce (Currie and Rossin-Slater 2015), and more likely to face chronic disease in adulthood (Risnes et al. 2011). These findings stem from the broader 'fetal origins hypothesis' literature, which finds that conditions in utero – often proxied by health at birth – can determine childhood and adult outcomes (see Almond et al. (2018) for a review).

But what determines a newborn baby's health? The answer can help policymakers and health care workers to anticipate which pregnancies are at risk of adverse outcomes and provide appropriate preventative care. In this chapter, I focus on one important factor associated with newborns' health: the time of year that the baby is born, and hence the seasonal influences they are exposed to in utero.

A large body of research shows that birth outcomes are seasonal. However, the precise patterns vary across locations, suggesting that the reasons for seasonality depend on both the climate and on other contextual factors (Strand, Barnett, and Tong 2011b).

In this chapter I analyse seasonality in a region with both tropical and arid climates. Studies from these climates tend to focus on agricultural explanations for seasonality, highlighting the effect of rainfall on the agricultural production cycle, and hence on maternal labour, income and food consumption (Chodick et al. 2009; Maccini and Yang 2009). While agricultural cycles are an important explanation for seasonality, research from other climates finds other explanations, and these are also likely to be present in tropical and arid climates. For example, we may expect seasonality in the characteristics of parents who conceive at different times of year, in disease prevalence, and in biological responses to very hot or cold weather during pregnancy (Currie and Schwandt 2013; Strand, Barnett, and Tong 2012). Also, seasonal weather events such as flooding may cause maternal stress and limit access to healthcare and fresh food, impacting newborn health independently of any impact on agricultural production.

This study analyses seasonal patterns in birth outcomes in the Northern Territory (NT) of Australia. Birth outcomes in the NT exhibit very substantial seasonal variations, with the

average difference in birthweight between the highest and lowest month in the NT being twice as large as the difference that Currie and Schwandt (2013) find in the USA, more than five times as large as what Torche and Corvalan (2010) find in Chile, and between six and 30 times larger than what McGrath et al. (2005) find in other parts of Australia. It is important to understand the reasons behind these very substantial seasonal patterns so that these early life health inequalities can be prevented in future.

This chapter makes three contributions to our understanding of seasonal patterns in newborns' health.

First, it compares the magnitudes of effects of different seasonal risks. Many researchers have considered the effects of individual seasonal risk factors on birth outcomes – predominantly rainfall, disease or heat (Maccini and Yang 2009; Andalón et al. 2016; Dorélien 2019). Those studies provide valuable, detailed analysis on the nature of, and mechanisms behind, specific seasonal risks. This chapter brings together evidence from this large body of work to answer a higher-level question: which seasonal risks are the most important determinants of newborns' health in the NT? The answer can help policymakers to identify the most effective means of addressing early life inequalities.

Second, it contributes evidence on the seasonal risks in tropical and arid climates, which are home to around half of the world's population.[3] While the relative importance of seasonal risk factors is likely to vary from country to country, some of the risks I analyse are climate-specific, such as very hot weather in arid climates and annual flooding in tropical climates. My estimates may therefore be indicative of the impacts of risks experienced by much of the world's population. The NT is a valuable context to study to for this purpose: rich administrative data allow me to analyse relationships that likely exist in many parts of the world, but for which data are not available. Importantly, my estimates may be a lower bound on the magnitudes of these effects, given that Australia is a high-income country with free public healthcare, meaning the population have greater access to climate adaptations, treatment and preventative care than most residents in lower-income countries located within tropical and arid climate zones.

[3] Approximately 43% of the world population lived in tropical areas (State of the Tropics 2020), and 35% live in deserts and drylands (including hyper-arid, arid, semi-arid and subhumid climates) (United Nations Environment Management Group 2011, 27). However, there is some overlap between these definitions.

Third, the rich administrative dataset that I use allows me to make several contributions in measurement and methodology, which can inform future research. I analyse birth outcomes for all children born within the NT over a 10-year period. The data contain gestational age, birthweight and Apgar scores[4] for all children, which I use to analyse how seasonality differs across these outcomes, and throughout the distribution of these outcomes. In contrast, most other studies focus only on average birthweight or a binary indicator of preterm birth (see 2.3.2). Additional administrative datasets from the NT allow me to analyse potential explanations of seasonality that are predicted by theory (see 2.2), but thus far have not been documented empirically. I am also able to link together siblings to estimate sibling fixed-effects models, which allow me to disentangle environmental effects from the effects of time-invariant maternal characteristics.

2.2 Background and conceptual framework

Much of the literature on seasonal and environmental determinants of health at birth considers individual seasonal risk factors. However, seasonal risks are often correlated. For instance, some diseases are most prevalent during the hottest months of the year, and both disease and heat can contribute to poorer newborn health. Therefore, in an analysis of the effect of heat exposure which does not consider concurrent disease prevalence, we may mistakenly conclude that it is a physiological response to heat, and not disease, which drives patterns in newborn health. Furthermore, there is the challenge that pregnancy spans multiple seasons. Without careful analysis, this means it is difficult to know, for example, whether it may be cold weather during the first trimester, hot weather during the third trimester, or something else entirely, that is driving worse outcome for babies born in the summer/wet season.

When our goal is to predict patterns in newborn health, this potential misattribution may not be a problem. However, attribution becomes important when our goal is to identify

[4] Apgar scores were developed by Virginia Apgar, and represent an index combining the birth attendant's observations of: skin colour, heart rate, reflex, muscle tone and respiratory effort. A birth attendance scores each on a scale of 0 to 2, 5 minutes after the baby is born.

effective policy responses. The potential policy responses to disease prevalence are quite different from responses to heat exposure.

I aim to disentangle competing explanations for seasonality in birth outcomes and estimate their relative contributions to newborns' health in the Northern Territory of Australia. These explanations can be classified into four broad categories: weather-related, disease-related, economic conditions, and differential fertility patterns across socioeconomic groups. In the remainder of this section, I first outline some context around the Northern Territory and the seasonal risks present in this region, I then describe the existing evidence for each category of seasonal risk (see Appendix A for a graphical summary).

2.2.1 The Northern Territory

The Northern Territory (NT) is one of Australia's eight states and territories. It is a large and sparsely populated region, covering the central part of northern Australia. The NT had a population of around 210,000 people as at the 2011 Census, of whom just over half live in Darwin, the capital city in the tropical north of the NT. Residents have access to Medicare, Australia's universal public healthcare system which allows them to access primary care, perinatal care and public hospital services at no cost.

The NT has two distinct climate zones: tropical and arid. The north – which is more densely populated – has a tropical climate, with a wet season (November-April) and a dry season (May-October). During the wet season there are monsoonal rains, causing flooding and limiting road access to certain towns and communities. It is also particularly humid early on in the wet season, before heavy rain begins (sometimes called the 'build up').

The central and southern parts of the Northern Territory are made up of desert and grasslands; an arid climate, with very hot summers and mild or cold winters. While average temperatures are similar across the tropical and arid zones, the range varies, with temperatures reaching 40 degrees Celsius in the height of summer in the tropical zones, compared with up to 45 degrees in the arid zones. These climate zones contrast with the rest of the country, with all other major Australian cities falling into temperate or subtropical climates (Figure 1).

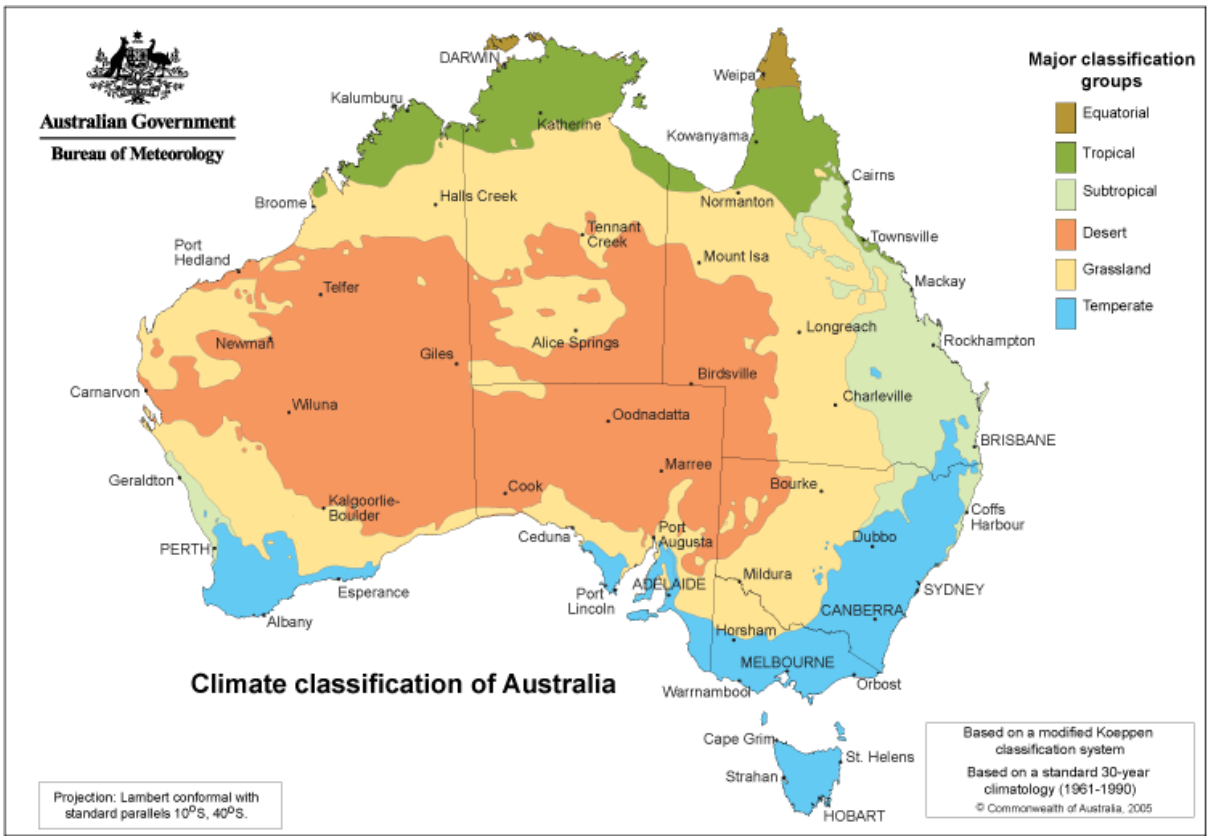


Figure 1: Map showing seasonal zones of Australia, based on major Köppen seasonal climate zones.
Source: Australian Bureau of Meteorology (http://www.bom.gov.au) product code: IDCJCM0000. Lines on map delineate Australian states and territories. The Northern Territory is in the upper middle with Darwin as its capital.

Around one-quarter of the NT population is Indigenous, the vast majority of whom identify as Aboriginal. There are significant differences between the Aboriginal and non-Aboriginal populations in the NT, in terms of geography, exposure to circulating disease and economic resources. 80% of Aboriginal residents in the NT live outside of Darwin, many in remote Aboriginal communities which experience more extreme weather conditions. In addition, just under half of Aboriginal people in the NT live in housing classified as 'overcrowded' (Australian Bureau of Statistics 2022a), which is associated with greater exposure to infectious disease (Memmott et al. 2022). Median income for Aboriginal people is one-quarter of the median for non-Aboriginal NT residents, and around 45% of Aboriginal households fall below the Australian poverty line, compared with 10% of non-Aboriginal NT households (Altman 2017). This means that Aboriginal people both have fewer financial resources that might allow them to adapt and avoid

seasonal risks (such as purchasing air conditioning), and that fluctuations in food prices and employment conditions are more likely to impact their consumption contemporaneously.

2.2.2 Weather

Exposure to hot weather in the late stages of pregnancy can bring forward labour. For example, in the USA, Barreca and Schaller (2019) find temperatures above 32 degrees Celsius can reduce gestation by up to two weeks. They suggest this occurs because heat increases pregnant women's oxytocin levels and cardiovascular stress, both of which can induce labour.

There is also evidence that heat exposure earlier in pregnancy can worsen newborns' health, though the mechanisms are less well understood. For example, Andalón et al. (2016) find that exposure to high heat any time during pregnancy lowers Apgar scores and gestational age in Colombia; for the USA, Sun et al. (2019) find that higher average temperatures throughout pregnancy increase the risk of babies being small for gestational age (SGA). Grace et al. (2015) find a similar pattern in 19 African countries, with exposure to additional days during pregnancy above 38 degrees Celsius corresponding to lower birthweight, particularly if that exposure was during the second trimester.

Some studies also consider the role of cold weather. They paint a mixed picture. Mathew et al. (2017) find that exposure to temperatures below freezing is associated with higher risk of preterm birth in central Australia. However, in Colombia, Andalón et al. (2016) find a coldwave during pregnancy has a small positive effect on Apgar scores.

Rainfall is often associated with healthier birth outcomes. This relationship is generally found in locations where a large share of the population is reliant on agriculture, since rainfall can contribute to stronger crop growth. For instance, higher rainfall during pregnancy is associated with higher birthweight in Mali and Kenya (Bakhtsiyarava, Grace, and Nawrotzki 2018), and with higher birthweight and longer gestation in Brazil (Rocha and Soares 2015).

However, this impact is likely context specific. In the NT, agriculture makes up a very small share of the economy (2.4% of employment in 2020 (ABS 2020)), meaning that most households are unlikely to be directly affected by agricultural cycles.

Besides contributing to agricultural cycles, rainfall can cause flooding. In the tropical north of the NT, flooding cuts off road access to some communities, limiting access to healthcare and fresh food during the wet season, sometimes for months at a time.[5] There is no analysis of the effect of flooding and seasonal migration on health outcomes in the NT. A small literature, summarised by Mallett and Etzel (2018), suggests flooding worsens birth outcomes and points to maternal stress to explain this finding. Chang et al. (2020) find long-term effects from such exposure in Southern India, with children exposed to heavy rainfall shocks in utero having poorer cognitive and non-cognitive skills at ages 5 and 15.

In this study, I include humidity as an additional seasonal risk. In most cases, humidity may not be expected to have any impact independent from rainfall. But in the tropical part of the NT, humidity is not directly linked to rainfall. It increases in the 'build up' to the wet season (September-December), which is a prolonged period with high humidity and temperatures but little or no rainfall. There could be two reasons for humidity to impact outcomes independently of heat and rainfall. First, the body's usual thermoregulation response is less effective in humid conditions (Oppermann et al. 2017), meaning that humidity may accentuate the physiological impacts of heat. Second, the humid 'build up' period is often anecdotally linked with domestic violence (e.g. see Australian Attorney General's Department (2004)). Though there is no quantitative evidence on this relationship, it is consistent, for instance, with Brunson et al.'s (2009) finding that in the UK, higher humidity is associated with a higher number of police calls for antisocial behaviour. If humidity does induce higher rates of domestic violence, this is likely to worsen newborn health (e.g. as found by Aizer (2011) in the USA).

[5] E.g. see https://securent.nt.gov.au/prepare-for-an-emergency/flooding. Anecdotally, residents avoid exposure to these risks by moving to larger population centres during the wet season. However, I cannot measure this response with my data.

2.2.3 Disease prevalence

Many infectious diseases circulate seasonally (Grassly and Fraser 2006), and pregnant women and their babies are at higher risk of complications if they become infected (Kourtis, Read, and Jamieson 2014). It is therefore likely that seasonality in disease transmission contributes to seasonality in newborns' health.

Influenza has received more attention than other diseases in explaining seasonality in birth outcomes. For instance, Dorélien (2019) and Schwandt (2019) find that maternal influenza exposure during pregnancy worsens birth outcomes in the USA and Denmark, respectively. Nunes et al.'s (2016) systematic review concludes that providing influenza vaccines to pregnant women lowers rates of preterm birth and low birthweight.

However, influenza is not the only seasonal disease. Using reportable disease data (described below), I identify three additional groups of diseases that are transmitted seasonally in the NT (see Appendix B). They are diarrheal and gastro-intestinal illnesses, mosquito-borne diseases, and sexually transmitted infections (STIs). I group together diseases transmitted in the same way, as they have similar seasonal patterns (i.e. mosquito-borne diseases spreading during the wet season).

These diseases have been associated with worse newborn health. Newman et al. (2019) find that diarrheal illness during pregnancy is associated with higher rates of SGA in Nepal. Mullick et al. (2005) summarise evidence from developing countries on the negative effects of STIs on birth outcomes. Moore et al. (2017) find that mosquito-borne diseases (notably malaria) reduce gestational age and intrauterine growth. Seasonal transmission of these diseases may therefore be expected to contribute to seasonality in birth outcomes.

2.2.4 Economic conditions

Parents' economic circumstances have important effects on birth outcomes. There is strong evidence that providing parents with additional economic resources during pregnancy can improve newborns' health (e.g. Chorniy et al. (2020) and Amarante et al. (2016)), likely because they are able to afford more or more nutritious food, and healthcare (in cases where care is not publicly provided). Given these findings, it is

possible that seasonal patterns in parents' incomes and purchasing power may also affect birth outcomes.

In Australia, average food prices fluctuate by between 2-3% with the seasons, and are highest in the June quarter. This means that the purchasing power of a constant income is lowest at that time of year.[6] For some households, this could affect the affordability of a sufficient caloric intake. For others, the effects may come through substitution between different foods. For instance Prasad et al. (2010) find that in Finland, pregnant women's vegetable consumption is highest in summer, and lowest in winter. They attribute these patterns to seasonality in prices and availability of fresh food. Watson and McDonald (2007) find similar patterns for pregnant women in New Zealand.

Seasonality in incomes may also affect affordability of nutritious food and healthcare. While there is no research on the effects of seasonal variation in household income on birth outcomes, Kyriopoulos et al. (2019) find evidence of a macroeconomic business cycle effect. They show that in Greece, business cycle fluctuations affect birthweight and gestational age, with effects largest for parents from low socio-economic backgrounds. In the NT, employment levels are highest in the final months of the year (September – December) and lowest in January and February.[7] The magnitude of seasonal fluctuations is modest, but large in comparison with the rest of Australia.

2.2.5 Fertility and parental characteristics

Parental characteristics are strongly related to birth outcomes, and in some contexts, the socioeconomic characteristics of new parents differ throughout the year. That is, even without seasonal variation in individual parents' characteristics and economic circumstances, we may observe a seasonal pattern in aggregate because of differential fertility patterns between socioeconomic groups. The nature and magnitude of these

[6] Over the 40 year period from 1980 to 2020, the average level of the Consumer Price Index in Darwin (the capital city of the NT) was 1.1% higher in the June quarter than the year-average, roughly in line with patterns in other capital cities in Australia (Sydney: 1.3%, Melbourne, 1.3%, Perth: 1.0%).

[7] Over the 20-year period from 2000 to 2020, the average employment-to-population ratio in the NT was 70.4 from September to December, and 68.5 from January to February. On average, there are 3,000 fewer jobs in the February than in November, out of an average workforce of 117,000. Total hours worked increase by around 5% within a year from trough to peak in the NT, while the equivalent number for Australia overall is 1%.

seasonal patterns varies across countries (Dorélien 2016), and even within countries such as Australia (Wilson, McDonald, and Temple 2020).

Seasonal patterns in parents' characteristics may result from preferences to give birth at a particular time of year, and/or from some socioeconomic groups being better able to realise such preferences. For example, Clarke et al. (2019) find that women in the USA have a preference for giving birth in spring. These preferences are consistent with Currie and Schwandt's (2013) findings of a seasonal pattern in conceptions in the USA. Furthermore, Currie and Schwandt find that white, college-educated and married women – whose babies are generally less likely to be preterm or low birthweight – are more likely to give birth in spring. Their evidence suggests that a large part of the seasonality in birth outcomes in the USA is because of these differences in fertility patterns between socioeconomic groups.

However, this finding is not universal. Torche and Corvalan (2010) find parental characteristics play a more muted role in driving aggregate seasonality in Chile, and Dorélien (2016) finds that while some countries in Sub-Saharan Africa show strong seasonal patterns in conception, others exhibit very limited seasonality. In the NT, there is a seasonal pattern in the characteristics of mothers who conceive at different times of year, but no existing evidence on whether this drives seasonality in birth outcomes; I address this question in the Results section.

2.3 Methods

2.3.1 Data

I use de-identified administrative birth records for all children born in the NT from 2005-2014, which are available through the Child and Youth Development Research Partnership (CYDRP). For this analysis, I link in:

- Observations of rainfall, humidity, maximum temperature and minimum temperature, from NASA's Power Data Access Viewer. I use daily observations

measured at intervals of ½ x ½ degrees of latitude and longitude (roughly 50x55km), giving a total of just over 500 locations throughout the NT.

- Records of the date and location of flooding-related road closures or roads deemed impassable due to flooding.[8] These data were provided by the Northern Territory Department of Infrastructure, Planning and Logistics and are available from 2005 onwards. I construct a variable with the number of road closures in each month, by region.[9]
- Region-specific disease prevalence, extracted from the Northern Territory Government's quarterly *Disease Control Bulletins*. These Bulletins report the number of confirmed cases of a range of reportable diseases, including influenza, mosquito-borne diseases, gastrointestinal-related diseases, and STIs. I converted case numbers to rates per 1,000 residents, using population estimates from the Australian Bureau of Statistics.
- Quarterly consumer price indices for Darwin (the capital city of the NT), and monthly employment-to-population ratio data for the NT, both from the Australian Bureau of Statistics.

My analysis sample is limited to babies born in the NT from 2005 to 2014, to mothers whose usual place of residence was in the NT, and whose place of residence could be geo-coded.[10] This gives a sample of 35,199 observations, out of a total population of

[8] The NT Department of Infrastructure, Planning and Logistics lists a road as impassable if it frequently floods in the wet season, but staff are unable to reach the location to verify whether it is flooded. I refer to both flooded and impassable roads as 'road closures' in the remainder of this chapter. Roads closed for reasons unrelated to flooding are not included in my measure.

[9] Regions are: Darwin and surrounds, East Arnhem, Barkly, Katherine, and Alice Springs and surrounds. These are the same regions for which the disease prevalence data are available.

[10] There are a small number of births for which the mother's place of residence as entered in the perinatal data could not be found, either using a fuzzy match with the R package 'geonames', or through manual search on Google, the NT Place Names Register (https://www.ntlis.nt.gov.au/placenames/) and BushTel (https://bushtel.nt.gov.au/).

38,866 babies born in the NT over this period. I include stillbirths (making up under 1% of births) and plural births (under 2.5% of births) in the analysis sample.

2.3.2 Measures

Table 1 shows the mean, standard deviation, minimum and maximum values of all variables that I use in analysis.

	Mean	SD	Minimum	Maximum
Outcomes				
Birthweight (g)[a]	3,289	534	2,160	4,180
Preterm birth (probability)	0.10	0.31	0	1
SGA (probability)	0.13	0.34	0	1
Apgar 5 score	8.9	1.2	0	10
Other characteristics of births				
Born in hospital (probability)	0.99	0.09	0	1
Aboriginal or Torres Strait Islander mother (probability)	0.36	0.48	0	1
Number of antenatal visits	9	4	0	38
Urban area (probability)	0.65	0.48	0	1
Arid (vs tropical) climate (probability)	0.20	0.40	0	1
Weather in 39 weeks to birth				
Average maximum temperature (C)	31	2	24	37
Avg days with max>35 C	38	44	0	190
Average minimum temperature (C)	22	4	11	28
Avg days with min<0 C	0	1	0	5
Average daily rainfall (ml)	4	2	0	11
Avg days with>50ml	2	2	0	8
Avg relative humidity (%)	61	14	21	80
Avg days with relative humidity>85%	25	24	0	140
Avg road closures in pregnancy, per month	6.7	3.9	0.0	24.4
Disease lab cases per 1,000 residents				
Influenza	0.6	0.9	0.0	5.0
Sexually transmitted infections (STIs)[b]	7.4	5.2	2.4	19.9
Mosquito-borne diseases[c]	0.5	0.3	0.0	1.7
Gastro-intestinal related diseases[d]	1.4	0.6	0.6	3.2
Economic conditions				
Food price growth in year to birth (%)	2.2	2.2	-2.9	7.1
Average employment-to-population ratio during pregnancy (%)	69.8	2.0	64.3	72.5

Table 1: Summary statistics – births in the NT, 2005-2014

a Variable is top- and bottom-coded at the 2.5th an 97.5th percentile to reduce the influence of extreme outliers.
b Includes: Chlamydia, gonococcal, trichomoniasis and syphilis.
c Includes: Malaria, Dengue, Kunjin virus, Chikungunya, Zika, Murray Valley encephalitis, Barmah Forest virus, Ross River Virus, Chikungunya and arbovirus – not elsewhere specified.
d Includes: Rotavirus, Salmonella, Shigellosis, Campylobacteriosis and Cryptosporidiosis.
Sources: Author's calculations based on Australian Bureau of Statistics; NASA; NT Centre For Disease Control Bulletins; NT Department of Infrastructure, Planning and Logistics; CYDRP.

The advantage of these outcomes is that they are routinely collected and are therefore widely available for research. However, recent research has criticised the focus on birthweight, gestational age, and derived variables as primary measures of health at birth. Conti et al. (2020) show that birthweight provides only a limited picture of infant health. They show that other measures, including Apgar scores, perform better in predicting postnatal outcomes. Recent evidence confirms that seasonal risks can affect fetal development in ways that are not evident from the baby's weight and gestational age. For example, Schwandt (2019) finds that babies exposed to influenza in the second trimester did not have significantly higher rates of preterm birth or low birthweight, but they had significantly lower education attainment and adult wages.

I use 5 minute Apgar scores (Apgar 5) in addition to the measures more commonly used in previous research: birthweight (in grams), preterm birth and SGA.[11] In Appendix C, I show that findings are similar for other commonly-used transformations of these variables (probability of low birthweight, gestational age in weeks, and low Apgar scores).

Apgar 5 is an index with values from 0 to 10, based on the birth attendant's judgement of the newborn's skin colour, heart rate, reflex, muscle tone and respiratory effort, 5 minutes after birth. Only one of the papers cited above (Andalón et al. (2016)) uses Apgar scores as an outcome measure. However, they are commonly used in the medical literature, and lower scores are strongly associated with poor neonatal outcomes (Thavarajah, Flatley, and Kumar 2018), even among full-term pregnancies that are otherwise judged to be low-risk (H.-Y. Chen, Blackwell, and Chauhan 2020).[12]

Seasonal risk factors

I draw on the evidence discussed in Section 2.2 to construct measures of exposure to seasonal risks in utero. They are minimum and maximum temperature, rainfall, humidity, disease prevalence, road closures due to flooding, food price growth and the employment-to-population ratio. I standardise these variables to have a mean of zero and standard

[11] As defined by Dobbins et al. (2012)'s national birthweight percentile tables for Australia.

[12] A common distinction is between Apgar scores of 7-10 which are considered in the 'healthy' range, and scores of 6 or below. But a large population-linkage study from Sweden suggests that even newborns at the lower end of the 'healthy' range face higher health risks (Razaz, Cnattingius, and Joseph 2019). Therefore, I analyse Apgar scores in levels. Appendix M shows estimates for a binary measure of Apgar score.

deviation of one when used in regression analysis, in order to compare the relative strengths of these competing explanations for seasonality in birth outcomes.

I focus primarily on exposure to seasonal risks throughout the full pregnancy. Therefore, I construct measures of average conditions that each mother was exposed to in the 39 weeks to birth (for weather variables), nine months to birth (for road closures and employment variables), or in the three calendar quarters after conception (for regional disease prevalence). Food price growth is measured as year-ended growth in the quarter of birth. Some effects, and in particular temperature exposure, may be non-linear and may depend on the timing of exposure. Therefore, I also construct measures of exposure to very hot and very cold temperatures by trimester, and for road closures and employment rates by trimester. While timing of exposure can be important for disease prevalence as well (Schwandt 2019), analysis is not possible with my data, as disease prevalence is available only at a quarterly frequency.

This chapter seeks to explain seasonality based on the baby's date of birth. Some researchers instead analyse seasonality by date of conception (Currie and Schwandt 2013). I define the start and end date of the sample based on date of conception, to avoid methodological issues with definitions based on date of birth (Strand, Barnett, and Tong 2011a), but in my regression analysis I use month of birth. Before undertaking this analysis, I compared seasonal patterns when measured based on month of conception and month of birth. I found that seasonality is slightly larger when measured based on month of birth (see Appendix D). Therefore, my analysis seeks to explain these patterns, and all descriptive statistics below are shown based on month of birth. As date of birth is directly observed and date of conception is estimated, this approach also reduces the scope for measurement error. In my regression analysis the distinction makes little difference, because I measure total exposure to seasonal risks throughout the pregnancy.

2.3.3 Analytical methods

The goal of this chapter is to analyse the relationships between seasonal risk factors and newborns' health. To do this, I use ordinary least squares to estimate the following model, which is based on the associations I described in Section 2.2, adapted for data availability and specificity.

$$outcome_{it_{yqmrj}} = b_0 + b_1 W_{tj} + b_2 R_{myr} + b_3 D_{qyr} + b_4 P_{qy} + b_5 E_m + \gamma_m + \delta_y + \theta_j + \epsilon_i \quad (1)$$

In this model, I estimate birth outcomes for baby *i*, born on date *t* (year *y*, month *m*, quarter *q*), and in town/city/community *j*, within the broader region *r*. I distinguish between small-area location *j* and region *r* because some explanatory variables are only measured at the region level. Birth outcomes are estimated as a function of weather variables *W*, road closures *R*, disease prevalence *D*, food price growth *P* and the employment-to-population ratio *E*. I calculate confidence intervals using standard errors clustered at the level of the fixed effects (i.e. either location-level, or sibling-level).[13]

The model includes fixed effects by location *j*, and the month *m* and year *y* of birth. Location fixed-effects are based on geographic clusters, which have a maximum distance of 50 miles between any two points. There are 96 clusters in the sample; these are not necessarily nested within the larger regions at which the disease prevalence and road closure measures vary. I derive clusters using the mother's community, town or suburb of residence at the time of birth. In the NT, there are many communities with multiple names or 'aliases', and larger towns surrounded by suburbs and town camps. I first geocode each location to find the latitude and longitude coordinates for the centre of the community, town or suburb. I then use cluster analysis to convert these coordinates into geographic groupings based on distance between points, so that I group together all suburbs within the same city or town, and all 'aliases' for the same community.[14] Given the dispersed population of the NT, these clusters are small, with on average 360 births per cluster over the 10-year period.

[13] The reason for this is that location is not constant within all sibling groups (a small number of mothers move between cities/towns/communities), and therefore I cannot cluster standard errors in the sibling fixed effects model by location.

[14] The data I have on location is a text string with the suburb or community name, provide by the mother to the hospital at the time of birth. This may be the name of the community, town or city (e.g. Darwin) or the name of a suburb (e.g. within greater Darwin). By using geographic clusters, I group together points that are geographically close, to define a single fixed effect for each city, town or community. I chose a 50 mile distance as this would allow one cluster to cover all of greater Darwin (the capital city). However, my estimates are little changed if I instead define fixed effects using shorter distances such as 10 miles or 25 miles – see Appendix E.

Time fixed effects control for month and year of birth. I choose to include month and year separately in my main model estimates because my measures of economic conditions do not vary within a given month-year pair, and therefore could not be estimated using month-year fixed effects. However, my findings for the remaining seasonal risks are robust to inclusion of month-year fixed effects (see Appendix F).

The location and time fixed effects reflect my conceptualisation of seasonal risks relative to the average values across individuals within a specific location, year and month. This allows for the fact that residents adapt to typical conditions in their location and anticipate and adapt to regular seasonal variation. Importantly, this means that the seasonal effects in my analysis are identified by idiosyncratic variation, not from typical seasonal patterns. That is, for example, that my estimated impact of higher temperatures is identified from cases where temperatures were warmer than usual, given the location and time of year. This is the most common approach in the literature, with many studies using location and time fixed effects to identify the causal impact of in utero exposure to weather conditions or disease prevalence (Andalón et al. 2016; Bakhtsiyarava, Grace, and Nawrotzki 2018; Barreca and Schaller 2019; X. Chen et al. 2020; Dorélien 2019; Grace et al. 2015; Hu and Li 2019; J. Kim, Lee, and Rossin-Slater 2021; Li et al. 2018; Maccini and Yang 2009; Mrejen, Perelman, and Machado 2020; Rocha and Soares 2015; Schifano et al. 2016; Wilde, Apouey, and Jung 2017). In Section 6, I discuss the implications of this modelling approach, and the assumptions required to interpret my estimates as evidence of the impacts of seasonal risks.

The model allows me to estimate the effects of a broad range of environmental factors on birth outcomes, holding other factors constant, using data on all babies born in the NT from 2005 to 2014. However, as described above, parental characteristics may contribute both to the timing of conception (and therefore seasonal risk factors that the fetus is exposed to) and birth outcomes. I do not observe parental characteristics in the data, but I can link siblings together. This allows me to hold constant any time-invariant shared family background, so that I can analyse how important these unobserved characteristics are in explaining seasonality. Analysis of siblings is based on smaller sample size, with approximately 50% of the children in the sample having a sibling born during the 2005-2014 period. Therefore I present results with two versions of the model: one using the

full sample without sibling fixed effects, and one with the addition of sibling fixed effects for this subsample.

2.4 Descriptive statistics

In the NT, babies born during the winter/dry season are healthier on average. They have higher average birthweight, higher Apgar 5 scores, and are less likely to be preterm (Figure 2). In contrast, babies born in the wet season/summer – particularly in the middle of the wet season (January and February) – have lower average birthweight and Apgar 5 scores, and are more likely to be SGA or preterm. While birthweight and Apgar 5 are shown here as averages, their seasonal patterns are accentuated at the lower end of the distribution (see Appendix G).

We can compare the magnitudes with other studies based on birthweight, as this is the most commonly used outcome. Compared with other contexts, seasonality in the NT is large. In the NT, there is a standard deviation of 14 grams in birthweight across the month averages. This is large relative to other influences on birthweight: for instance, the variation in birthweight based on month of birth in the NT is slightly larger than the impact that Reader (2023) finds of a modest cash transfer during pregnancy in the UK (of 8-12 grams). The difference between birthweight in its lowest average month (January) and its highest average month (October) is around 55 grams. This is twice as large as the difference that Currie and Schwandt (2013) find in the north-eastern USA (25 grams), and more than 5 times as large as what Torche and Corvalan (2010) find in Chile (9 grams). These seasonal patterns are also much larger than those in other parts of Australia. McGrath, Barnett and Eyles (2005) analyse seasonality in birthweight in the eastern states of Australia, finding small seasonal differences, ranging from 1.4 grams (in Tasmania) and 7.7 grams (in Queensland).[15]

Interestingly, seasonal patterns vary across the four birth outcomes. Seasonal variation is larger in magnitude for birthweight and preterm birth than for the other outcomes. The

[15] Some of these studies consider only term births, with gestational age of 37 weeks and above. In the NT data, the 50 gram difference from highest average to lowest-average month remains, even if we also limit analysis to babies born from the 37th week onwards (see Appendix G).

timing also differs across outcomes. Birthweight and preterm birth show more favourable outcomes first in April and again in October. In contrast, April is the month with the least favourable outcomes for SGA, while Apgar scores are highest in June and November. The different timing of these patterns suggests that not all birth outcomes are influenced by the same seasonal risks.

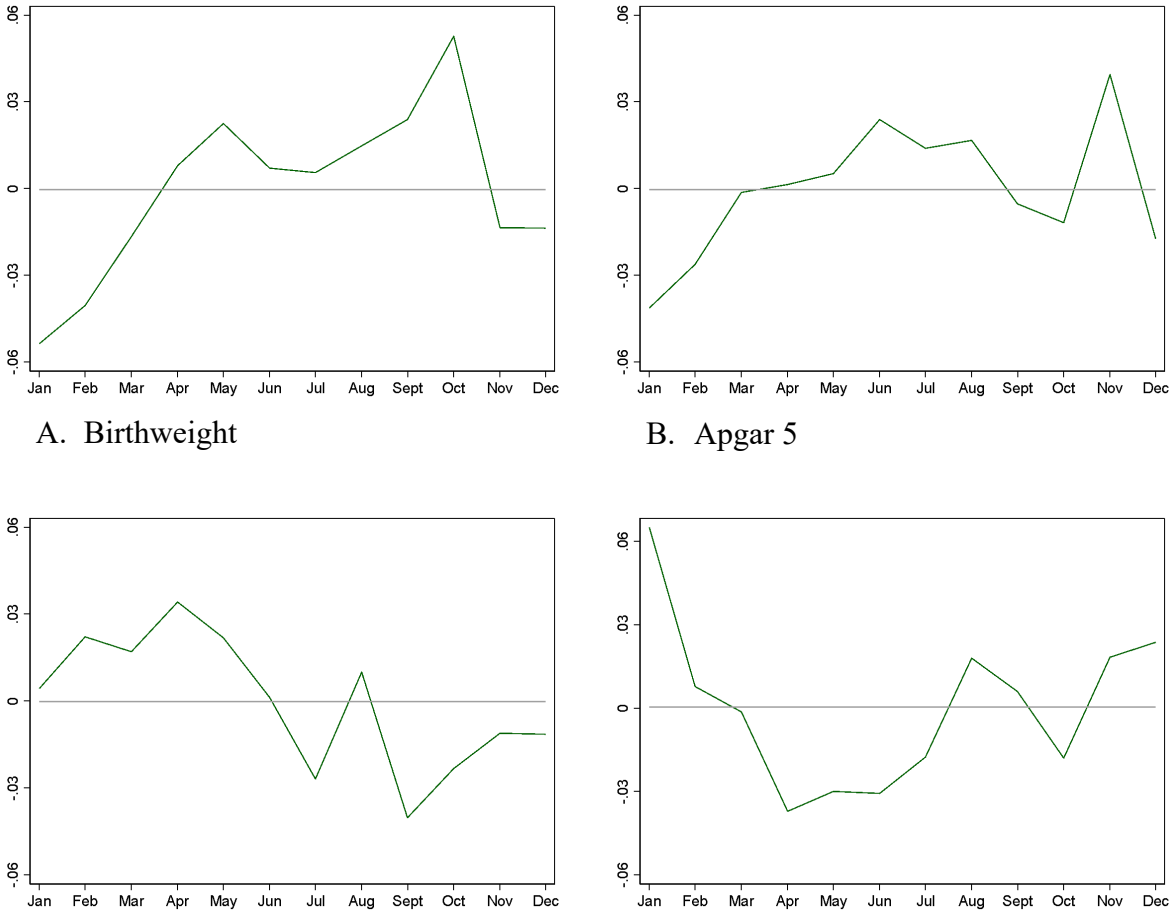


Figure 2: seasonal variations in birth outcomes, by month of birth, 2005-2014
Horizontal line represents the year-average for each outcome (standardised to zero). Vertical axis shows the number of standard deviations from the variable's average, to allow comparison in magnitudes across the four outcomes. Appendix D shows these same graphs in levels of the variable, and Appendix G shows these patterns with 95% confidence bands.

The pattern across the NT also obscures variation across climate zones and between urban and rural locations. There is more seasonal variation in arid ('desert') regions in the central and southern part of the NT, and in rural areas, than in tropical and urban areas. The timing is also different, with gestational age and birthweight peaking in May in

tropical areas, and in October in arid regions, generating the two in peaks seen in Figures 2A and 2D (see Appendix H). I discuss these differences further in Section 2.5.3.

2.5 Results

2.5.1 Which seasonal risks are associated with birth outcomes?

Almost all seasonal risks are associated with birth outcomes to a meaningful degree before adjusting for time and location fixed effects (Appendix I). It is unlikely that these associations represent causal effects. Instead, they may reflect variation between locations in exposure to seasonal risks, and the fact that many seasonal risks are correlated with each other in time – highlighting the importance of an analytical approach that disentangles these effects. The remainder of this section does so by jointly analysing all seasonal risk factors and controlling for time and location-specific effects.

2.5.2 Regression analysis

Figure 3 summarises the primary estimates with 95% confidence intervals (see Appendix J for regression tables). The higher points indicate the coefficients on each seasonal risk, in a regression that includes fixed effects but does not control for other seasonal risk factors. The middle points show the coefficient when all seasonal risk factors are included in a single regression model (as specified in equation (1)). The lower points show the sibling fixed effects model. With some exceptions (discussed below), the estimates are similar in all three models, though confidence intervals are wider in the sibling fixed effects model, possibly reflecting the smaller sample size. Given the large number of effects estimated, in Appendix K I show a table of the benchmark model estimates, adjusted for multiple hypothesis testing using Anderson's (2008) Sharpened Q values. This adjustment does not affect my conclusions, but I note below where estimated effects are no longer statistically significant after this adjustment.

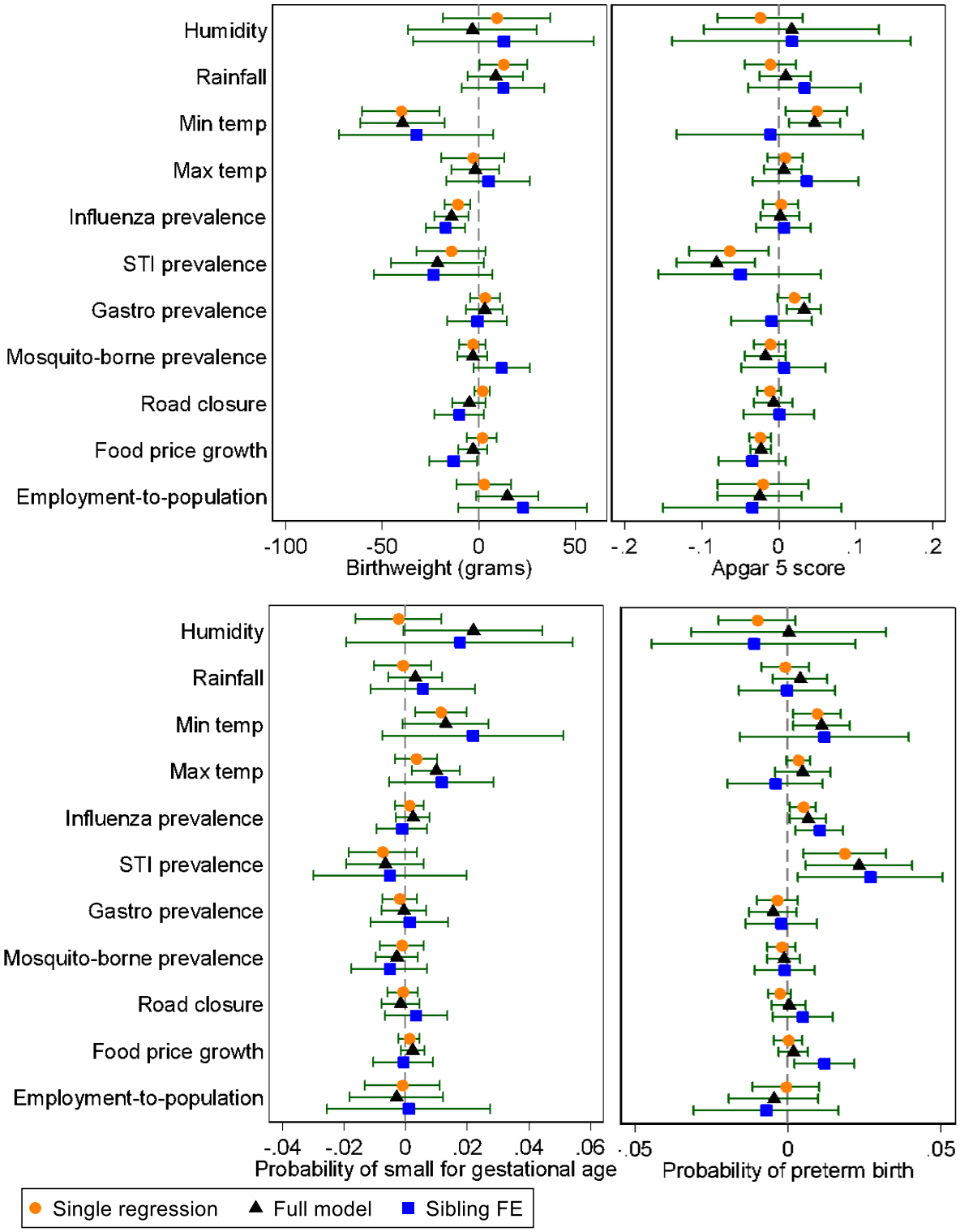


Figure 3: Main regression coefficients with 95% confidence intervals.
All explanatory variables are standardised to have a mean of zero and standard deviation of 1. The higher points (circles) represent estimates from a regression with fixed effects but without controlling for other seasonal factors (so that each point is an estimate from a different regression). The middle points (triangles) represent the coefficients from the full model, controlling for all other seasonal factors. The lower (squares) points represent coefficients from the full model with sibling fixed effects. Sample size 35,199 for full model, and 18,359 for sibling fixed effects model.

Relative humidity

We see no clear relationship between average humidity and birth outcomes in Figure 3. One channel through which I proposed that humidity would affect birth outcomes in the NT was domestic violence – with the potential for high humidity to increase interpersonal violence. To analyse this mechanism, I examined whether there is a seasonal pattern in domestic violence using child protection notifications.[16] However, I do not find evidence of seasonality in the number of incidences (see Appendix L).

Rainfall and road closure

There is no apparent relationship between average rainfall and birth outcomes. All coefficients are close to zero and not statistically significant. This is unsurprising given that the studies that have found a relationship between rainfall and birth outcomes have been from locations where a large share of households rely on agriculture, and this is not the case in the NT.

There is, however, some evidence that extreme values of rainfall – and resultant road flooding – can worsen birth outcomes. For the estimates shown in Figure 3, the coefficients on road closures are small and not statistically significant. However, my measure of road closures is a simple count of the number of road closures within the region. It does not identify individuals who were affected by each closure, the length of the closure, or the importance of the road. When I restrict analysis to Aboriginal mothers, who are more likely to live in the remote communities that are most inconvenienced by road closures, the coefficient on birthweight is negative and statistically significant at the 5% level (see Section 2.5.3). In addition, in the benchmark model with the full sample, when I run the model splitting exposure into the first two trimesters and the third trimester, more road closures in the first two trimesters are associated with statistically significantly lower birthweight and lower Apgar scores (see Appendix M).[17] Therefore, I interpret the small negative coefficients as suggestive evidence of an effect on the subset of pregnancies for which road closures affect access to healthcare and fresh food.

[16] In the NT, it is mandatory for certain professionals to report suspected child abuse and neglect, and this includes physical violence and exposure to family violence.

[17] It may be that women expecting to give birth in a community prone to flooding during the wet season leave the community before floods begin, leading to little effect in the third trimester. In addition, almost all pregnant women in remote communities are transferred to their closest hospital at 37 weeks.

However, my measure of closures is at the region level. In a month with a high number of closures, some communities will be severely affected, while most mothers within the region will be unaffected. Further research using a community-level measure of road closures would be needed both to verify this finding, and if it is verified, to quantify its magnitude more precisely.

Temperature

There are different effects of temperature across the four birth outcomes. Higher average minimum temperatures are associated with higher Apgar scores, but also with higher risk of SGA, preterm birth, and lower birthweight (Figure 3). Higher average maximum temperatures are also associated with higher risk of SGA but have no statistically significant effect on the other outcomes. These effects remain statistically significant at the 5% or 10% level, after adjusting for multiple hypothesis testing (see Appendix K).

These somewhat conflicting findings may reflect the shortcomings of a linear analysis of temperatures based on averages over a 39-week period. Higher minimum temperatures in the dry season/winter means absence of very cold weather in some parts of the NT, while high minimum temperatures in the wet season/summer can reflect a heatwave. We may therefore expect non-linear effects. In addition, it is possible that exposure at different times during pregnancy, or exposure to extreme cold or heat may be more important than average exposure throughout the whole pregnancy.

I therefore construct measures of exposure to very high and very low temperatures. These are indicators of the number of days during pregnancy with maximum temperatures above 35 degrees Celsius, and with minimum temperatures below freezing (see Table 1 for summary statistics of these measures). I disaggregate these by trimester.

Figure 4 reports estimates from regressions including those measures. It shows that additional days above 35 degrees during pregnancy – in any trimester – reduce birthweight and Apgar scores and increases risk of preterm birth. Exposure to additional days below freezing, in contrast, has relatively small effects, but helps to clarify the independent effect of average minimum temperature. After controlling for exposure to very warm and very cold days, higher average minimum temperature is associated with poorer birth outcomes, consistent with this variable indicating heatwaves. In addition,

higher average maximum temperature is more consistently associated with healthier birth outcomes. This suggests that in the absence of heatwaves, warmer weather may improve birth outcomes.

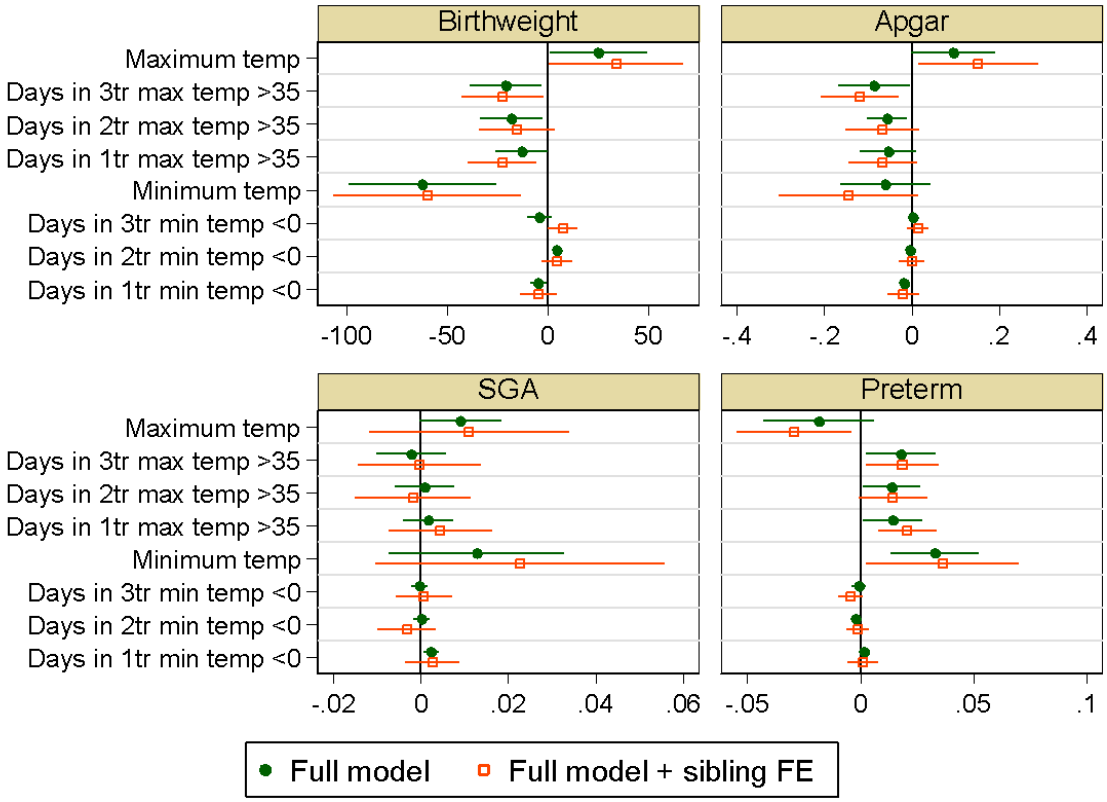


Figure 4: Regression estimates controlling for extreme weather conditions, by trimester.

These reflect coefficients from a single regression using the full sample, with all variables included – including other variables from the full model specified in equation (1) (circles), and from the sibling fixed effects model, using the smaller sample (squares). Sample size 35,199 for full model, and 18,359 for sibling fixed effects model.

Disease prevalence

Figure 3 reveals a large and statistically significant relationship between prevalence of STIs during pregnancy and birth outcomes. One standard deviation increase in STI prevalence is associated with a 0.08 point decrease in Apgar 5 scores and a 2.4 percentage point increase in preterm births, statistically significant 10% and 5% levels respectively, after adjusting for multiple hypotheses. The confidence intervals around these estimates are wider in the sibling fixed effects model and are therefore not statistically significant at the 5% level, but the estimates themselves are similar for birthweight and preterm birth (though smaller for Apgar scores).

To analyse this relationship in more detail, I rerun the model separately for mothers who did not regularly attend antenatal care visits during pregnancy, because antenatal care appointments provide opportunities for testing and treatment of STIs.[18] I find a substantially larger association in this sub-sample (Appendix N). This relationship is suggestive of the role of antenatal testing and treatment of STIs in attenuating seasonality. However, it could also be that fewer antenatal visits reflects access to care more generally or the mother's socio-economic status (Australian Institute of Health and Welfare 2020), both of which could be associated with greater exposure to STIs relative to the region-level case rate that is captured by the prevalence measure I use.

Influenza prevalence significantly increases the risk of preterm birth and lowers birthweight, consistent with findings from other contexts. After adjusting for multiple hypotheses, the relationship with birthweight is statistically significant at the 5% level, though the relationship with preterm birth is not statistically significant.

The coefficients on prevalence of mosquito-borne diseases are small and not statistically significant.

Surprisingly, Figure 3 shows that prevalence of gastrointestinal-related diseases is associated with significantly higher Apgar scores after controlling for other factors. One potential reason for this could be that disease exposure may lead to fetal loss, such that only the fetuses that were already healthier survive (and appear in the birth records). There are no data on miscarriages, but birth records contain information on stillbirths – defined in the NT as fetal losses after 20 weeks of gestation. I find no statistically significant association between prevalence of gastrointestinal disease and live birth status (Appendix O), but it remains possible that fetal loss before 20 weeks may explain the positive relationship between gastrointestinal disease and Apgar scores.

Economic conditions

As may be expected, I find some evidence that higher price inflation is associated with worse birth outcomes and higher employment rates associated with better birth outcomes. In the sibling fixed effects model, the effects are large, with a standard deviation increase

[18] Pregnant women in the NT are recommended to have 10 antenatal visits, and on average have 9 visits (see Table 1). In this analysis, I limit the sample to those who had 5 or fewer visits.

in food price growth associated with a 13 gram decrease in birthweight and a 1 percentage point increase in preterm births. The fact that the effect is larger in this model may indicate that it is certain families (e.g. those with lower incomes) whose birth outcomes are most affected by seasonality in food affordability. For those household, seasonality in food affordability may impact intrauterine growth and general infant health.

However, while the coefficients move in the expected direction, they are small and not statistically significant within most models. Therefore, while other research tells us that income and food affordability are important for infant health (see Section 2.4 above), variation in aggregate economic conditions does not appear to be a major explanation for the seasonality we observe in population-average birth outcomes.

2.5.3 Heterogeneity of effects

By Aboriginal status

During the sample period, just over one-third of births in the NT were to Aboriginal mothers. On average, Aboriginal women face different environments from non-Aboriginal women during pregnancy – 72 percent of Aboriginal mothers lived outside of the two main cities (Darwin and Alice Springs), compared with just 14 percent of non-Aboriginal mothers. Many live in remote Aboriginal communities, where residents are likely to face challenges in accessing healthcare and fresh food, particularly during the wet season, as described above. In addition, Aboriginal people are disproportionately affected by other disadvantages that may influence maternal and infant health, including substantially lower-than-average economic resources, more crowded living conditions, and poorer health (Australian Bureau of Statistics 2017). To test how these differences impact my estimates, I run the model for Aboriginal and non-Aboriginal mothers separately (Figure 5).

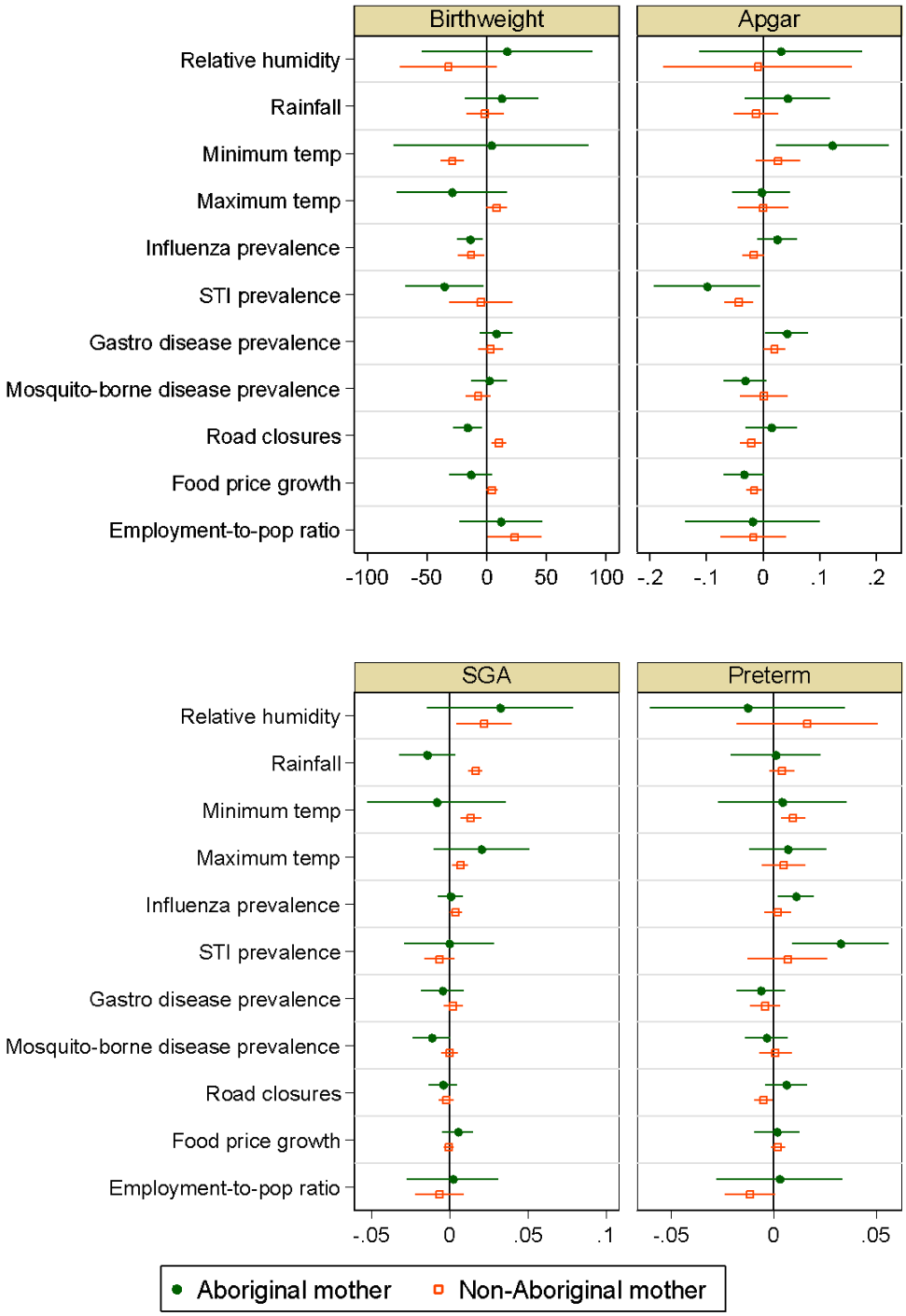


Figure 5: Model estimates by Aboriginal status of mother.
Estimates for mothers identified as Aboriginal or Torres Strait Islander are the higher, circular points (35% of all births) and estimates for all other mothers are the lower, square points (65% of all births). These models do not include sibling fixed effects. All explanatory variables are standardised to have a mean of zero and standard deviation of 1. Note that 99% of people in the NT who identify as Aboriginal or Torres Strait Islander identify as solely Aboriginal or as both Aboriginal and Torres Strait Islander, hence the discussion centres around Aboriginal status. Sample size: 12,523 (Aboriginal) and 22,607 (non-Aboriginal).

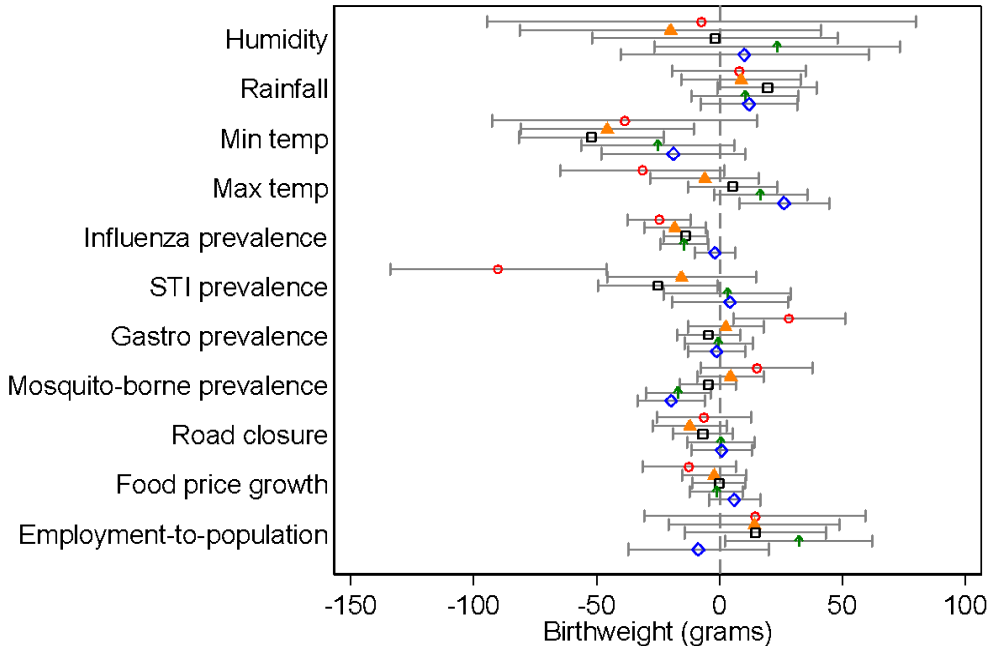
Some findings stand out – for instance, there is a negative effect of road closures on birthweight for Aboriginal mothers, but not for the rest of the sample. This makes sense given that Aboriginal people are more likely to live in remote communities with just one major road in and out of the community, and therefore are more severely affected when roads are inaccessible. The point estimates on STI prevalence are also larger for Aboriginal mothers, potentially reflecting higher STI prevalence in Aboriginal communities (Gooley 2021). However, in most cases, these estimates are consistent with the results from the aggregated sample – that is, showing that exposure to STIs and flu, and less favourable economic conditions worsen birth outcomes for Aboriginal and non-Aboriginal mothers alike.

The confidence intervals around the estimates in the Aboriginal subsample are larger. This is likely to reflect both the smaller sample, and diversity among Aboriginal peoples; Aboriginal communities are located throughout the NT, meaning there is greater diversity among Aboriginal mothers in their exposure to climactic and environmental conditions, and in access to healthcare. In contrast, most non-Aboriginal mothers live in Darwin, the capital city of the Northern Territory.

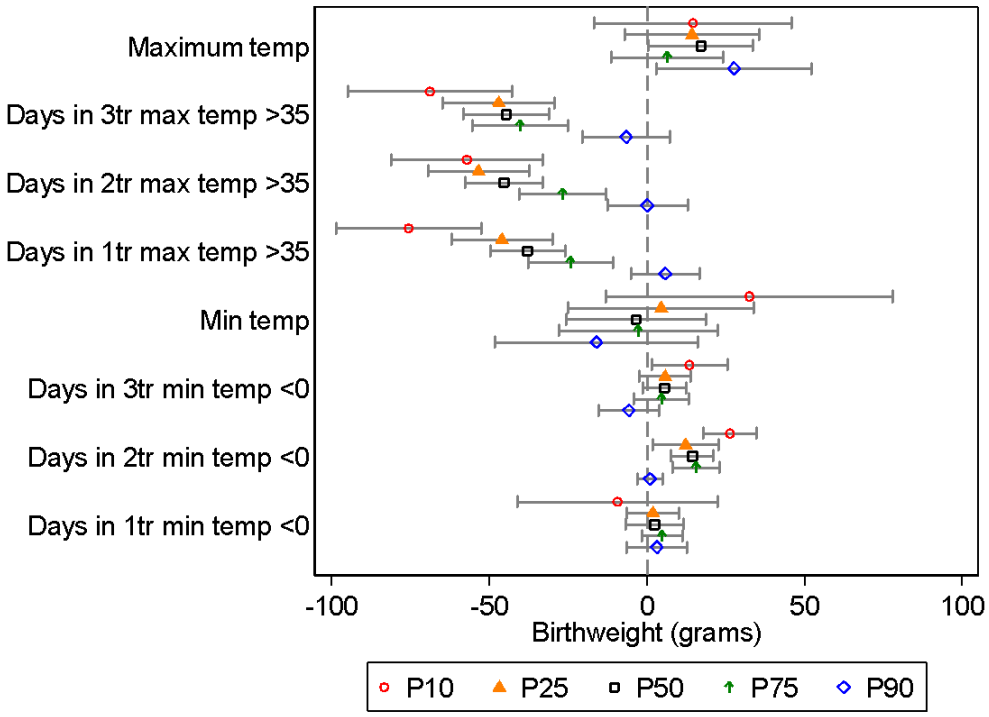
By health status

Figure 6 shows quantile regression estimates for birthweight. They are based on the same specification as equation (1) but instead of estimating the effect on mean birthweight, they are estimated for babies at the 10^{th}, 25^{th}, 50^{th}, 75^{th} and 90^{th} percentiles of the birthweight distribution.

These estimates suggest that seasonal risk factors may have small or no effects on babies that are otherwise healthy, but may interact with other health complications, for babies who may have had low birthweight even in the absence of seasonal risks. The effects of STI prevalence are especially large for babies at the 10^{th} percentile of birthweight and are smaller or not statistically significant for heavier babies. While heat and influenza prevalence have effects throughout most of the distribution, the effects are largest for babies at lower percentiles.



A. Main model



B. Extreme temperature exposure by trimester

Figure 6: Quantile regressions of birthweight.
These figures show the full model as in Figures 3 and 4, with all variables included, but without sibling fixed effects. These are estimated using quantile regressions for the 10th, 25th, 50th, 75th and 90th percentile of the birthweight distribution, with robust standard errors. Sample size 35,199.

By urban/rural location

Birth outcomes are more seasonal in rural areas than urban areas (Appendix H). There could be two potential explanations for this. One is different exposure to seasonal risks – e.g., that it is hotter in rural areas than urban areas. A second explanation is that people living in urban locations are better able to adapt to the same level of risk – for instance, by spending less time outdoors during the hottest part of the day or having more access to air conditioning.

To investigate these explanations, I fit the model separately on the urban and rural subsamples (Figure 7). The estimates suggest that the second explanation – that urban location moderates seasonal effects – better explains the data for most seasonal risk factors. The effects of the same unit of change in average weather conditions are substantially larger – and in some cases, work in different directions – in rural than in urban locations. A higher average maximum temperature, for instance, appears to have a large negative effect on birthweight in rural areas, but no effect in urban areas.

As an additional test, I compare the share of variance explained by the benchmark model with and without including seasonal risks (Appendix P). I find that within the rural subsample, the full set of seasonal risks explains a higher share of variance than in the urban sample for birthweight, Apgar 5 scores and SGA. I interpret this as additional suggestive evidence that the impacts of the same seasonal risks are slightly greater in rural than urban areas. The exception is for preterm birth, for which the explanatory power of the seasonal risks is greater in the urban sample. Based on the estimates shown in Figure 7, this appears to be driven by the large coefficient on STI prevalence, perhaps reflecting the first explanation, that exposure to this particular seasonal risk (which is measured at the region level, and therefore measured prevalence rates are the same for urban and rural locations within a region, while actual prevalence may vary) is greater in urban area, and in particular in Alice Springs – see Section 2.6.

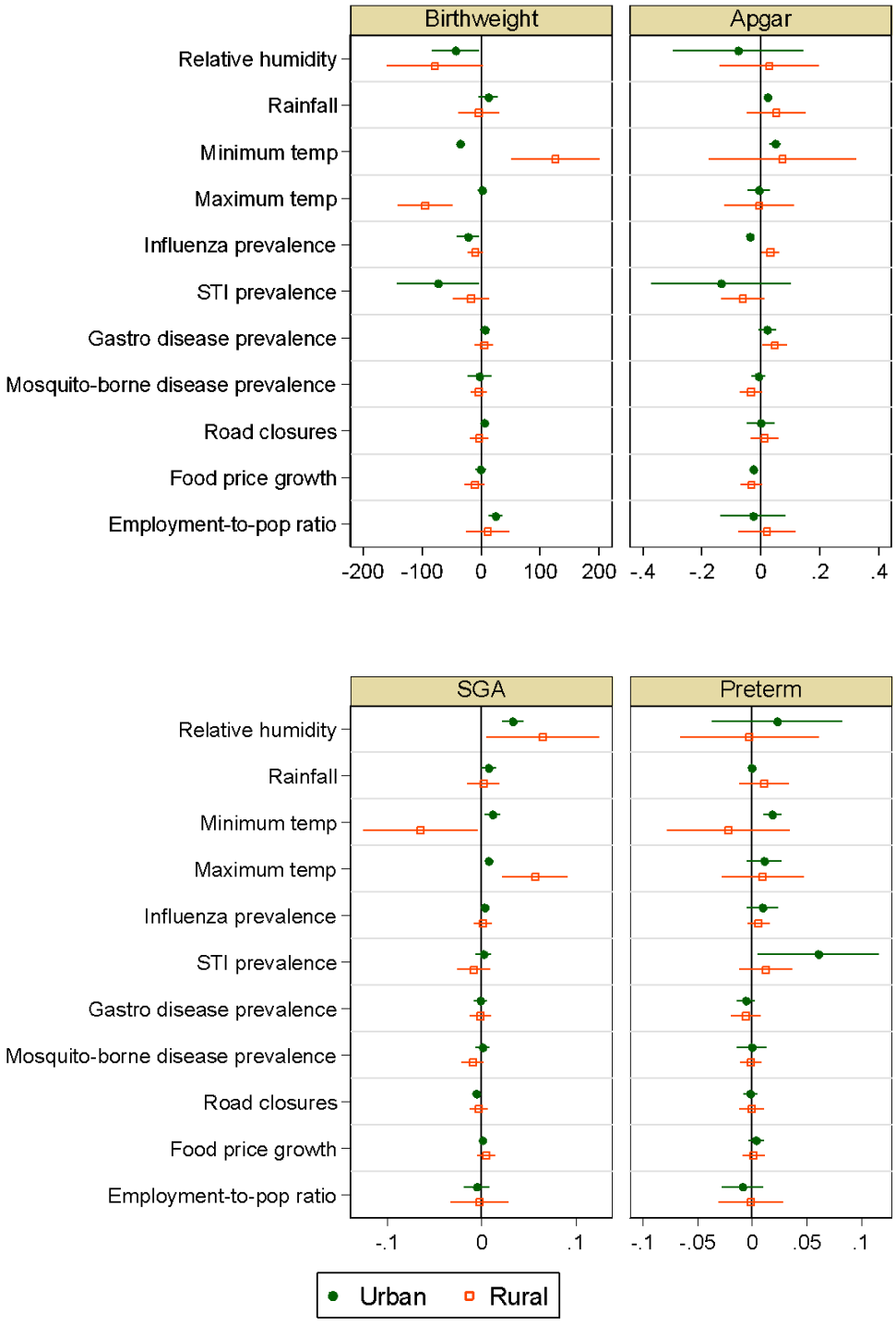


Figure 7: Model estimates by urban/rural location.
All explanatory variables are standardised to have a mean of zero and standard deviation of 1. Sample size: 23,053 (urban) and 12,146 (rural).

2.5.4 Parental characteristics

In the NT, there is some variation in maternal characteristics over the year (see Appendix Q).[19] To analyse whether this variation explains seasonality in birth outcomes, I revisit the sibling fixed-effects model. I focus on the subset of the sample living in rural areas, as the analysis presented above shows there is much less seasonality in the urban sample.

I present models with and without sibling fixed effects, to analyse whether and how the month coefficients (i.e. the unexplained seasonality) change after including sibling fixed effects. If maternal characteristics are driving seasonality, I would expect the magnitude of the month coefficients to move closer to zero after adding sibling fixed effects.

Figure 8 shows the estimates graphically, while Appendix R presents a table with the share of variance explained by the month coefficients. The full model (equation (1), without sibling fixed effects) brings the month coefficients closer to zero. But this is not consistently the case after adding sibling fixed effects. The addition of sibling fixed effects produces larger month coefficients for Apgar scores and SGA, though it reduces the month coefficients slightly for preterm birth. Appendix R shows that the variance explained by the month coefficients increases with the addition of sibling fixed effects for birthweight, Apgar 5 and preterm birth. This is also the case in the full sibling population, and the urban subpopulation.

Based on these estimates, I conclude that it is unlikely that time-invariant parental characteristics drive overall seasonal patterns in birth outcomes in the NT.

[19] Appendix Q shows variation over the year in the share of mothers who smoke or drink at their first antenatal visit, the number of antenatal visits, complications in the mother's medical history, age and Aboriginal status of mother. Some of these factors are time-varying. I do not have data on other maternal characteristics (e.g. socioeconomic status, education, employment or income).

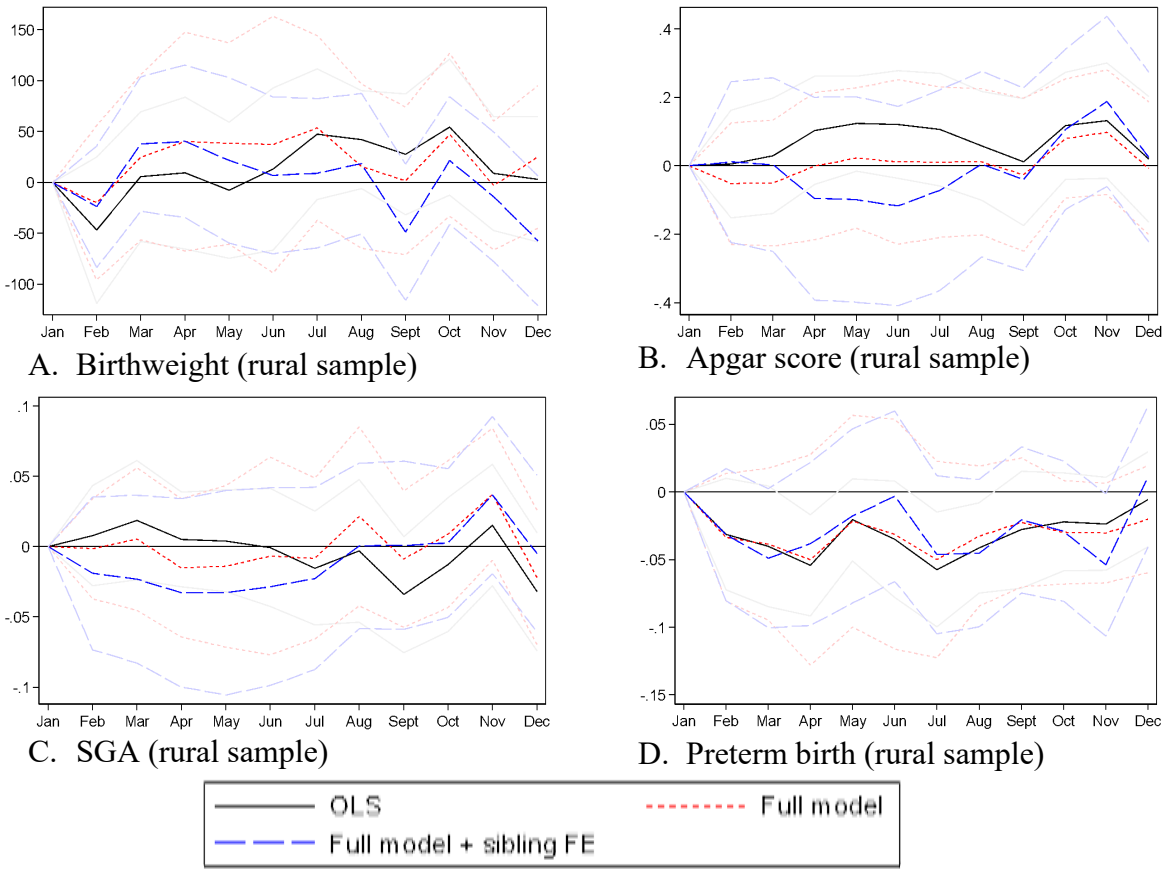


Figure 8: The role of fixed parental characteristics.
Estimated month coefficients and 95% confidence intervals from an OLS regression with only month dummies as covariates (solid line); the full model as in equation (1) (dotted line); the full model with sibling fixed effects (dashed line). Models are estimated on the subset of the sample with siblings, and therefore the coefficients in the OLS model and the full model are different from other estimates, and confidence intervals are wider. Because most mothers do not move between location clusters, the sibling fixed effects model includes controls for region instead of location cluster fixed effects. Sample size: 7,016.

2.6 Implications – what is the effect of typical seasonal risks on newborns' health?

I have estimated the effects on newborns' health of in utero exposure to seasonal variation in weather, disease prevalence, economic conditions, and of seasonal patterns in fertility. I find that heat exposure and disease prevalence are associated with particularly large effects on birth outcomes in the NT (reductions in average birthweight of 23, 21 and 14 grams per standard deviation change in exposure to very hot days in any trimester, STI prevalence, and influenza prevalence, respectively). I also find smaller, but economically and statistically significant effects of food price growth, employment conditions and road

flooding for some outcomes and subpopulations. Seasonal fertility, rainfall and humidity do not appear to explain seasonality in birth outcomes in the NT.

While a thorough analysis of the causal mechanisms behind each of these findings is beyond the scope of this chapter, my results are suggestive of causal effects, because these relationships are found after controlling for small-area location, month and year fixed effects, and after controlling for other seasonal risks. They are also little changed even after controlling for sibling fixed effects. My use of fixed effects allows for the fact that people adapt to usual conditions within their location, and abstracts from any additional unobserved monthly patterns unrelated to those I study here, which may influence month-average birth outcomes.

My use of month fixed effects does, however, mean that while I seek to explain regular seasonal patterns in birth outcomes across months, my coefficients are identified using variation within months.

In this section, I extrapolate from these coefficients, to estimate the effects of the seasonal risks that a baby is typically exposed to in utero, given their month of birth. I focus on the three main seasonal risks I have identified: heat exposure, influenza and STI prevalence. Such extrapolation assumes that within-month variation has the same effect as between-month variation. This means, for instance, that because August is typically slightly warmer than July, my estimates of the effect of heat exposure, identified using unseasonably warm days in July (and other months), are informative about the effects of typical heat exposure for a baby born in August.

Table 2 shows these extrapolated estimates. I show the effects for babies born in the month with peak exposure to each risk, compared with average exposure for babies born in any other month of the year.[20] The magnitude of these estimates depends on two factors: the strength of the association in my regression estimates, and the magnitude of typical seasonal variation of the risk factor.

[20] That is, for example, I show the effect of being born in May, and hence in utero during the hottest months of the year, compared with average exposure for babies born in any other month of the year.

Seasonal risk	Estimated impact on birthweight	Estimated impact on preterm birth rate	Estimated impact on Apgar scores	Estimated impact on SGA rates
Pregnancy spans hottest months of the year (born in Apr)	-8.7 to -11.4 grams	0.008 to 0.009 pr	-0.03 to -0.04 pts	No effect
Pregnancy spanning peak period of Influenza prevalence (born in Dec)	-1.9 to -2.3 grams	0.001 to 0.001 pr	No effect	No effect
Pregnancy spanning peak period of STI prevalence (born in Aug)	-1.1 to -1.2 grams	0.001 to 0.001 pr	0 to -0.004 pts	No effect
Pregnancy spanning peak period of STI prevalence region with highest STI rates (born in Aug)	-4.0 to -4.5 grams	0.004 to 0.005 pr	-0.01 to -0.015 pts	No effect

Table 2: Estimated effects of major seasonal risks on birth outcomes in the NT, derived from regression estimates

Note: estimated impacts are calculated as the estimated impact in the full model with and without sibling FE (giving the range), divided by one standard deviation, multiplied by the peak month minus the year-average excluding that peak. Estimates are shown for the three seasonal influences for which I find statistically significant effects across multiple birth outcomes, but are not shown for the SGA outcome or impact of influenza on Apgar scores, as these estimates were not statistically significant in either model. The estimates for heat exposure represent the sum of the effects of typical exposure in each trimester.

Across all outcomes, the effects are largest for exposure to warm weather. Lending a causal interpretation, this would suggest that a baby for whom the pregnancy does not span the cooler winter/dry season months would be, on average, 8-11 grams lighter, have a 1 percentage point higher probability of preterm birth, and Apgar scores between 0.03 to 0.04 points lower than the year-average *because* of exposure to very warm weather. For a baby in utero during the peak influenza period, the effect is smaller but still substantial: reducing birthweight by 2 grams, and increasing the chance of preterm birth by 0.1 percentage point.

In contrast, while the impact of one standard deviation in STI rates as shown in Figure 2 is large, much of the variance in STI rates is not seasonal. Therefore the impact of typical seasonal exposure is small, reducing birthweight by only 1 gram on average for babies in

utero with the highest seasonal exposure. However, STI rates vary substantially by region. If I instead scale my regression estimates by STI rates in the region with greatest seasonality (Alice Springs) these are much larger. They suggest that a baby in utero in Alice Springs during the June quarter would be, on average, 4 grams smaller because of the higher risk of exposure to STIs, have a 0.5 ppt higher chance of preterm birth, and have Apgar scores 0.01 points lower.

This chapter seeks to explain the reasons behind worse birth outcomes for babies born in the summer/wet season, compared with those born in the winter/dry season. A baby born in the summer/wet season is exposed to all three of the risks set out in Table 2 – they are in utero during the hottest months of the year, and during peak periods of flu and STI prevalence. While the seasonal pattern implied by these three risk factors do not perfectly match up with the seasonal patterns we observe, they come close, and move in the expected direction.

The magnitudes of these estimates are large, within the context of other influences on birth outcomes. For instance, they are comparable in magnitude to the impact that Almond, Hoynes and Schanzenbach (2011) find from the introduction of the US food stamps program during pregnancy, which increased average birthweight by 1-8 grams, depending on the subpopulation studied. Reader (2023) finds that the UK's Health in Pregnancy grant increased birthweight by 8-12 grams. I find effects from annual seasonal variation in heat and disease prevalence that are about the same as the impact from these substantial increases in income, but moving in the opposite direction. In contrast, my estimates are smaller than those of Bakhtsiyarava et al. (2018), who find that an additional month of temperatures above 35 degrees in food cropping communities in Kenya and Mali reduces average birthweight by 71 grams (more than double the 33 grams reduction from an additional 44 days (1 SD) of hot weather in the NT). This may be because relatively high average incomes and access to public healthcare may help to shield newborns in the NT from larger effects.

If readers do not accept the assumptions that produce the estimates in Table 2, my results remain informative. Idiosyncratic, a-seasonal weather events are increasingly common, and disease outbreaks are possible outside of their typical seasonal patterns. I find that these common events experienced in utero can have large impacts on birth outcomes.

2.6.1 Limitations

I identify three limitations to this study which could be addressed with further research.

First, my findings raise the possibility of interaction effects between seasonal risks, parents' characteristics and economic conditions. However, I am unable to explore this relationship because I do not have data on parents' socio-economic status. This may be an avenue for future research.

Second, while I analyse a wide range of explanatory variables, many are not measured precisely; exposure to disease, road closures and economic conditions are measured at a broad regional level or higher. The fact that we still observe effects of these imprecisely measured seasonal risks suggests that their effects may be particularly large for some sub-groups. More detailed individual- or community-level data on these factors would help us to better understand these relationships.

Finally, I have estimated the effect of a standard deviation change in each seasonal risk, and my findings suggest that some seasonal risks – in particular, heat exposure and disease prevalence – are more important determinants of infant health than others. However, an equally important consideration is the cost-effectiveness of any intervention to alleviate the effects of these seasonal risks. Future analysis may consider potential policy responses in more detail, alongside the cost and likely impacts of those responses.

2.7 Conclusion and policy applications

I draw three important conclusions from my analysis.

First, analysing multiple risk factors together is important for policy, but not necessarily for research. In this chapter I have demonstrated the importance of considering many pieces of the puzzle together.

On one hand, my analysis indicates that it is possible to obtain reliable estimates of the effects of one seasonal risk without measuring other risks: within my fixed effects models, the estimated effects of each seasonal risk factor generally do not change after including

other seasonal risks in the model. This suggests that any future research seeking to explore a specific seasonal risk in detail can do so without requiring data on all other seasonal risks.

On the other hand, my analysis provides more actionable policy implications than analysis of a single seasonal risk. Without analysing these factors together, we would not know, for instance, that policymakers seeking to reduce inequalities at birth may gain more traction at the population-level by finding ways to shield mothers from heat exposure than from influenza.

Second, we need more analysis on how STIs impact newborn health. Parts of Northern Australia have experienced an epidemic of STIs (Gooley 2021), but such an epidemic is not unique to Australia (L. Newman et al. 2013). The association between STI prevalence and birth outcomes is large, even if the average impact of STI seasonality shown in Table 2 is small. A one standard deviation increase in regional prevalence of STIs reduces birthweight by 21 grams, reduces Apgar scores by 0.08 points, and increases the risk of preterm birth by 2 percentage points. It is not clear whether this relationship is causal.

Linkage of individual-level data on STI diagnosis with birth records would help greatly in disentangling the impacts of STIs themselves from other co-varying risks. Another option for further research in helping to address this challenge is to analyse the reasons for non-attendance at antenatal care. Antenatal care is a key opportunity for testing and treatment of STIs, and I find a stronger link between STI prevalence and poorer birth outcomes for women with low rates of attendance. But we do not know whether these worse outcomes are the result of non-attendance, or the result of other factors correlated with both STI prevalence and non-attendance.

Third, seasonal effects may be greater in poorer parts of the world with tropical and arid climates. It is remarkable that there is such large seasonal variation in birth outcomes in the NT, especially in contrast with the smaller variation that McGrath et al. (2005) find in the sub-tropical and temperate climates of Australia's eastern states.

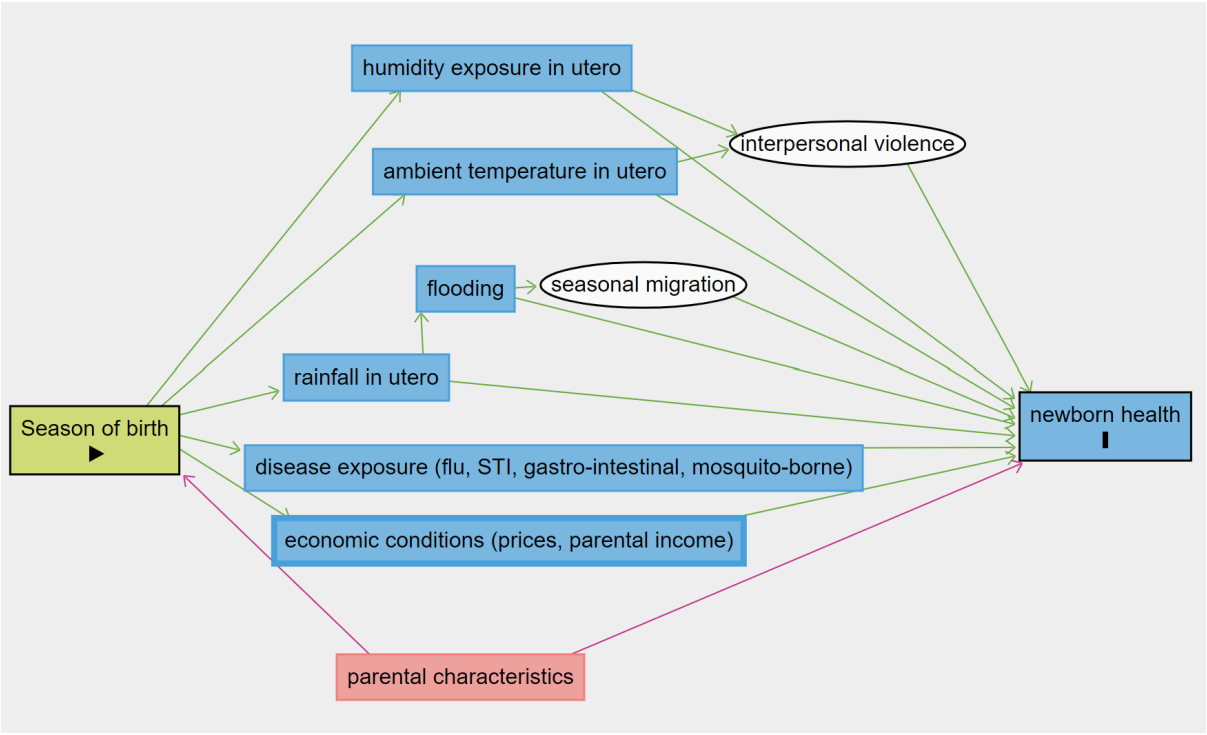
Climate can explain some of these differences. My estimates show large effects of exposure to very hot days and high minimum temperatures, indicative of heatwaves. These high temperatures are more common in tropical and arid climates.

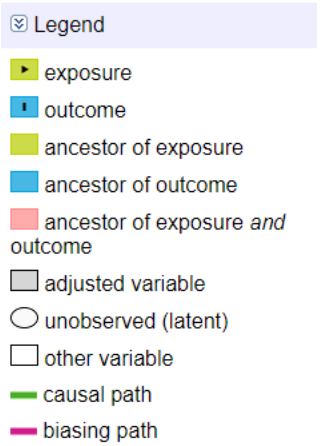
It is likely that these climate effects interact with social and economic factors. While I do not have data on parents' socio-economic status, we know that there are more families with low income in the NT than in the rest of the country,[21] many of whom live in remote Aboriginal communities. Previous research shows that mothers with lower socioeconomic-status tend to have more underlying health conditions, higher exposure to circulating diseases, and less access to nutritious food (Aizer and Currie 2014) – all of which would make them more vulnerable to seasonal risks (Cil and Kim 2022; Nyadanu et al. 2022).

Together, these findings suggest that climate-related seasonal risks may be even more important contributors to poor birth outcomes in lower-income parts of the world. The fact that the effect of heat is smaller in urban areas suggest that adaptations (such as air conditioning, more common in urban areas) may moderate this relationship. In light of the changing climate, improving access to such adaptations will be of increasing importance in reducing inequalities at birth.

[21] Around 9% of the NT population receives an unemployment benefit, compared with 6% of the total Australian population, and a higher share of people in the NT receive a tax benefit targeted at low-income parents. These are based on the Department of Social Services payment demographics data from June 2020, compared against ABS population estimates by state from June 2020.

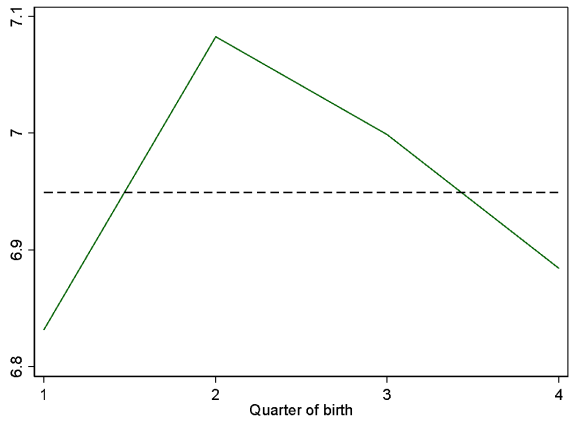
Appendix A – Diagram of conceptual framework



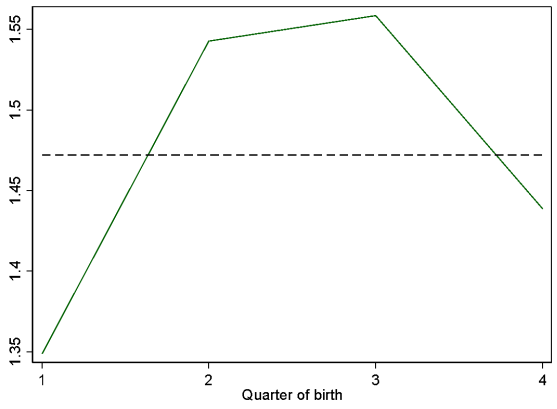


Note: This figure illustrates the links between season of birth and newborn health, as described in the Background and Conceptual Framework section. Diagram generated using daggity.net. Note: Parental characteristics may in theory affect economic conditions and disease exposure at the individual level. In this chapter I analyse circulating diseases and economic conditions at the aggregate level, and I do not expect individual families to influence these aggregate variables. In a more complex model, however, parental characteristics may affect the way that children respond to these aggregate variables (e.g. a social gradient, whereby children in more privileged families may experience a smaller impact).

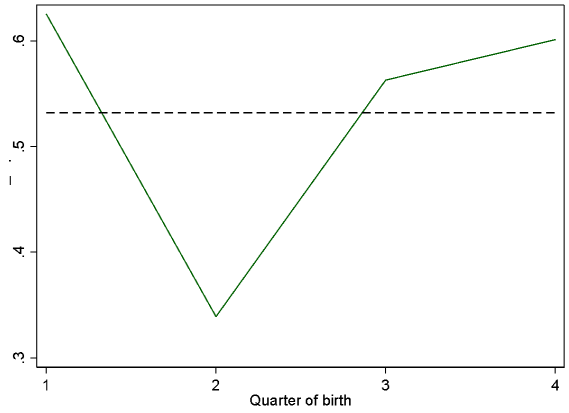
Appendix B – Seasonal pattern in disease prevalence in the NT



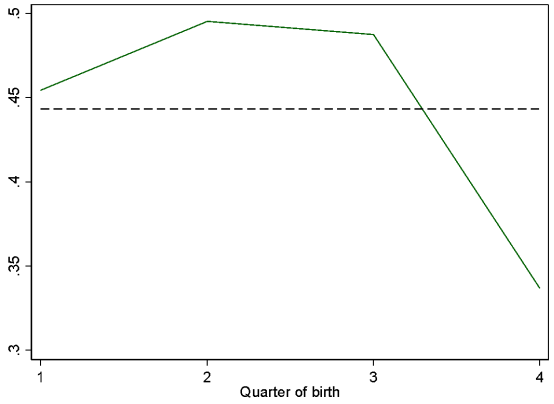
A. Average number of STI cases per 1,000 residents in 3 quarters to date



B. Average number of gastro-intestinal disease cases per 1,000 residents in 3 quarters to date



C. Average number of influenza cases per 1,000 residents in 3 quarters to date

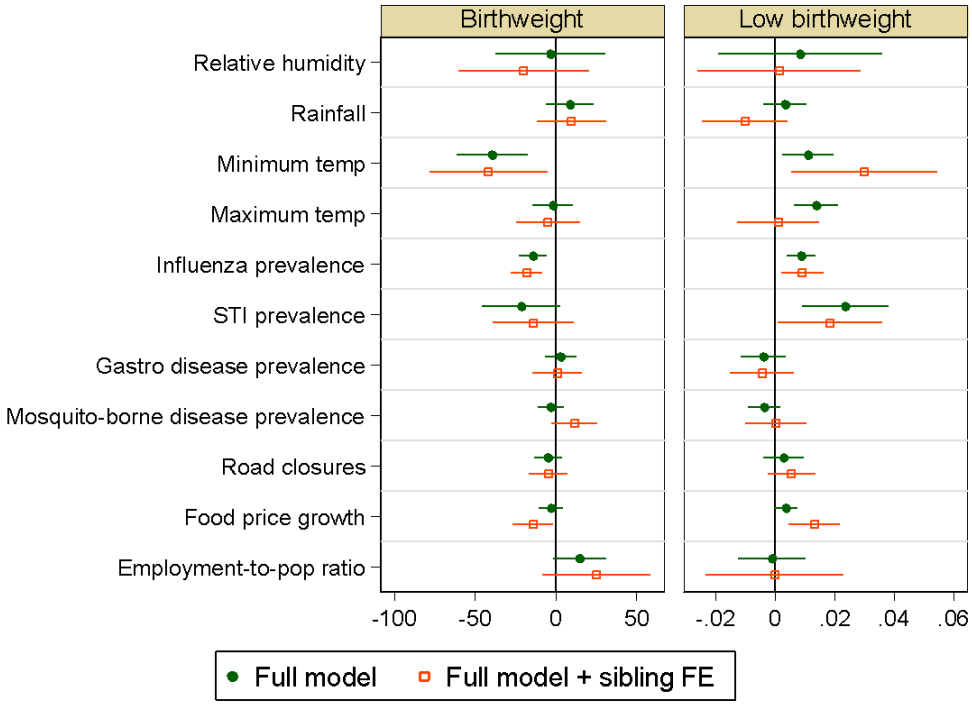


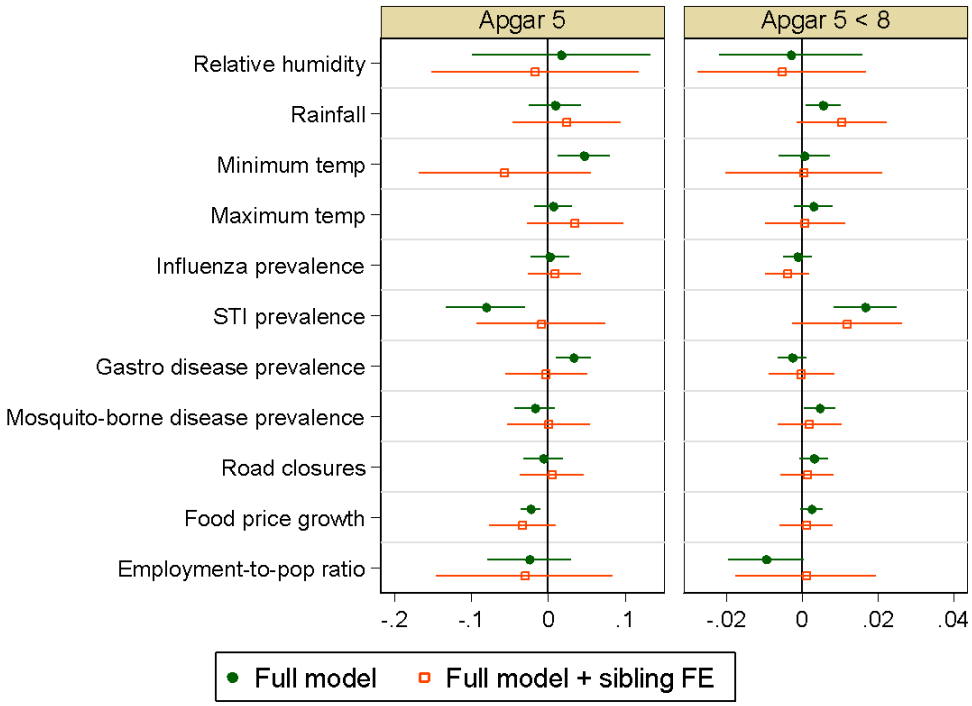
D. Average number of mosquito-borne disease cases per 1,000 residents in 3 quarters to date

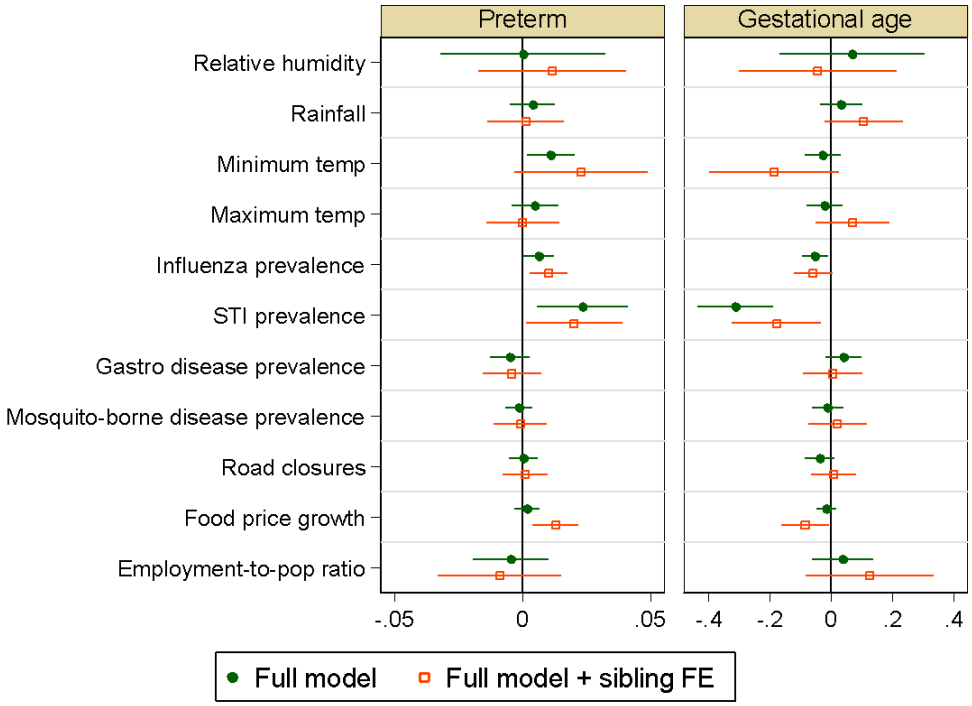
Note: case rates are based on cases within the region per resident of any age – a breakdown by sex and age is not available at the regional level. Seasonal patterns vary by region – for instance, while there is little seasonality in aggregate STI rates overall in the NT, this is because case rates are low and vary little in the capital city Darwin, where the population is largest. Seasonality is greater in other regions, and the seasonal patterns vary across regions (for instance, with a peak in STIs in June in Alice, in September in Barkly, and in March in Arnhem. Quarters are based on the calendar year, where Q1 is January-March.

Appendix C – Model estimates using alternative outcome measures

The below figures show model estimates are similar using definitions of the outcome variable, based on averages (birthweight, Apgar score, gestational age), as well as measures focusing on the lower end of the distribution (low birthweight, preterm birth, low Apgar scores).




Apgar 5
Apgar 5 < 8
Relative humidity
Rainfall
Minimum temp
Maximum temp
Influenza prevalence
STI prevalence
Gastro disease prevalence
Mosquito-borne disease prevalence
Road closures
Food price growth
Employment-to-pop ratio
-.2
-.1
0
.1
-.02
0
.02
.04
Full model
Full model + sibling FE

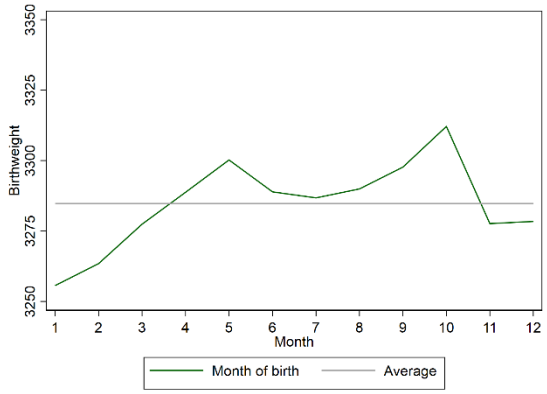

Preterm
Gestational age
Relative humidity
Rainfall
Minimum temp
Maximum temp
Influenza prevalence
STI prevalence
Gastro disease prevalence
Mosquito-borne disease prevalence
Road closures
Food price growth
Employment-to-pop ratio
-.05
0
.05
-.4
-.2
0
.2
.4
Full model
Full model + sibling FE

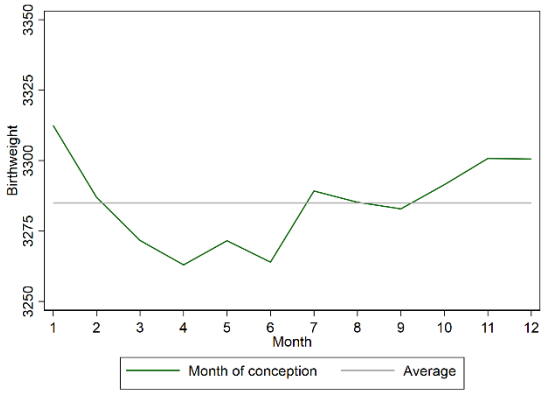
Appendix D– Seasonal patterns in birth outcomes by month of conception vs month of birth

By month of birth

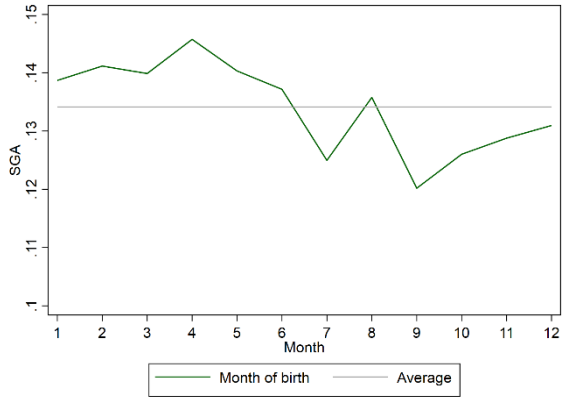
By month of conception

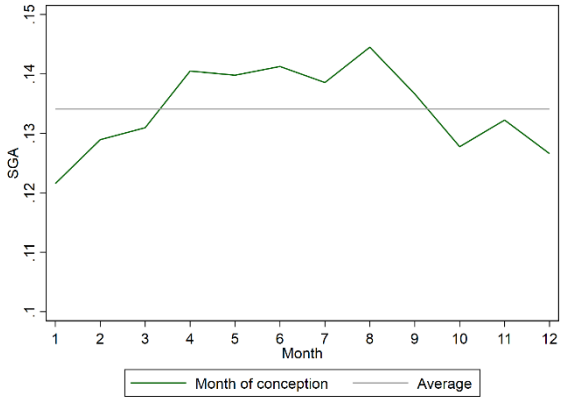
Birthweight



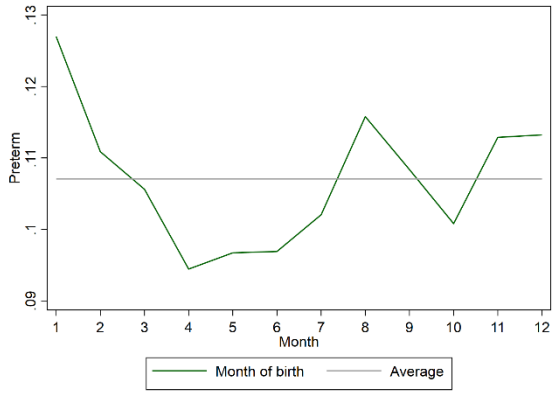


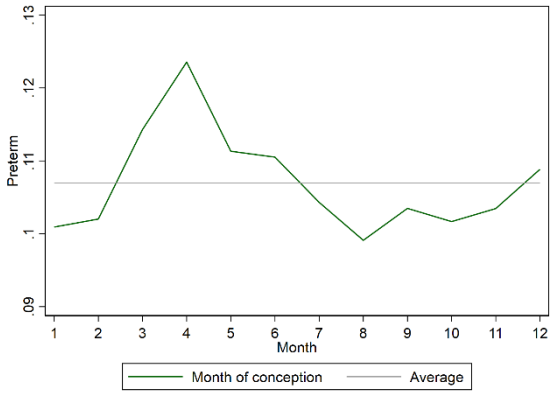
Small for gestational age



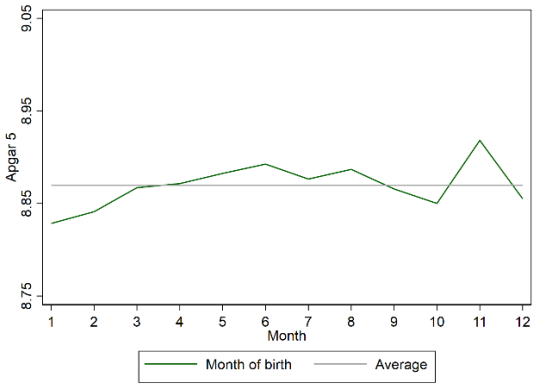


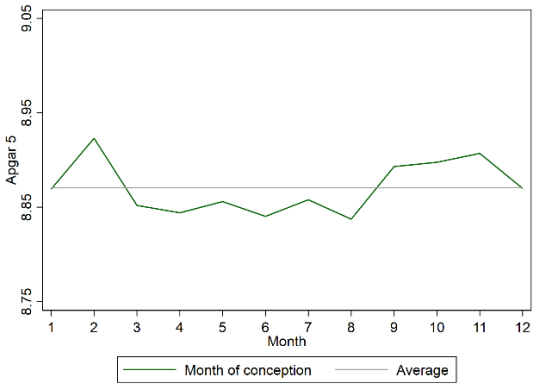
Preterm birth





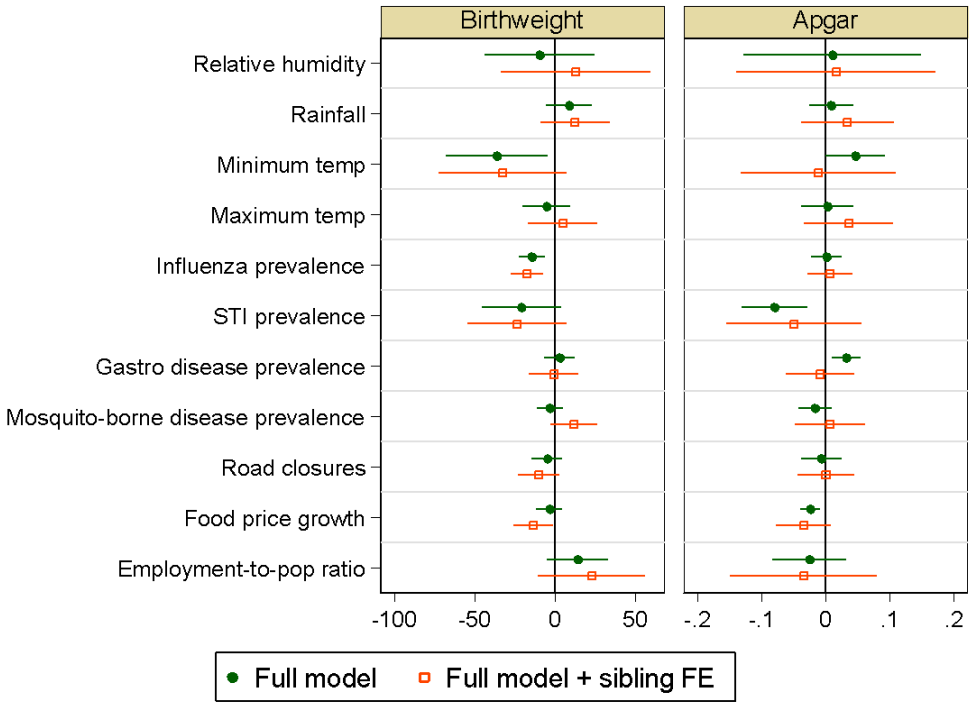
Apgar 5

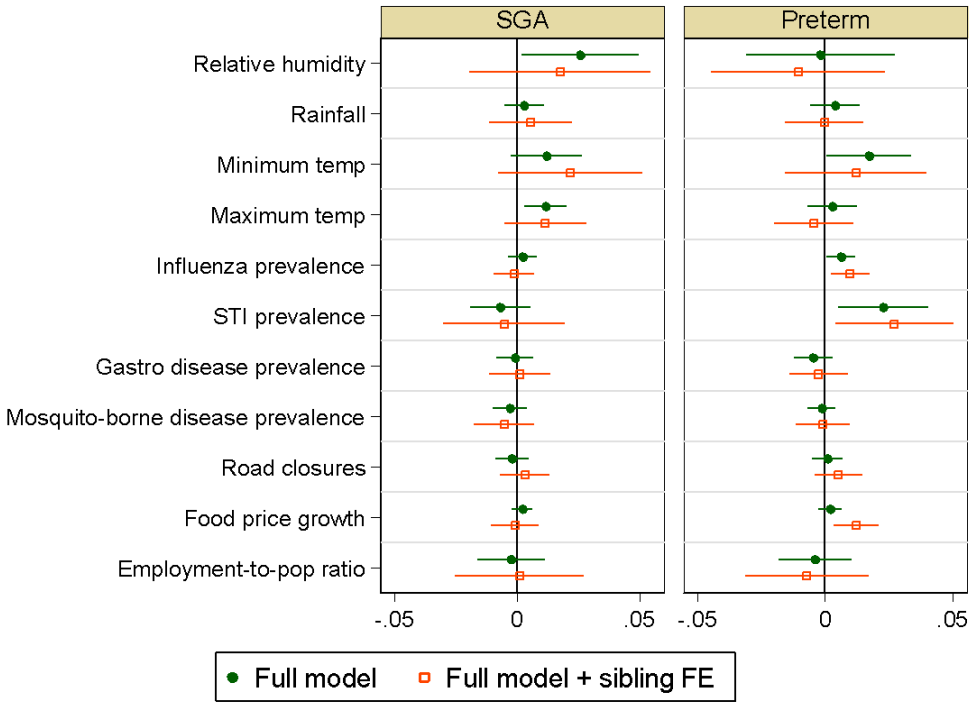

Apgar 5
9.05
8.95
8.85
8.75
1 2 3 4 5 6 7 8 9 10 11 12
Month
Month of birth
Average

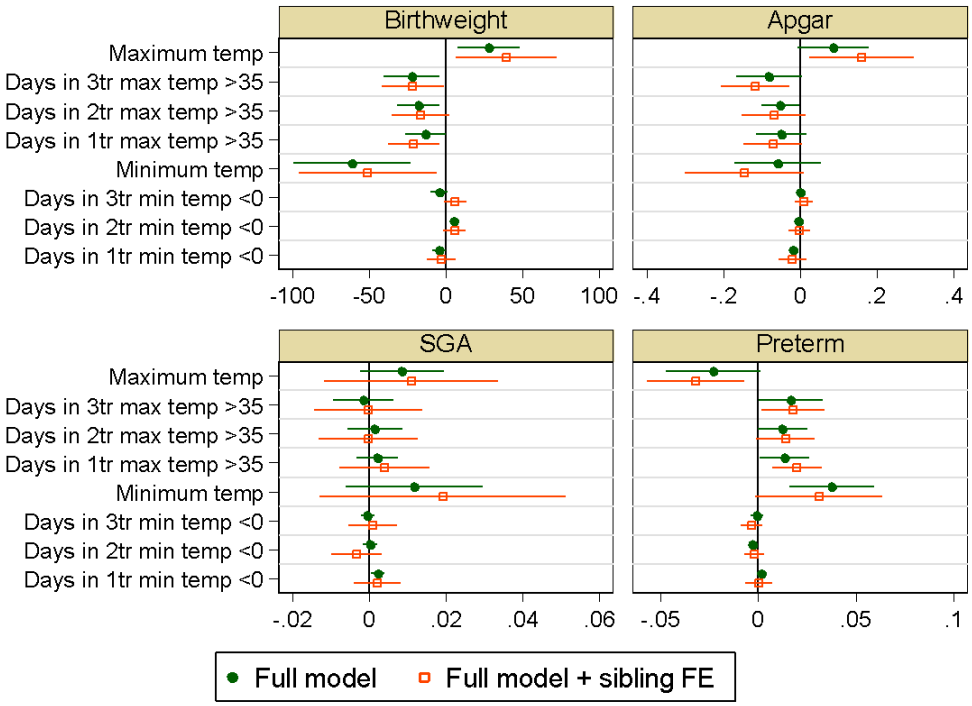

Apgar 5
9.05
8.95
8.85
8.75
1 2 3 4 5 6 7 8 9 10 11 12
Month
Month of conception
Average

Appendix E – Alternative cluster size for location FE

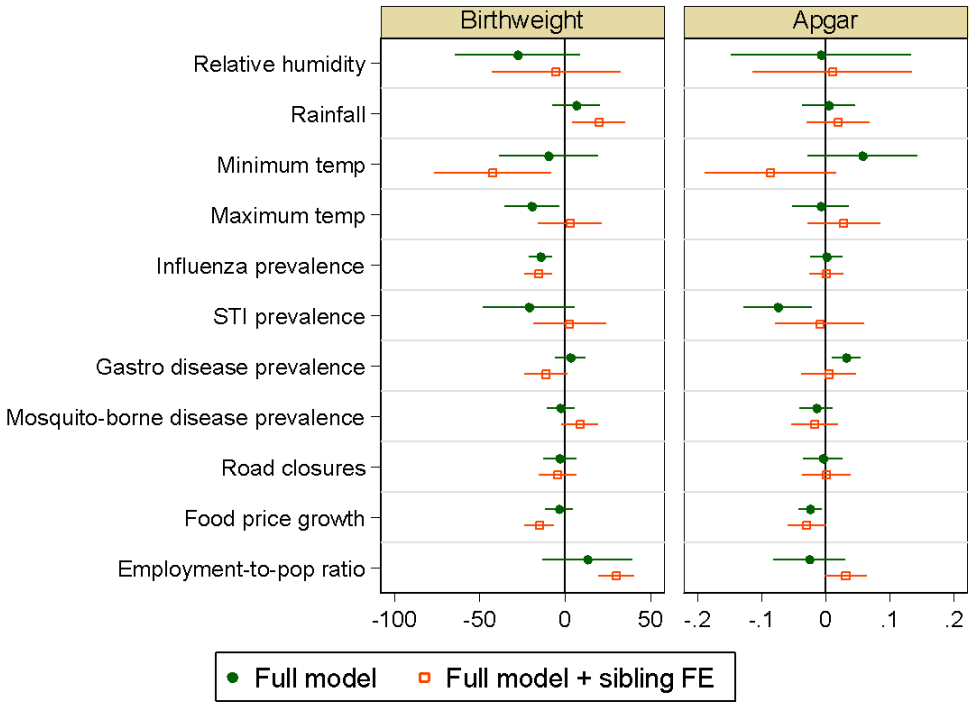
25 mile clusters:

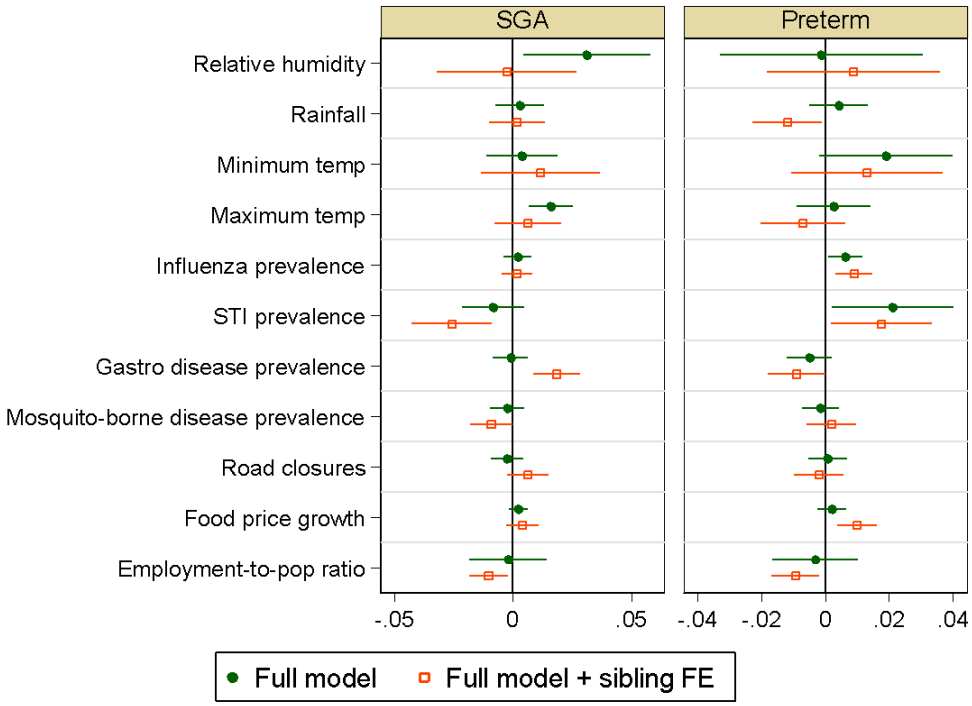


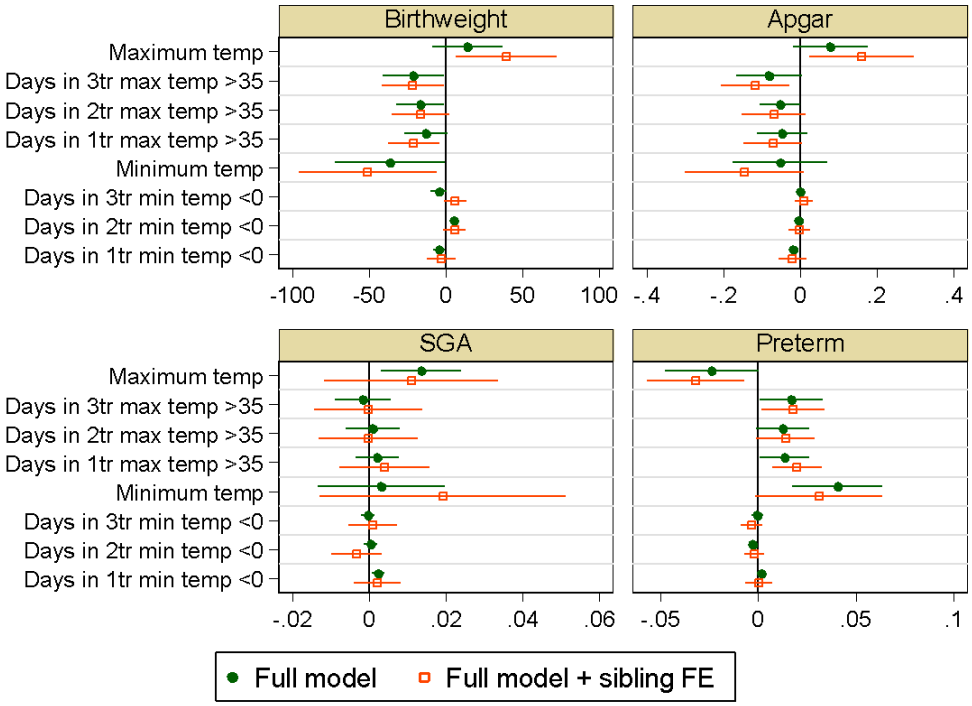




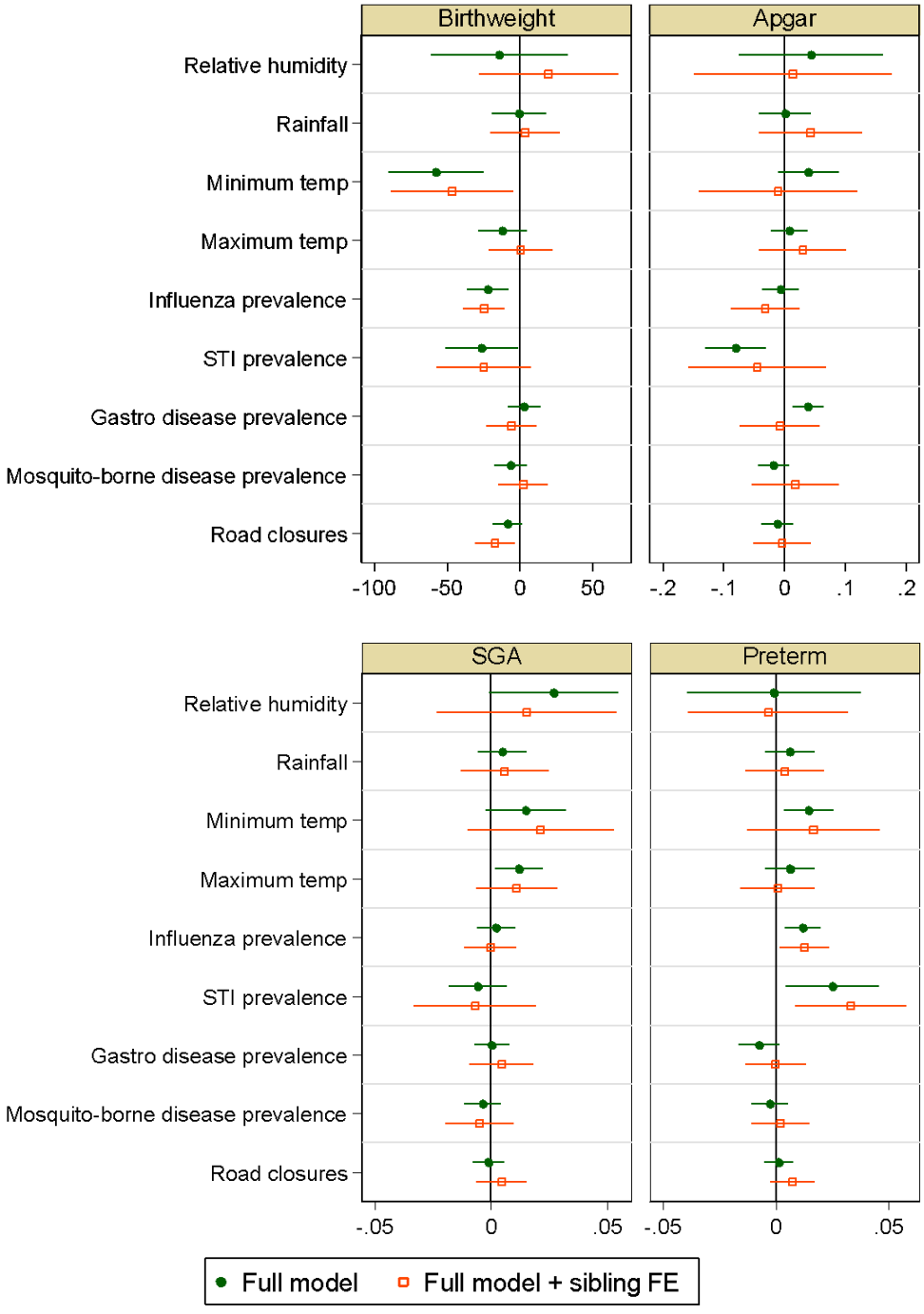
10 mile clusters:




SGA
Preterm
Relative humidity
Rainfall
Minimum temp
Maximum temp
Influenza prevalence
STI prevalence
Gastro disease prevalence
Mosquito-borne disease prevalence
Road closures
Food price growth
Employment-to-pop ratio
-.05
0
.05
-.04
-.02
0
.02
.04
Full model
Full model + sibling FE


Birthweight
Apgar
Maximum temp
Days in 3tr max temp >35
Days in 2tr max temp >35
Days in 1tr max temp >35
Minimum temp
Days in 3tr min temp <0
Days in 2tr min temp <0
Days in 1tr min temp <0
-100
-50
0
50
100
-.4
-.2
0
.2
.4
SGA
Preterm
-.02
0
.02
.04
.06
-.05
0
.05
.1
Full model
Full model + sibling FE

Appendix F – Regression results with month-year (interacted) fixed effects



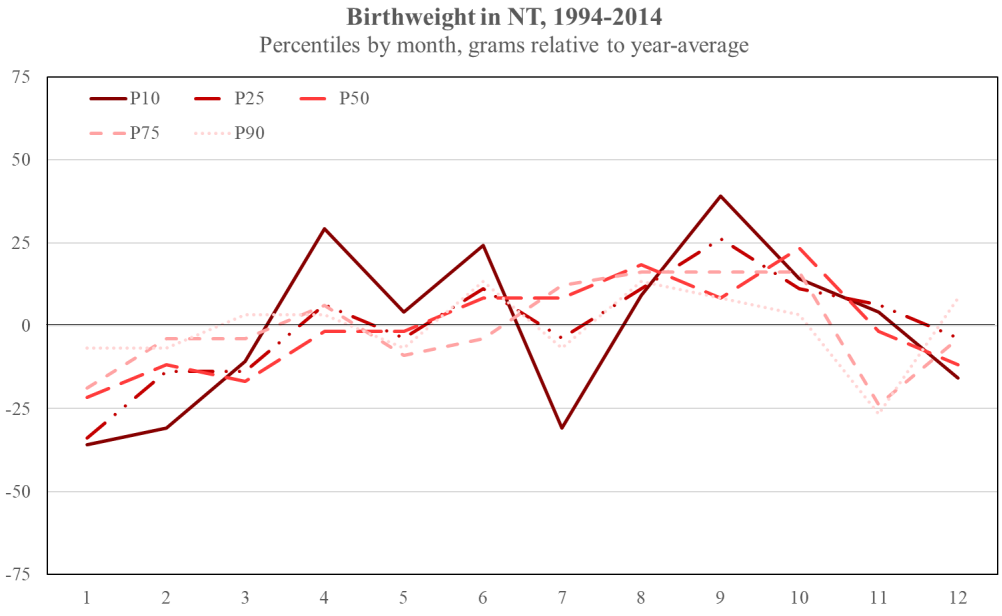
Regression coefficients with 95% confidence intervals. All explanatory variables are standardised to have a mean of zero and standard deviation of 1. The higher points represent estimates from the benchmark model, but with month-year interacted instead of entering into the regression separately. The lower points represent the same model, with the inclusion of sibling fixed effects, for the subpopulation with siblings within the dataset. Standard errors clustered at the level of the fixed effects.

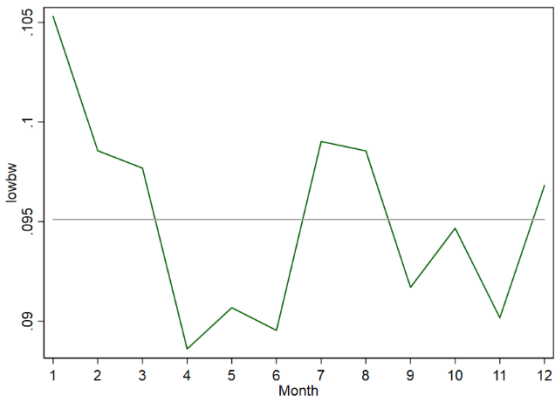
Appendix G – Alternative measures of seasonal patterns in outcomes

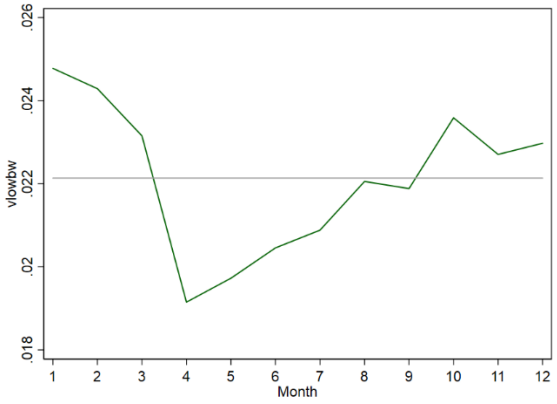
Seasonal patterns at mean values vs tails of distribution

The below graphs show the distribution of seasonal patterns in birthweight. They show that seasonality is most accentuated at the lower end of the distribution, with the 10th percentile showing more seasonal variation than the higher percentiles. The following graphs show that the seasonal pattern in the share of births that are low birthweight (<2500g) and very low birthweight (<1500g) is similar (but inverse) to the pattern in average birthweight. We see the same for average Apgar 5 scores compared with binary variables based on the score being below 7 (the standard cut-off for 'normal' scores (Razaz, Cnattingius, and Joseph 2019) – about 3.5% of births in the NT) or below 8 (about 5.5% of births in the NT), and for gestational age in weeks compared with probability of preterm birth.

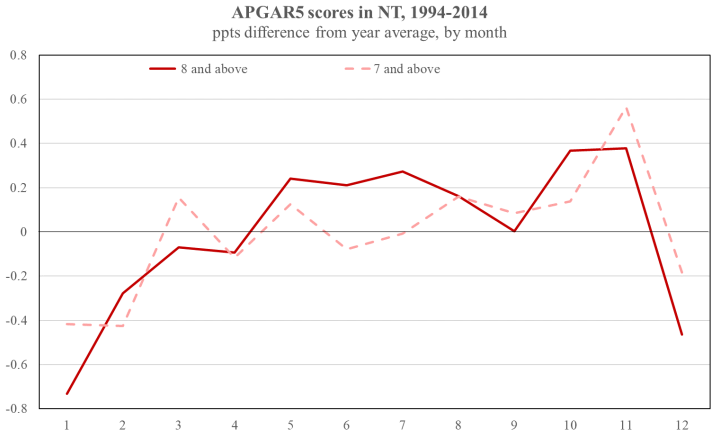
Birthweight



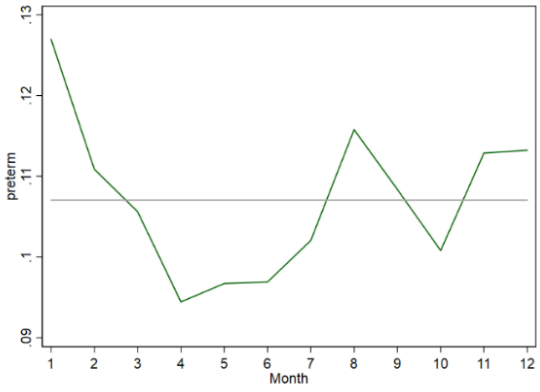


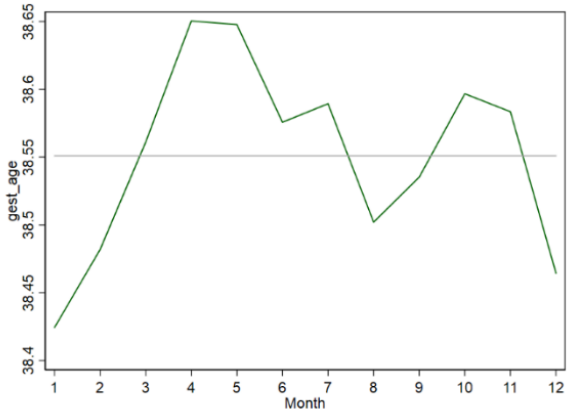


Apgar scores



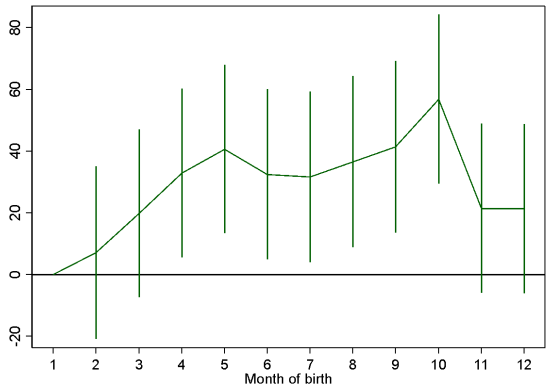
Gestational age and preterm birth



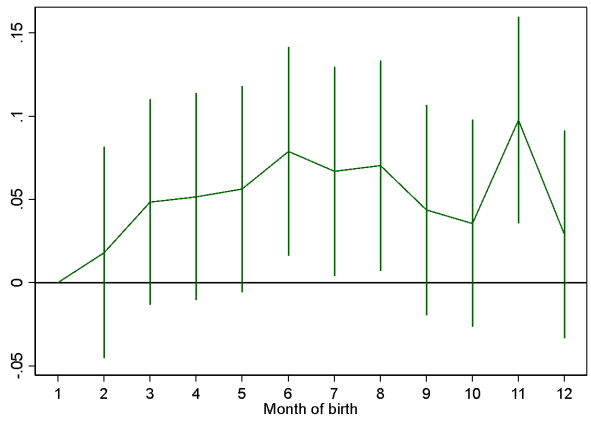


Seasonal patterns in analysis sample, with confidence intervals

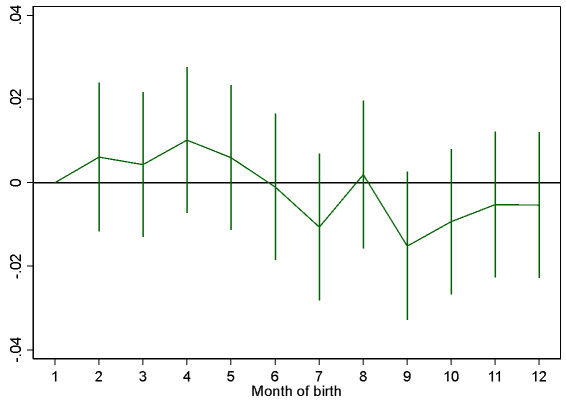
The graphs below show the seasonal patterns in birth outcomes for the analysis sample (births from 2005 to 2014). They show coefficients from a simple regression of each outcome on month of birth, with no other variables included in the regression. The bands represent 95% confidence intervals, and the base category is January. The outcomes variables are not standardised in these estimates.



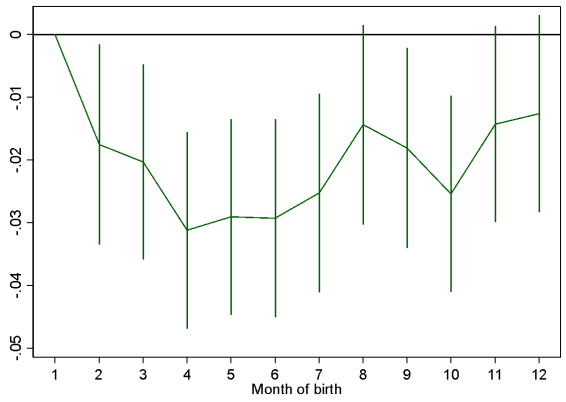
A. Birthweight



B. Apgar 5



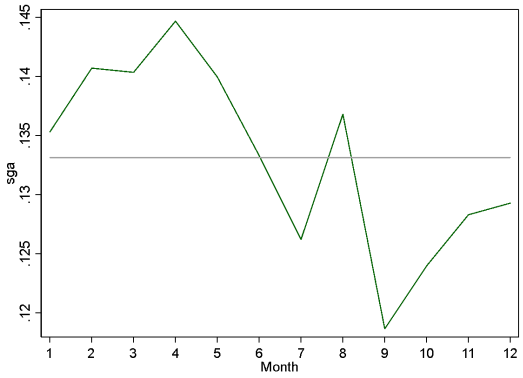
C. SGA



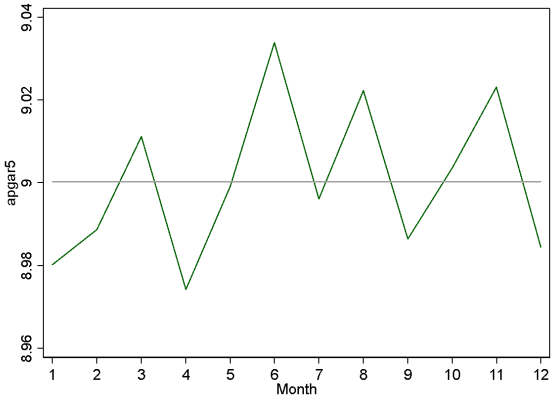
D. Preterm birth

Seasonal patterns in outcomes for full-term births

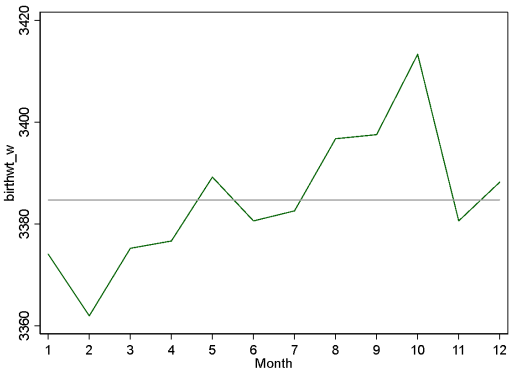
Seasonal patterns in birth outcomes, with sample restricted to babies born at or after week 37 of pregnancy (for comparison with McGrath et al. (2005)).



A. Small for gestational age

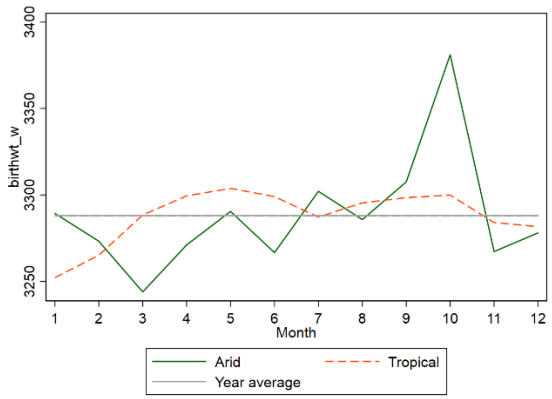


B. Apgar scores

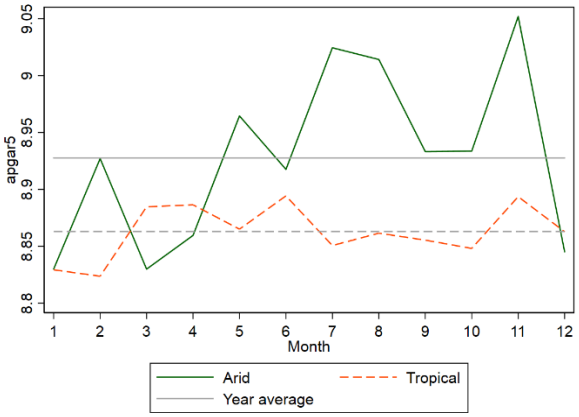


C. Birthweight

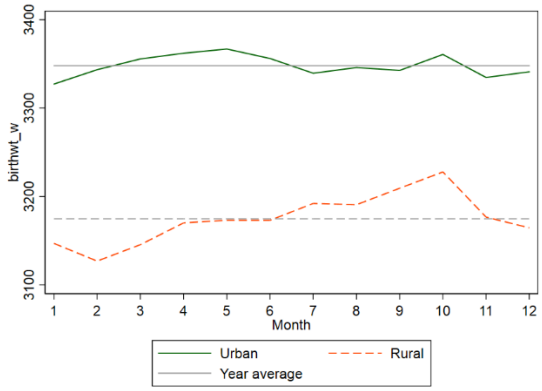
Appendix H – Further descriptive statistics



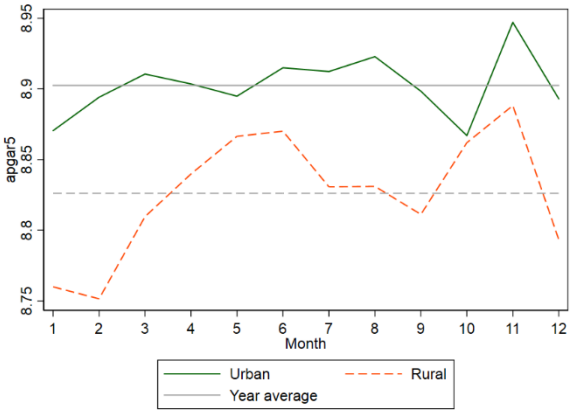
A. Birthweight in arid vs tropical regions of the NT



B. Apgar 5 in arid vs tropical regions of the NT



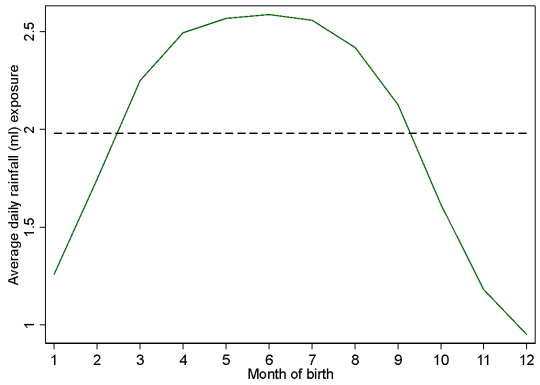
C. Birthweight in urban vs rural locations in the NT



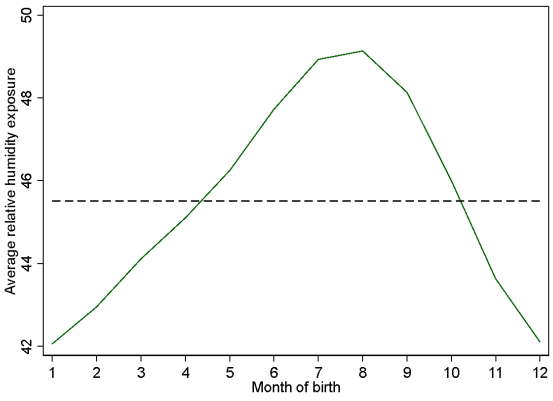
D. Apgar 5 in urban vs rural locations in the NT

Appendix I – Seasonal risk factors

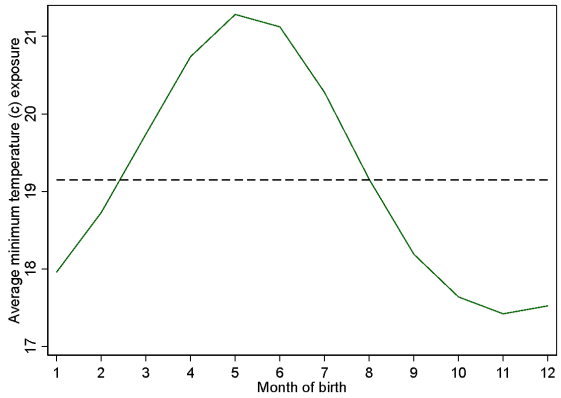
The figure below shows average levels of humidity, rainfall, minimum and maximum temperatures that babies born in each month would be exposed to during pregnancy. For example, panel A shows average relative humidity in the 39 weeks to each date, averaged across all years from 1995-2020, and across all locations in the NT. As may be expected, these measures are highly correlated with one another. All measures peak in the middle of the year (i.e. for babies that were born in the dry season, but were in utero throughout the whole wet season), and are lowest towards the end of the year.



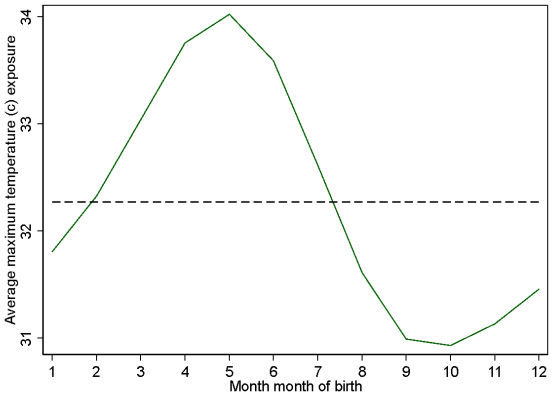
A. 9 month moving average relative humidity (%)



B. 9 month moving average daily rainfall (ml)



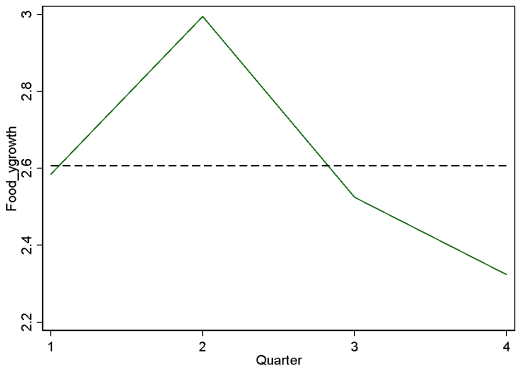
C. 9 month moving average daily minimum temperature (c)



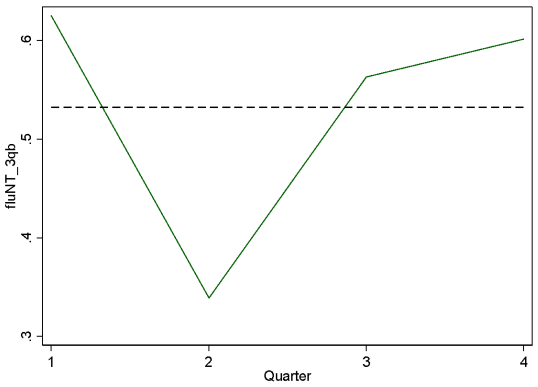
D. 9 month moving average daily maximum temperature

Babies born in late summer and early autumn are also exposed to stronger average growth in food prices during pregnancy (Panel A below), meaning that women who give birth at this time of year may be least able to afford nutritious food during pregnancy. But at the

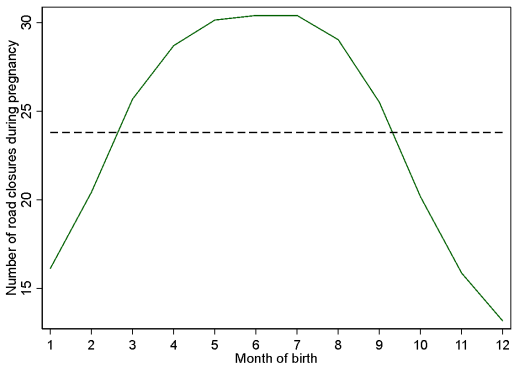
same time, flu exposure is lowest (Panel B below), suggesting two potentially counteracting forces. Babies born around this time are also more likely to be affected by road closures – with those born in June experiencing the highest average number of road closures in the 9 months to birth (Panel C below).



A. Year-end growth in food prices by quarter (%)

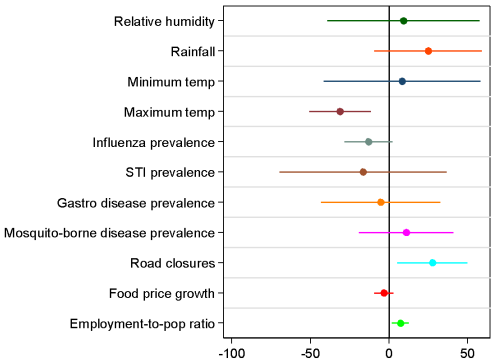


B. Reported influenza cases in previous 3 quarters, by quarter

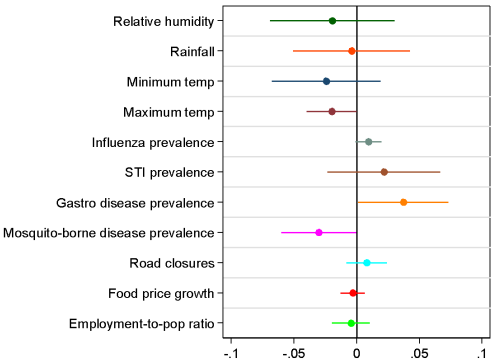


C. Average number of road closures in previous 9 months

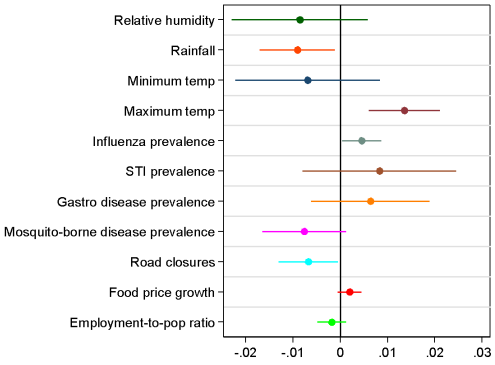
The below graph presents the coefficients and 95% confidence intervals from a regression of each seasonal risk factor on birth outcomes, with no fixed effects or other controls in the regression. All seasonal risk factors summarise exposure throughout the entire pregnancy and are standardised with mean zero and standard deviation of one, such that the point estimates show the effect associated with a one standard deviation change in each risk factor.



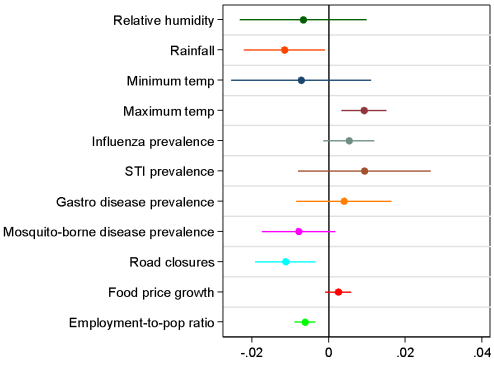
A. Birthweight (grams)



A. Apgar 5 (score from 0 to 10)



B. Probability of small for gestational age (SGA)



C. Probability of preterm birth

Note: These graphs show estimates from bivariate regressions of each birth outcome on each individual risk factor, without any controls, with standard errors clustered at the location level. Every line in the graph is from a different regression. All seasonal risk factors are standardised with mean zero and standard deviation of 1.

Appendix J – Tables of regression results from main model

Outcome: Birthweight (grams)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Relative humidity	9.382											-3.293	12.88
	(14.22)											(17.03)	(23.80)
Rainfall		12.88**										8.756	12.73
		(6.331)										(7.398)	(11.02)
Minimum temperature			-40.1**									-39***	-32.37
			(10.20)									(11.12)	(20.31)
Maximum temperature				-3.014								-1.828	4.992
				(8.287)								(6.297)	(11.08)
Influenza prevalence					-11***							-14***	-17***
					(3.361)							(4.456)	(5.15)
STI prevalence						-14.23						-21.44*	-23.47
						(9.163)						(12.30)	(15.71)
Gastro-related disease prevalence							3.265					3.079	-0.709
							(3.936)					(4.885)	(7.86)
Mosquito-borne disease prevalence								-3.165				-3.073	11.84
								(3.443)				(3.966)	(7.43)
Road closures									1.901			-4.865	-10.09
									(1.953)			(4.390)	(6.55)
Food price growth										1.812		-2.967	-13.1**
										(3.947)		(3.857)	(6.35)
Employment to population ratio											2.726	14.79*	22.96
											(7.169)	(8.253)	(17.03)
Constant	3,336***	3,345***	3,240***	3,314***	3,311***	3,331***	3,311***	3,307***	3,316***	3,315***	3,322***	3,303***	3,260***
	(31.67)	(21.40)	(27.22)	(15.04)	(15.03)	(22.10)	(13.32)	(15.39)	(15.14)	(15.36)	(24.49)	(61.74)	(61.39)
Observations	35,199	35,199	35,199	35,199	35,199	35,199	35,199	35,199	35,199	35,199	35,199	35,199	18,359
R-squared (within clusters)	0.058	0.058	0.059	0.058	0.058	0.058	0.058	0.058	0.058	0.058	0.058	0.059	0.037
Test of stat. significance of seasonal risks (F or T stat)	0.435	4.142	15.49	0.132	10.39	2.413	0.688	0.845	0.947	0.211	0.145	4.425	1.921
Time fixed effects	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Location cluster fixed effects	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Sibling fixed effects	N	N	N	N	N	N	N	N	N	N	N	N	Y
Control for baby's sex	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

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

Outcome: Apgar 5 score	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Relative humidity	-0.024											0.017	0.017
	(0.0282)											(0.0583)	(0.079)
Rainfall		-0.011										0.00904	0.034
		(0.0169)										(0.0172)	(0.037)
Minimum temperature			0.049**									0.05***	-0.011
			(0.0202)									(0.0170)	(0.062)
Maximum temperature				0.0084								0.0064	0.036
				(0.0116)								(0.0125)	(0.035)
Influenza prevalence					0.0032							0.0019	0.007
					(0.0116)							(0.0126)	(0.018)
STI prevalence						-0.06**						-0.08***	-0.050
						(0.0263)						(0.0260)	(0.054)
Gastro-related disease prevalence							0.0202*					0.03***	-0.009
							(0.0104)					(0.0113)	(0.027)
Mosquito-borne disease prevalence								-0.0107				-0.0172	0.007
								(0.0104)				(0.013)	(0.028)
Road closures									-0.011			-0.006	0.0007
									(0.0079)			(0.01)	(0.023)
Food price growth										0.02***		-0.02***	-0.034
										(0.007)		(0.0067)	(0.022)
Employment to population ratio											-0.020	-0.025	-0.034
											(0.03)	(0.0276)	(0.059)
Constant	8.83***	8.85***	8.97***	8.88***	8.88***	8.96***	8.86***	8.86***	8.87***	8.87***	8.8***	8.97***	8.38***
	(0.0630)	(0.0565)	(0.0560)	(0.0303)	(0.0304)	(0.0404)	(0.0278)	(0.0392)	(0.0314)	(0.0300)	(0.09)	(0.186)	(0.212)
Observations	35,154	35,154	35,154	35,154	35,154	35,154	35,154	35,154	35,154	35,154	35,154	35,154	18,330
R-squared (within clusters)	0.019	0.019	0.019	0.019	0.019	0.019	0.019	0.019	0.019	0.019	0.019	0.020	0.016
Test of stat. significance of seasonal risks (F or T stat)	0.715	0.420	5.990	0.523	0.0774	5.861	3.762	1.060	2.074	11.03	0.458	5.155	0.627
Time fixed effects	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Location cluster fixed effects	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Sibling fixed effects	N	N	N	N	N	N	N	N	N	N	N	N	Y
Control for baby's sex	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

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

Outcome: Pr SGA	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Relative humidity	-0.0023											0.0220*	0.0176
	(0.007)											(0.0115)	(0.019)
Rainfall		-0.0009										0.00310	0.0056
		(0.0047)										(0.004)	(0.009)
Minimum temperature			0.01***									0.0130*	0.0218
			(0.004)									(0.0071)	(0.015)
Maximum temperature				0.00349								0.01**	0.0117
				(0.0035)								(0.004)	(0.0086)
Influenza prevalence					0.00125 (0.0024)							0.0024	-0.0011
												(0.003)	(0.0042)
STI prevalence						-0.007						-0.007	-0.0051
						(0.006)						(0.006)	(0.013)
Gastro-related disease prevalence							-0.0019					-0.0006	0.0013
							(0.0029)					(0.004)	(0.0064)
Mosquito-borne disease prevalence								-0.001				-0.0028	-0.0052
								(0.0036)				(0.004)	(0.0063)
Road closures									-0.0008 (0.0025)			-0.002	0.0034
												(0.003)	(0.0052)
Food price growth										0.0013		0.0023	-0.0007
										(0.002)		(0.002)	(0.005)
Employment to population ratio											-0.001	-0.003	0.00108
											(0.006)	(0.008)	(0.0135)
Constant	0.13***	0.14***	0.16***	0.14***	0.14***	0.15***	0.14***	0.14***	0.14*** (0.0091)	0.14***	0.13***	0.21***	0.15***
	(0.019)	(0.015)	(0.013)	(0.0094)	(0.01)	(0.013)	(0.01)	(0.0120)		(0.0093)	(0.021)	(0.036)	(0.046)
Observations	35,205	35,205	35,205	35,205	35,205	35,205	35,205	35,205	35,205	35,205	35,205	35,205	18,365
R-squared (within clusters)	0.014	0.014	0.014	0.014	0.014	0.014	0.014	0.014	0.014	0.014	0.014	0.014	0.013
Test of stat. significance of seasonal risks (F or T stat)	0.107	0.0327	7.167	1.015	0.280	1.711	0.451	0.106	0.111	0.529	0.0295	1.745	0.739
Time fixed effects	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Location cluster fixed effects	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Sibling fixed effects	N	N	N	N	N	N	N	N	N	N	N	N	Y
Control for baby's sex	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

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

Outcome: Preterm birth	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Relative humidity	-0.009											0.0004	-0.011
	(0.006)											(0.016)	(0.017)
Rainfall		-0.001										0.004	-0.0002
		(0.004										(0.004)	(0.008)
Minimum temperature			0.01**									0.01**	0.012
			(0.004)									(0.005)	(0.014)
Maximum temperature				0.004*								0.0049	-0.004
				(0.002)								(0.005)	(0.008)
Influenza prevalence					0.005**							0.007**	0.010**
					(0.002)							(0.003)	(0.004)
STI prevalence						0.02**						0.02***	0.027**
						(0.007)						(0.009)	(0.012)
Gastro-related disease prevalence							-0.003					-0.005	-0.002
							(0.003)					(0.004)	(0.006)
Mosquito-borne disease prevalence								-0.002				-0.001	-0.001
								(0.002)				(0.003)	(0.005)
Road closures									-0.003			0.0006	0.005
									(0.002)			(0.003)	(0.005)
Food price growth										0.0003		0.002	0.01***
										(0.002)		(0.002)	(0.005)
Employment to population ratio											-0.0005	-0.004	-0.007
											(0.006)	(0.007)	(0.012)
Constant	0.14*** (0.0205)	0.16*** (0.0190)	0.18*** (0.0117)	0.16*** (0.0127)	0.16*** (0.0129)	0.14*** (0.0172)	0.17*** (0.0121)	0.16*** (0.0126)	0.16*** (0.0132)	0.16*** (0.0132)	0.16*** (0.0173)	0.16*** (0.0470)	0.21*** (0.004)
Observations	35,205	35,205	35,205	35,205	35,205	35,205	35,205	35,205	35,205	35,205	35,205	35,205	18,365
R-squared (within clusters)	0.021	0.021	0.021	0.021	0.021	0.021	0.021	0.021	0.021	0.021	0.021	0.021	0.007
Test of stat. significance of seasonal risks (F or T stat)	2.323	0.0391	5.790	3.001	5.668	6.885	0.974	0.626	1.742	0.0126	0.00707	2.366	1.593
Time fixed effects	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Location cluster fixed effects	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Sibling fixed effects	N	N	N	N	N	N	N	N	N	N	N	N	Y
Control for baby's sex	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

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

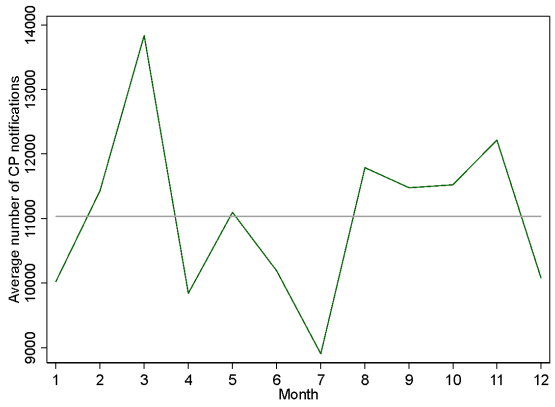
Appendix K – Correction for multiple hypothesis testing

	Birthweight	Preterm	SGA	Apgar 5
Relative humidity				
Coefficient	-3.293	0.000364	0.0220*	0.0169
P-value	0.85	0.98	0.06	0.77
Sharpened Q	1.00	1.00	0.22	1.00
Rainfall				
Coefficient	8.756	0.00413	0.00310	0.00904
P-value	0.24	0.35	0.49	0.60
Sharpened Q	0.67	0.91	1.00	1.00
Minimum temperature				
Coefficient	-39.43***	0.0111**	0.0130*	0.0466***
P-value	0.00	0.02	0.07	0.01
Sharpened Q	0.02	0.09	0.24	0.05
Maximum temperature				
Coefficient	-1.828	0.00492	0.00994**	0.00644
P-value	0.77	0.29	0.01	0.61
Sharpened Q	1.00	0.76	0.07	1.00
Influenza prevalence				
Coefficient	-14.09***	0.00653**	0.00241	0.00194
P-value	0.00	0.03	0.39	0.88
Sharpened Q	0.03	0.13	0.91	1.00
STI prevalence				
Coefficient	-21.44*	0.0236***	-0.00662	-0.0809***
P-value	0.08	0.01	0.30	0.00
Sharpened Q	0.26	0.06	0.76	0.03
Gastro-related disease prevalence				
Coefficient	3.079	-0.00479	-0.000583	0.0327***
P-value	0.53	0.23	0.88	0.00
Sharpened Q	1.00	0.67	1.00	0.04
Mosquito-borne disease prevalence				
Coefficient	-3.073	-0.00119	-0.00281	-0.0172
P-value	0.44	0.66	0.43	0.20
Sharpened Q	0.91	1.00	0.91	0.67
Road closures				
Coefficient	-4.865	0.000556	-0.00162	-0.00643
P-value	0.27	0.85	0.61	0.62
Sharpened Q	0.75	1.00	1.00	1.00
Food price growth				
Coefficient	-2.967	0.00189	0.00230	-0.0230***
P-value	0.44	0.44	0.24	0.00
Sharpened Q	0.91	0.91	0.67	0.02
Employment to population ratio				
Coefficient	14.79*	-0.00447	-0.00295	-0.0246
P-value	0.08	0.55	0.70	0.38
Sharpened Q	0.25	1.00	1.00	0.91
Constant	3,303***	0.158***	0.213***	8.974***
Observations	35,199	35,205	35,205	35,154
R-squared (within clusters)	0.059	0.021	0.014	0.020
Time fixed effects	Yes	Yes	Yes	Yes
Location cluster fixed effects	Yes	Yes	Yes	Yes
Sibling fixed effects	No	No	No	No
Control for baby's sex	Yes	Yes	Yes	Yes

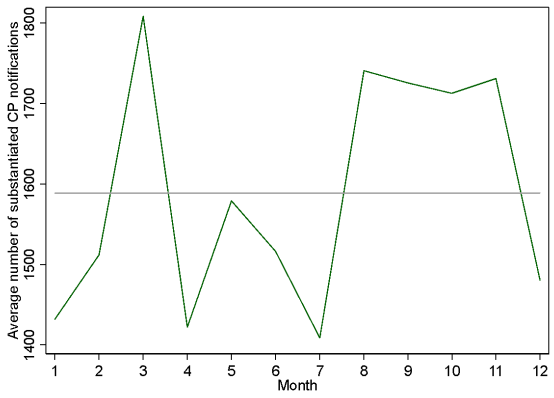
Clustered standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Appendix L – Seasonality in child protection notifications

There is a pattern in child protection notifications and substantiations, with report numbers stepping up in March, and again in the months around the time of the build up to the wet season. However, this increase in notifications around the time of the build-up to the wet season (i.e. when humidity is highest) is driven almost entirely by reports from school staff. Therefore, it is likely that this pattern reflects seasonality in reporting (to coincide with school term dates), and not seasonality in incidences. This is not necessarily evidence that there is not seasonality in domestic violence, but evidence that there is no clear seasonal pattern in notifications and substantiated incidences.

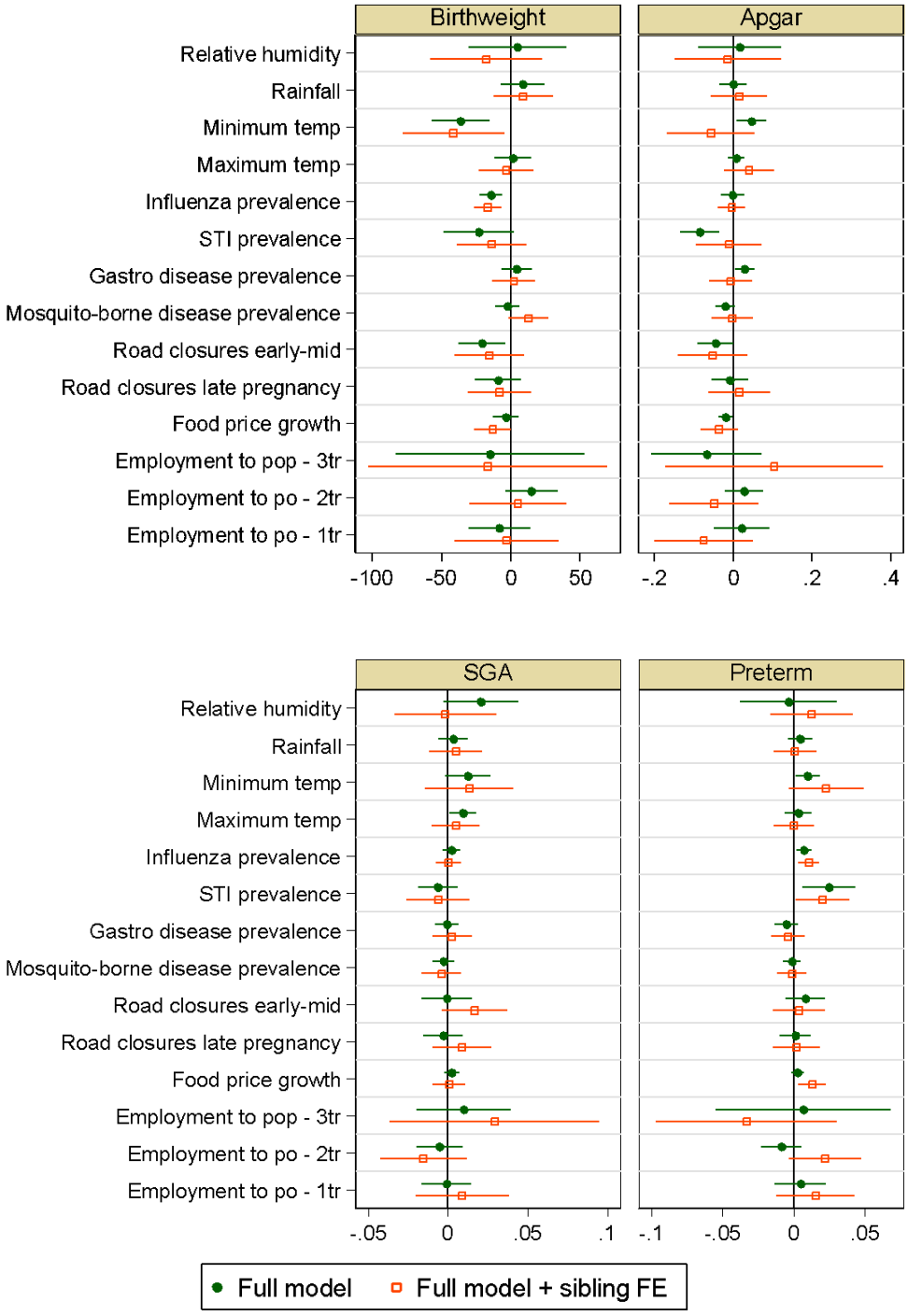


Notifications per month, 2003-2018



Substantiated cases per month, 2003-2018

Appendix M – Effects of road closures and employment by timing during pregnancy



Appendix N – Impact of STI prevalence by frequency of antenatal visits

	Birthweight	Apgar 5	SGA	Preterm birth
STI prevalence	-16.18	-0.251**	-0.0172	0.0444**
	(29.45)	(0.0981)	(0.0173)	(0.0215)
Constant	2,933***	8.164***	0.176***	0.385***
	(55.58)	(0.167)	(0.0363)	(0.0296)
Number of observations	5,381	5,356	5,383	5,383
R-squared (within clusters)	0.011	0.032	0.007	0.008
Number of clusters	87	87	87	87
Time fixed effects	Y	Y	Y	Y
Location fixed effects	Y	Y	Y	Y
Control for sex of baby	Y	Y	Y	Y

Regression results for subsample of births with <=5 antenatal visits (2005-2014)

Appendix O – Effect of disease prevalence on live birth status

	(1)	(2)	(3)	(4)	(5)
Flu prevalence	0.00058				0.0005
	(0.0008)				(0.0008)
STI prevalence		-0.004***			-0.004***
		(0.001)			(0.0016)
Gastro-related disease prevalence			-0.0002		0.0006
			(0.0015)		(0.0016)
Mosquito-borne disease prevalence				-0.00115	-0.0014*
				(0.0007)	(0.0008)
Constant	0.99***	0.988***	0.99***	0.989***	0.987***
	(0.0032)	(0.0030)	(0.0036)	(0.0033)	(0.0030)
Observations	35,205	35,205	35,205	35,205	35,205
R-squared (within clusters)	0.008	0.008	0.008	0.008	0.009
Number of clusters	96	96	96	96	96
Time fixed effects	Y	Y	Y	Y	Y
Location fixed effects	Y	Y	Y	Y	Y
Control for sex of baby	Y	Y	Y	Y	Y

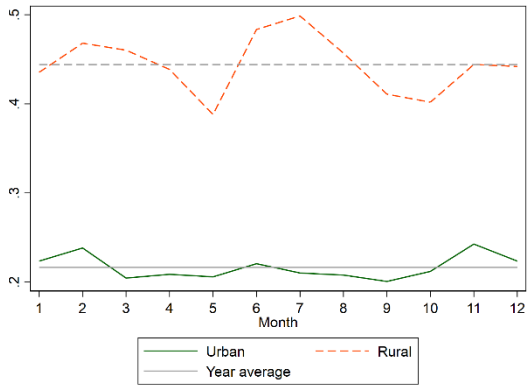
Regressions on live birth status (vs stillbirth)

Appendix P – Share of variance explained by seasonal risks

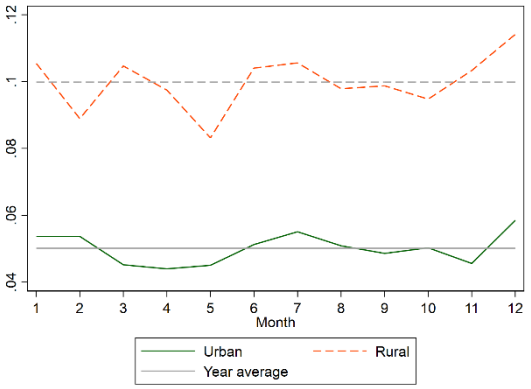
	Full population		Rural		Urban		
	R-squared without seasonal risks	Change in R-squared after including seasonal risks	R-squared without seasonal risks	Change in R-squared after including seasonal risks	R-squared without seasonal risks	Change in R-squared after including seasonal risks	Change in R-squared: ratio of rural to urban
Birthweight	0.0582	0.0008	0.0709	0.0022	0.0168	0.0015	1.4314
Preterm birth	0.0206	0.0008	0.0232	0.0006	0.0040	0.0015	0.4137
SGA	0.0139	0.0002	0.0201	0.0016	0.0018	0.0003	5.6830
Apgar 5	0.0191	0.0007	0.0015	0.0015	0.0158	0.0011	1.3097

Note: This table shows the R-squared in the benchmark mode, not including sibling fixed effects. 'R squared without seasonal risks' is the model including only time and location fixed effects, and controls for the baby's sex. The 'Change in R-squared from seasonal risks' is the R-squared from the full benchmark model, minus the 'R-squared without seasonal risks'. This shows the additional explanatory power of the seasonal risks. The final column shows the ratio of additional explanatory power from seasonal risks in the rural compared with the urban populations. A ratio above one suggests that the seasonal risks have higher explanatory power in the rural subpopulation.

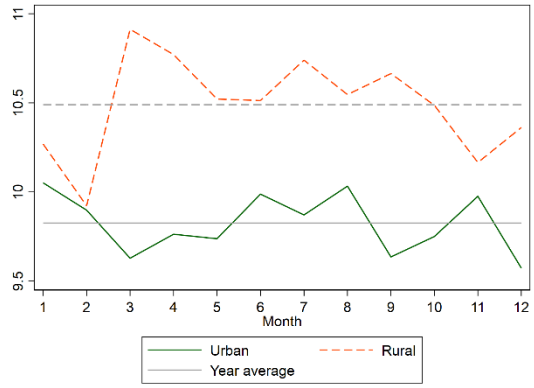
Appendix Q – Seasonal patterns in maternal characteristics, by month giving birth



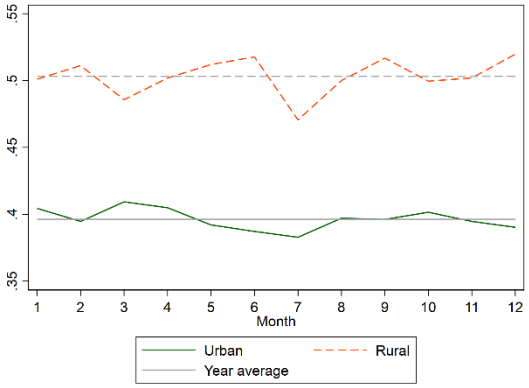
A. Smoking at first antenatal visit



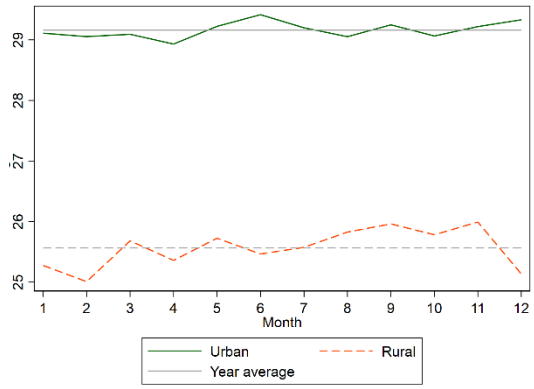
B. Drinking alcohol at first antenatal visit



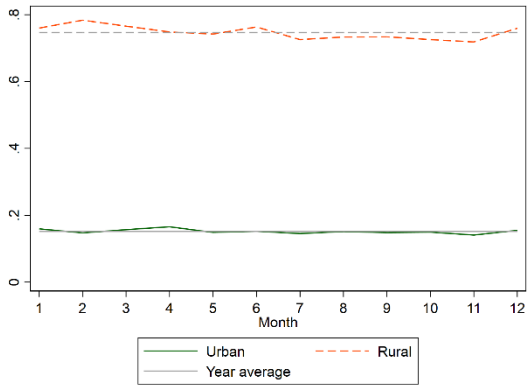
C. Number of antenatal visits



D. Whether any complications in the mother's medical history (before current pregnancy)



E. Mother's age at birth



F. Aboriginal status of mother (1 = Aboriginal, 0 = non-Aboriginal)

Appendix R – Explanatory power of month coefficients

Rural	OLS	Full model	Full model + sibling FE
Birthweight	0.00231	0.0011	0.0016
Apgar 5	0.00117	0.0007	0.0016
SGA	0.00185	0.0015	0.0011
Preterm	0.00225	0.0011	0.0012
Urban			
Birthweight	0.00081	0.0009	0.0012
Apgar 5	0.00045	0.0003	0.0015
SGA	0.00124	0.0014	0.0024
Preterm	0.00161	0.0015	0.0019
Total			
Birthweight	0.00063	0.0006	0.0005
Apgar 5	0.00027	0.0002	0.0008
SGA	0.00070	0.0005	0.0006
Preterm	0.00112	0.0007	0.0007

Added explanatory power of month fixed effects

Note: Each cell shows the R-squared for the model shown, minus the R-squared for that same model without controls for month of birth. For the sibling fixed effects model, the within-R-squared is shown. A higher number indicates that more of the variation in the outcome is explained by the month fixed effects (as opposed to the seasonal risks included in the model). We may expect that if the model does a good job of explaining seasonality, the added explanatory power of month coefficients would be smaller.

Chapter 3: The relationship between prenatal heat exposure and birth outcomes: how much does the heat metric matter?

Abstract

The impact of prenatal heat exposure on birth outcomes is well-established, but what is it about heat that affects prenatal development? Is it exposure to extreme temperatures, to moderate heat, or the confluence of heat and humidity? Despite the large body of research on heat exposure and birth outcomes, the literature lacks consistent measurement. This means we cannot extract practical recommendations around which heat conditions pose the greatest risk, and hence should be avoided during pregnancy. It also means we cannot predict the implications of climate change on neonatal health and healthcare needs at a population level. This chapter has two goals: first, to demonstrate that our conclusions around the existence and magnitude of the impact of heat exposure vary dramatically with the choice of heat exposure metric, and second, to make general recommendations for how heat exposure should be measured in future. We present analysis from Australia's Northern Territory — a region spanning tropical and arid climates. We compare commonly used heat exposure metrics, alongside additional metrics supported by theory. We find that metrics based on 'bands' of exposure and incorporating daily minimum as well as maximum measures provide the best fit; this is consistent with our theoretical understanding that the consequences of heat exposure do not increase linearly with higher temperatures, and instead, both moderate and extreme heat affect fetal development in different ways. Our preferred heat metric estimates that the impact of prenatal heat exposure on preterm birth is orders of magnitude larger than what would be implied by some metrics commonly used in the literature. Our findings underscore the importance of getting the measure of heat right, particularly in tropical climates.

3.1 Introduction

There is a wealth of evidence that ambient heat exposure adversely affects human health (Khosla, Jani, and Perera 2021; Ebi et al. 2021). The impact of prenatal heat exposure on a baby's health at birth is particularly well-studied (Nyadanu et al. 2024). Chersich et al.'s (2020) systematic review finds strong evidence that heat exposure is associated with a higher risk of preterm birth, lower average birthweight, and a higher risk of stillbirth. These effects may have long-term consequences: prenatal heat exposure has been linked to lower levels of physical health, mental health, education and earnings into adulthood (Isen, Rossin-Slater, and Walker 2017; Hu and Li 2019; Fishman, Carrillo, and Russ 2019; Ai and Tan 2023).

But what exactly is the nature of the relationship between heat and newborns' health? Is fetal development damaged by short-lived exposure to extreme temperatures? Or should we be more concerned about prolonged exposure to moderately high temperatures? Or is it instead the interaction of heat with humidity that leads to poorer health at birth?

We are unable to answer these questions because, to date, research on this topic has lacked consistency in the way that heat exposure is defined and measured. But answers to these questions are important: In order to anticipate the likely impact of climate change on healthcare needs, and mitigate the impacts of heat exposure on population health, it is important to understand what level and duration of heat exposure leads to poorer health at birth.

In this chapter, we analyse how much the way we measure heat exposure matters for our conclusions. To do this, we first set out options for measuring heat exposure, including the three measures of heat exposure that are most common in the literature, alongside two additional measures motivated by our conceptual framework: we call these our five 'heat metrics'. We estimate regressions with each of the five metrics, and combinations of multiple metrics together. We identify a preferred metric in our context, based on standard measures of goodness-of-fit. We then demonstrate how reliance on non-preferred metrics would affect our conclusions about the impact of prenatal heat exposure on birth outcomes.

In our analysis, we use data from the Northern Territory of Australia – a region spanning tropical and arid climates. The fact that we use data from these climate zones is important, because half of the world's population lives in climates like these, yet most empirical research on this topic uses data from cooler, less humid climates.[22] Using data from the Northern Territory means we can analyse exposure to very high temperatures (e.g. above 35 degrees Celsius), which are either not experienced at all in other climates, or are experienced so rarely that their effects cannot be reliably estimated.

There is also good reason to expect that the choice of heat metric is particularly important in tropical climates. In temperate climate zones, which cover much of the USA, Europe, and East Asia, correlations between various measures of heat exposure are high, meaning that one measure of heat exposure (e.g. maximum daily temperature) is a good proxy for others (e.g. minimum daily temperature, or wet bulb temperature). Figure 1 demonstrates this, showing the correlations between daily minimum and daily maximum temperatures in a handful of major cities – the correlations are high in cities with temperate climates.

However, in tropical and subtropical climates – home to around 40% of the world's population (State of the Tropics 2020) – this is not the case. For instance, in Lagos, Mexico City, Mumbai, Jakarta, and Darwin (the capital city of the Northern Territory), Figure 1 shows that there are relatively low correlations between maximum and minimum daily temperature. The same is true for the correlation between air temperature and wet bulb temperature (see Appendix A). In tropical climates, one measure is not a good proxy for all others, and therefore we must carefully consider which aspect of heat exposure we want to measure.

This chapter makes two main contributions to the knowledge base on the link between prenatal heat exposure and birth outcomes.

[22] 43% of the global population live in tropical and sub-tropical zones (State of the Tropics 2020) and 35% in deserts and drylands (including hyper-arid, arid, semi-arid and subhumid climates) (United Nations Environment Management Group 2011, 27). Summed together, this would make 78%, however there is some overlap between these definitions, meaning the total is likely closer to 50%. In contrast, in Chersich et al.'s (2020, 1a (supplementary)) systematic review, they find only 15 of 47 studies that analyse the association between heat exposure and preterm birth come from subtropical, arid or semi-arid climates. Of these, just one study relates to a hot arid climate – predominant in much of Africa, the Middle East and central Asia. They find no studies from equatorial/tropical climate zones, such as those covering northern Australia, the Indian Subcontinent, south-east Asia, central and south America, and much of central Africa (Kottek et al. 2006).

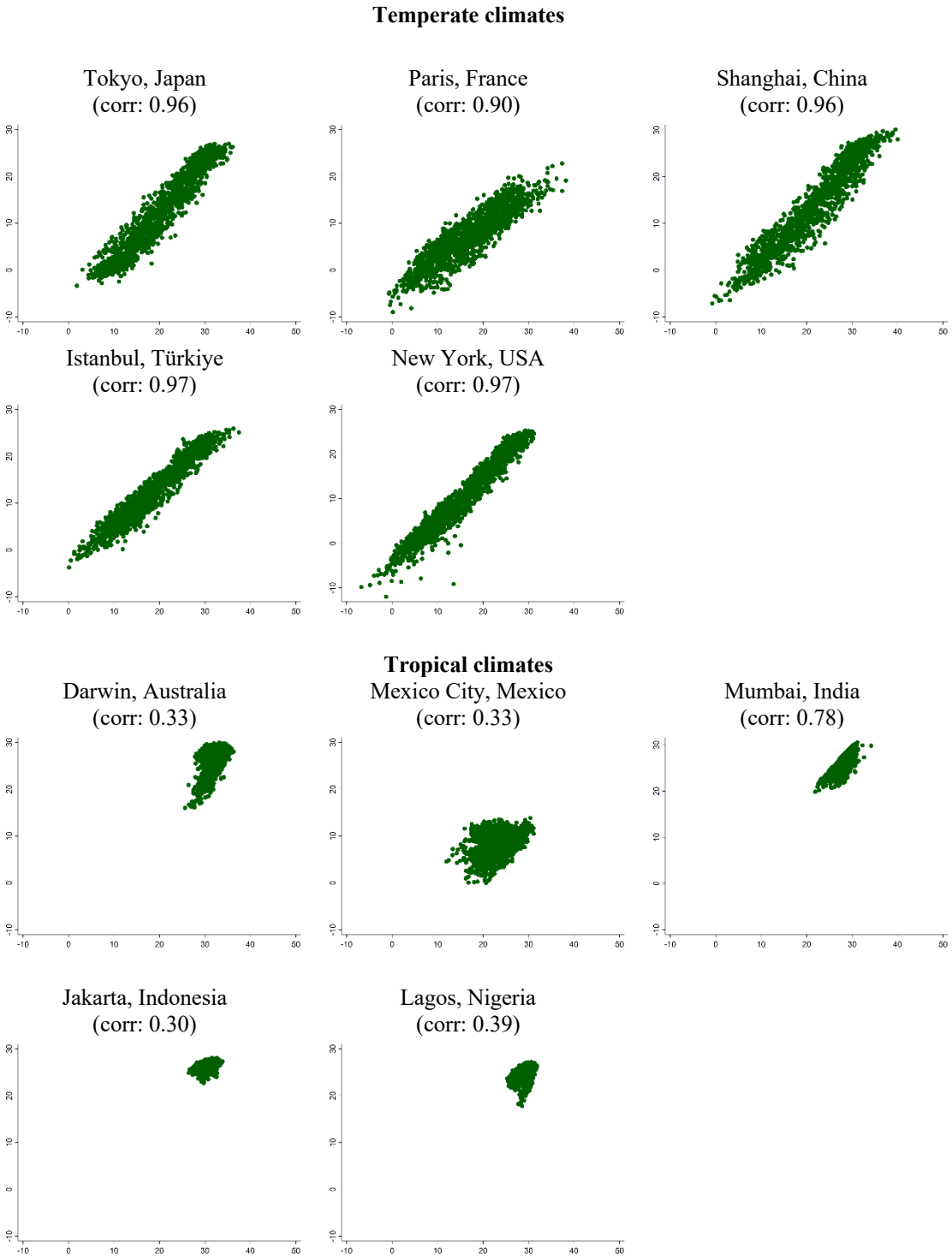


Figure 1: scatter plot of daily maximum and minimum temperature, 2020-2023
Source: NASA reanalysis data

First, we quantify how much the way we measure heat exposure matters. Given the large body of research that finds an impact of heat exposure on birth outcomes, we are confident that this causal relationship exists. But every recent review on this topic has highlighted

a lack of consistency in choice of heat metric (Yunquan Zhang, Yu, and Wang 2017; Wondmagegn et al. 2019; Nawaro et al. 2023; Baharav et al. 2023; Nyadanu et al. 2024; Chersich et al. 2020). The problem with this inconsistency is that when different metrics are applied in different populations, we do not know whether it is the population, the metric, or something else entirely, that explains differences in findings. Our analysis compares alternative heat metrics within a single population. We find that both the existence and size of the relationship between heat exposure and birth outcomes can vary dramatically depending on the metric chosen. In the Northern Territory, we find that a metric based on average maximum daily temperatures alone – as is common in the literature – captures less than half of the impact that our preferred metric estimates.

Second, we analyse which measurement choices matter most. Recent research has questioned the use of air temperature as a default in this literature (Brimicombe et al. 2024; Baldwin et al. 2023; Leung 2023). Their reasoning is that metrics which measure both air temperature and humidity together (e.g. wet bulb temperatures and heat indices), will more accurately reflect people's experiences of heat than air temperature alone. However, we find that wet bulb temperatures have limited explanatory power in the population we study, and instead, other metrics – e.g. including both maximum and minimum daily temperatures – provider better explanatory power. Unlike wet bulb and heat indices, these measures are also more readily available, meaning there is little barrier to their use.

This chapter also provides new evidence on how heat exposure affects birth outcomes in the Northern Territory of Australia. We estimate that typical seasonal variation in heat exposure contributes to a 4.5 percentage point higher risk of preterm birth at some times of year. This is large – for example, close in magnitude to the risks of frequent smoking during pregnancy (Liu et al. 2020).

In our analysis we focus on preterm birth, as this is the outcome most commonly studied in previous research. Preterm birth is associated with poorer health and developmental outcomes later in childhood (Kelly and Griffith 2020) and leads directly to an increase in healthcare costs by triggering admission to neonatal intensive care units (Jacob et al. 2017). However, as we will discuss in our conceptual framework, heat exposure may affect fetal development in a range of ways, most of which do not necessarily lead to

preterm birth. It is therefore important to note that we find impacts of heat exposure that are consistent across five different measures of health at birth; this confirms that heat exposure affects fetal development and health at birth in general, the effect is not isolated to a single outcome.

The rest of this chapter proceeds as follows. The next section sets out a conceptual framework, outlining the possible causal pathways that may explain the impacts of heat exposure on prenatal development, and how they may be measured. The Data section provides context on the Northern Territory and the administrative data we use. The Methods section outlines the heat metrics estimation method we use. We then turn to our data analysis, in which we identify our preferred heat metric and discuss conclusions we may draw if we instead used alternative heat metrics. In the Discussion, we reflect on the practical implications of using non-preferred heat metrics, and recommendations for future analysis.

3.2 Conceptual framework

In order to select the most appropriate heat exposure metric, we must understand which aspects of heat exposure matter for prenatal development. The difficulty here is that there are multiple mechanisms through which heat exposure during pregnancy can affect the mother and developing fetus (Brink et al. 2024), and we do not yet know which of these mechanisms is most important (Samuels et al. 2022). In this section, we explain these mechanisms (summarised in Figure 2) and discuss how this informs our selection of the heat exposure metrics which we test in our analysis.

3.2.1 What is heat exposure?

Heat exposure it often discussed interchangeably with air temperature – that is, the temperature that can be measured with a standard thermometer. However, while air temperature is a major contributor to a person's experience of heat, there are additional contributors. As McGregor and Vanos (2018) explain in their primer on the physiological impacts of heat on the human body, these include other weather conditions (humidity,

windspeed, radiation), individual factors (levels of exertion, pre-existing medical conditions, and medications), and the built environment.

Many researchers argue that humidity is particularly important (Brimicombe et al. 2024; Leung 2023; Davis, McGregor, and Enfield 2016). This is because one way that the body cools itself is through having sweat evaporated from the skin, and sweat evaporates more slowly when it is humid. Therefore, for a given temperature, higher humidity increases the body's risk of overheating.

Here we discuss the effects of heat in general on the fetus, before returning to the question of which measures may best capture these effects.

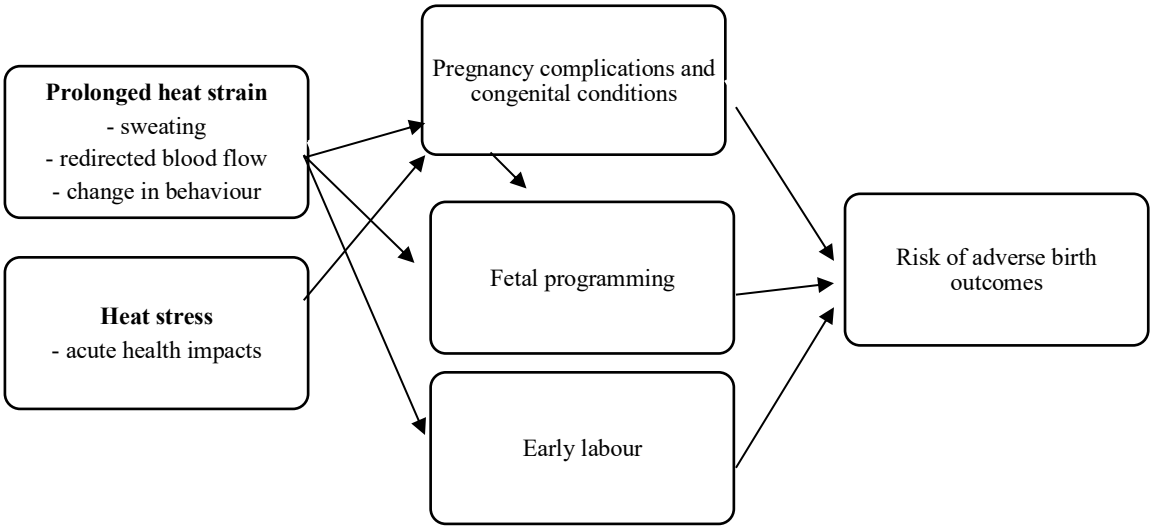


Figure 2: Summary of mechanisms through which heat stress and heat strain may affect birth outcomes
Source: Authors' analysis

3.2.2 How does heat exposure affect the fetus?

We set out three channels through which heat exposure may affect prenatal development: heat stress, heat strain and maternal behaviour.

It is important to note that while these mechanisms can increase the risk of adverse birth outcomes like preterm birth – which is the focus of this study – they will not necessarily do so. They may affect fetal development in ways that impact a child's health and development later in life, but which do not necessarily lead to preterm birth.

Heat stress

Research on the impact of heat exposure on health tends to focus on heat stress; that is, the acute health conditions that can result from an increase in the core body temperature (McGregor and Vanos 2018). These effects include heat stroke, cardiovascular stress, respiratory stress, and acute kidney failure (Ebi et al. 2021; Clark et al. 2024). Heat stress typically occurs at air temperatures around or above the average body temperature (37 degrees C). However, it can occur at lower air temperatures (Ebi et al. 2021), for instance due to high humidity or high levels of exertion.

We know from both animal studies and epidemiological evidence that maternal heat stress during pregnancy can cause birth defects (Auger et al. 2017; Samuels et al. 2022). This may happen either through the fetus itself overheating, or indirectly through the heat stress-induced condition that the mother is facing. Some suggestive evidence also indicates that heat stress early in pregnancy may affect the development of the placenta, increasing the likelihood of pre-eclampsia (Part et al. 2022), and neural tube formation, leading to neural tube defects (Bruckner et al. 2024; LaPointe et al. 2024).

Heat strain

Even when conditions are not hot enough to cause heat stress, heat may still impact fetal development because it causes 'heat strain' for the mother (McGregor and Vanos 2018). Heat strain involves the normal workings of the thermoregulatory system to cool the body, and hence avoid heat stress: that is, by sweating (as described above) and by redirecting blood flow from the core of the body to the skin to dissipate heat (Tansey and Johnson 2015).

If heat strain is experienced for a prolonged period, it may affect the fetus. This can happen because of sweat-induced dehydration. It can also happen because when the mother's blood flow is redirected to her skin, this means reduced blood flow to the placenta, and hence reduced flow of nutrients to the fetus (Samuels et al. 2022). Both

channels have been evidenced in animal studies, showing long-term 'fetal programming' effects on offspring health (Ross and Desai 2005; Johnson et al. 2020). In addition, there is strong epidemiological evidence that exposure to moderate heat, if experienced towards the end of pregnancy, can bring forward labour (Barreca and Schaller 2019).

Maternal behaviour

Beyond physiological effects, prolonged heat exposure can also affect maternal behaviour, with flow-on effects to the fetus. Heat exposure reduces metabolism and appetite (Tansey and Johnson 2015), which may mean changed food and drink consumption patterns (Charlot et al. 2022). Heat may also reduce quality and length of sleep (Zhang X. et al. 2022) and lead to changes in exercise patterns (Samuels et al. 2022).

3.2.3 Implications for heat exposure measurement

Given the mechanisms outlined above, we draw three takeaways for how we might construct heat exposure metrics.

First, that it is not obvious which underlying measure of heat we should use. Within the literature, heat exposure is almost universally measured based on air temperature. But as described above, a person's experience of heat is a result of a range of factors, including temperature, humidity, wind and radiation. Therefore, in theory, a data series incorporating these variables may provide a more accurate picture of true heat exposure. One option is wet bulb temperature, which can be observed directly, and reflects air temperature and humidity.[23] However, the reason that this is not an obvious choice is that, empirically, we do not see a consistent link between high humidity and poorer health outcomes (Baldwin et al. 2023).[24]

Second, that the effects of extreme heat – that is, heat at levels around or above normal core body temperatures – may be different from the effects of moderate heat, because extreme heat is more likely to cause heat stress.

[23] A range of additional indices are also available, such as Apparent Temperature, Heat Index, Universal Thermal Climate Index, and Wet Bulb Global Temperature (Brimicombe et al. 2024).
[24] As Baldwin et al. suggest, there may be a range of explanations for this, including the fact that a build up of humidity when it is not raining has different effects on heat than high humidity while raining.

Third, that even moderate levels of heat, when experienced for a prolonged period, may affect fetal development by causing heat strain or changes in maternal behaviour. If this is the case, then it may be important to choose metrics that reflect prolonged heat exposure – such as high daily minimums – in addition to the metrics used in the overwhelming majority of studies which measure peak levels of heat exposure (i.e. high daily maximums). This may be particularly important in tropical climates where daily maximums and minimums are not highly correlated with one another (as shown in Figure 1).

In our analysis, we estimate the impact of heat exposure on preterm birth, using a range of different heat metrics, to test which fits the data best. In line with this conceptual framework, we test metrics that include both air temperature and wet bulb temperature, metrics based on daily maximums and minimums, and metrics allowing for non-linear effects of heat exposure. We compare these metrics with those most commonly used in prior research (described below).

3.3 Study context and data

3.3.1 Study context

The Northern Territory (NT) is one of Australia's eight states and territories. The NT has a population of around 233,000, 60 percent of whom live in or around the capital city of Darwin (Australian Bureau of Statistics 2022). The NT is a large region covering the central part of northern Australia, in which residents face high levels of heat exposure. Previous research shows that heat exposure plays a major role in explaining month-to-month variation in average birth outcomes in the NT (Doyle 2023).

The tropical north of the NT, where Darwin is located, is hot and humid. Temperatures vary within season but, on an average day in the wet season (November to April), temperatures range between 25 to 33 degrees Celsius, and in the dry season (May to October), between 20 to 30 degrees. The hottest time of year is October-December, in the 'build up' to the wet season – when temperatures and humidity are high, and there is little

rainfall. Heavy rainfall usually begins in late December, though the timing can vary from year to year – this is reflected in the weather data we use in our analysis.

The central and southern parts of the NT have an arid climate, with very hot summers and mild or cold winters. In Alice Springs, the largest town in the region, temperatures on an average summer day range between 20 and 35 degrees, and on an average winter day, between 4 and 20 degrees.

Around one-third of the NT population identify as Aboriginal. There are significant differences between the Aboriginal and non-Aboriginal populations in the NT, in terms of geography, heat exposure and economic resources. Eighty percent of Aboriginal residents in the NT live outside of Darwin, many in remote Aboriginal communities which experience more extreme weather conditions. In remote communities, many houses are poorly insulated, and many residents face energy poverty (Quilty et al. 2022).

3.3.2 Data

We combine administrative birth records with daily weather data.

Our analysis sample includes all babies who were born in the NT and conceived[25] between March 2000 and September 2009, a sample of 34,258 children. We include only births to mothers whose usual place of residence is in the NT, and could be geo-coded.[26] We include stillbirths (making up under 1% of births) and plural births (under 2.5% of births) in the analysis sample.

We link these records to NASA's daily weather reanalysis data, which is measured at intervals of 0.5 x 0.625 degrees of latitude and longitude (roughly 50x55km). This gives a total of just over 500 cells throughout the NT, 213 of which are inhabited.

[25] We define our sample based on date of conception instead of date of birth because definitions based on date of birth will systematically exclude children born preterm at the beginning of the sample, and exclude those born late term at the end of the sample period – this could lead to bias, especially when analysing the impact of seasonal exposures like heat (Strand, Barnett, and Tong 2011a). We determine date of conception by subtracting gestational age (in weeks) from birthdate.

[26] There are a small number of births for which the mother's place of residence as entered in the perinatal data could not be found, either using a fuzzy match with the R package 'geonames', or through manual search on Google, the NT Place Names Register (https://www.ntlis.nt.gov.au/placenames/) and BushTel (https://bushtel.nt.gov.au/).

Reanalysis data comes from NASA's model using ground station and satellite data. Some may be concerned about accuracy of these data compared with traditional weather station observations. However, in our context, the NASA data lines up very closely with observations from the Australian Bureau of Meteorology's weather stations. We use the reanalysis data because they are highly localised and have no missing observations – which is not the case for weather stations across remote parts of the NT (see Appendix B for details).[27]

3.4 Methods

3.4.1 Heat exposure metrics

We construct five different heat exposure metrics. The first three are based on common metrics used in prior research, while the next two build on the conceptual framework outlined above. Our metrics are:

1. Benchmark

This is a count of the number of days in each trimester with maximum daily temperatures within 5-degree ranges (e.g. under 20, 20-24.99, 25-29.99, 30-34.99, 35-39.99 and 40+). The omitted category is 25-29.99 degrees. This metric imposes a general functional form, and it is the approach that Dell et al. (2014) recommend in cases when researchers are agnostic about how the heat metric should be specified. It has been widely used e.g., by Barreca and Schaller (2019), Kim et al. (2021), Chen et al. (2020), He et al. (2018), Deschênes, Greenstone and Guryan (2009), Isen et al. (2017), Hu and Li (2019).

We take this as our 'benchmark' model, because, out of the metrics that are most commonly used in the literature, this best allows for the non-linearities that we expect to see, based on our conceptual framework.

2. Average temperatures

[27] In contrast, over this period some weather stations are decommissioned, and others have systematic missing observations (i.e. in the wet season when they become inaccessible due to flooding and therefore observations are systematically missing during periods of extreme weather).

This is a simple average of maximum daily temperatures in each trimester in pregnancy. This metric is used, for example, in Wilde et al. (2017), Strand et al. (2012), Sun et al. (2019), Wolf and Armstrong (2012), Randell et al. (2020) and Thiede and Gray (2020). It assumes linear effects of each additional degree, and hence does not allow for different effects for higher temperatures most likely to cause heat stress.

3. Heatwave count

This is a count of the number of heatwaves in each trimester of pregnancy. We use the Australian Bureau of Meteorology's heatwave definition.[28] Counts of extreme heat events (whether defined as heatwaves or not) are common in the literature (Dadvand et al. 2011; Andalón et al. 2016; Bakhtsiyarava, Grace, and Nawrotzki 2018; Grace et al. 2015; Mathew et al. 2017). Such a metric allows for nonlinearities in the effects of heat exposure, but imposes a specific structure on that nonlinearity; it assumes a threshold over which effects occur, and that there is no impact of exposure to more moderate heat conditions below that threshold.

4. Max and min

This is an enhanced version of the benchmark metric, where, in addition to maximum temperatures, we include counts of the number of days in each trimester with minimum air temperatures within 5-degree ranges (under 5, 5-9.99, 10-14.99, 15-19.99, 20-24.99, 25+). The omitted category for minimum temperatures is 15-19.99. This metric reflects the implication from our Conceptual framework that daily minimum temperatures – which, if high, reflect prolonged exposure to heat – may indicate higher risk of heat strain.

5. Wet bulb

This metric is analogous to the 'Max and min' metric described above, but using wet bulb temperatures as the underlying data series, instead of air temperature. We construct counts of the number of days within each trimester with daily average wet bulb temperatures

[28] We calculate this based on the Excess Heat Factor, which measures the extent to which daily air temperatures are unusually high for a given location and time of year (Nairn and Fawcett 2015). It is calculated by combining a) average daily temperatures in the previous 3 days, compared with the past 30 days, and b) average daily temperatures in the previous 3 days compared with the long-term location average. Full details of how this is constructed are available in Nairn and Fawcett (2015). We use this measure as it is the one used by the Australian Bureau of Meteorology, who define heatwaves as three or more consecutive days with a positive Excess Heat Factor.

with ranges of <10, 10-14.99, 15-19.99, 20-24.99, 25+, and the number of days with daily maximum wet bulb temperatures within ranges of <10, 10-14.99, 15-19.99, 20-24.99, 25-29.99 and 30+.[29] This metric reflets the possibility highlighted in our Conceptual framework that wet bulb temperatures may measure heat exposure more directly than air temperatures.

Table 1 sets out the mean, standard deviation, maximum and minimum of these measures over the full pregnancy, for babies in our sample. During an average pregnancy, there are 2.4 heatwaves and just over one week with maximum temperatures above 40 degrees – but this varies greatly, with some pregnancies experiencing up to 9 heatwaves, and 98 days with maximum temperatures above 40 degrees.

3.4.2 Outcome measure

The birth outcome measure we focus on in our analysis is preterm birth, which is defined as birth before 37 complete weeks of pregnancy. Around 10 percent of babies are born preterm (Table 1). We focus on preterm birth for comparability with previous research: it is the most commonly studied birth outcome. However, our conclusions hold across four additional measures of health at birth: birthweight, small for gestational age, Apgar scores,[30] and admission to a special care nursery. Our main regression estimates using these additional outcomes are presented in Appendix D.

Preterm birth is an important intermediary outcome: a large body of research tells us that children born preterm are more likely to face worse health and educational outcomes (Moreira, Magalhães, and Alves 2014; Kelly and Griffith 2020; Allotey et al. 2018). In the NT specifically, children who were born preterm are more likely to be assessed as developmentally vulnerable at age 5 (Guthridge et al. 2016).

[29] Maximum wet bulb temperature is not available within the data, but we construct this measure using daily maximum combined with relative humidity, using Stull's (2011) equation. We use average wet bulb temperature instead of minimum because this measure is directly available. However, as daily average is a linear transformation between maximum and minimum, the relationship between maximum and minimum (for air temperature), and maximum and average (for wet bulb temperature) should be constant. Appendix Tables D6 and D7 show models using average wet bulb and maximum wet bulb separately.

[30] Birthweight is measured in grams. SGA is a binary indicator, which equals 1 if birthweight is below the 10th percentile for a given gestational age. I define these relative to Australian national percentiles, separately for male and female infants (Dobbins et al. 2012). Apgar 5 is an index with values from 0 to 10, based on the birth attendant's judgement of the newborn's skin colour, heart rate, reflex, muscle tone and respiratory effort, 5 minutes after birth.

	Mean	Standard deviation	Minimum	Maximum
Birth outcomes				
Preterm birth (probability)	0.10	0.31	0.00	1.00
Birth weight (grams)[a]	3282.29	532.71	2160.00	4180.00
SGA (probability)	0.14	0.35	0.00	1.00
Apgar 5 score	8.89	1.22	0.00	10.00
Special care nursery (probability)	0.18	0.39	0.00	1.00
Heat exposure in 39 weeks to birth				
Maximum temp (average)	31.23	1.74	24.60	37.84
Minimum temp (average)	22.43	4.21	10.33	28.03
Daily average wet bulb (average)	21.67	4.57	9.47	26.41
Maximum wet bulb (average)	24.43	3.84	11.72	30.00
Number of heatwaves	2.36	1.80	0.00	9.00
Excess Heat Factor (average)	-9.31	10.45	-46.29	-0.60
Daily maximum temperature bands				
Days max temp <15	0.41	1.68	0.00	21.00
Days max temp 15-20	4.53	11.18	0.00	64.00
Days max temp 20-25	10.17	18.45	0.00	83.00
Days max temp 25-30	85.75	47.81	0.00	266.00
Days max temp 30-35	129.41	51.03	6.00	253.00
Days max temp 35-40	34.78	36.90	0.00	161.00
Days max temp 40+	7.96	14.02	0.00	98.00
Daily minimum temperature bands				
Days min temp < 5	5.19	12.74	0.00	59.00
Days min temp 5-10	10.83	21.12	0.00	90.00
Days min temp 10-15	16.30	23.12	0.00	87.00
Days min temp 15-20	34.88	28.83	0.00	128.00
Days min temp 20-25	87.65	33.28	3.00	225.00
Days min temp 25+	119.15	76.18	0.00	271.00
Daily average wet bulb temperature bands				
Days avg wet bulb <10	21.14	40.33	0.00	168.00
Days avg wet bulb 10-15	19.42	27.37	0.00	122.00
Days avg wet bulb 15-20	36.07	23.44	0.00	121.00
Days avg wet bulb 20-25	81.62	32.12	2.00	204.00
Days avg wet bulb 25+	114.75	70.02	0.00	247.00
Daily maximum wet bulb temperature bands				
Days max wet bulb <10	0.78	2.57	0.00	17.00
Days max wet bulb 10-15	13.90	28.27	0.00	114.00
Days max wet bulb 15-20	25.86	35.25	0.00	142.00
Days max wet bulb 20-25	65.24	28.50	0.00	167.00
Days max wet bulb 25-30	163.61	73.37	0.00	272.00
Days max wet bulb 30+	3.61	8.29	0.00	53.00

Table 1: Descriptive statistics

Source: NASA and Analysis dataset, average birth outcomes and heat exposure for all babies conceived in the NT from March 2000 to September 2009.

a: variable is top- and bottom-coded at the 2.5th and 97.5th percentiles, to reduce the influence of extreme outliers.

Recent research demonstrates that this relationship is not fixed: in advanced economies, the predictive power of preterm birth has declined over time (Goisis, Özcan, and Myrskylä 2017), and many children born preterm may face no detectable long-term effects (Baranowska-Rataj et al. 2023). This makes sense, for two reasons. First, advances in neonatal healthcare have greatly improved outcomes for preterm-born infants (Crump 2020). Second, because preterm birth has many causes (Romero, Dey, and Fisher 2014). The aetiology of preterm birth, and hence the long-term outcomes associated with it, may differ over time and across contexts.

However, even in the desirable situation where preterm birth does not increase the risk of long-term developmental vulnerabilities for an individual, it remains an outcome of interest given the high cost of providing remedial neonatal care (Jacob et al. 2017).

3.4.3 Analytical methods

Our goal in this analysis is to assess whether and how our estimates of the causal impact of heat exposure on preterm birth change when we use different heat metrics. A challenge in doing so is that heat exposure may be correlated with birth outcomes for reasons which do not necessarily reflect the causal impact of heat itself.

Of primary concern is omitted variable bias. Heat exposure varies both over time, and across locations. But most of the variation in birth outcomes over time and across locations is not due to heat exposure. For instance, socioeconomic status is known to affect birth outcomes, but is also likely to affect both place of residence and timing of conception (Currie and Schwandt 2013): this link with place of residence may be particularly important in the NT, given that most non-Aboriginal people live in the capital city of Darwin, whereas most Aboriginal peoples live in remote communities, many of which face higher levels of heat exposure than Darwin (Australian Bureau of Statistics 2022b). Similarly, we know that there are seasonal risks in addition to heat exposure, such as disease prevalence and economic conditions, which contribute to month-to-month variations in birth outcomes (Doyle 2023).

Our analytical approach is to estimate the causal effect of heat exposure by isolating fluctuations in heat that are exogenous to these (possibly endogenous) sources of time- and place-based variation. Exposure to such exogenous variation is beyond individuals'

control, and cannot be anticipated more than a few days in advance. This variation in exposure is, therefore, as good as random.

To do this, we specify a regression model with fixed effects for time and place. This approach is discussed in detail in Dell et al.'s (2014) review, and has been widely used in prior research.

As is common practice, we allow for the effects of heat exposure to vary by trimester; this allows for the possibility, as discussed in the Conceptual framework, that the effects of heat exposure may depend on its timing – for instance, exposure in the first trimester may increase the risk of pre-eclampsia (Part et al. 2022), and exposure in the third trimester may bring forward labour (Barreca and Schaller 2019).

We therefore estimate the following linear fixed effects model:

$$preterm_{itj} = b_0 + \sum_{k}^{3} b_1^k heat_{tj}^k + b_2 X_i + \gamma_{my} + \theta_{jmi} + \epsilon_{ij} \qquad (1)$$

where $preterm_{itj}$ is an indicator of whether baby *i*, conceived on date *t* (in month *m,* year *y*), in location *j*, was born preterm or not. *k* is an index of the three trimesters of pregnancy. *X* is a set of individual-level covariates, which includes an indicator for whether it is the mother's first pregnancy, the age of the mother (in 5-year age bands), and an indicator for whether the baby is Aboriginal. Standard errors are clustered at the location level.[31]

Heat represents any one of the five heat metrics outlined above. They are based on daily weather observations from small-area location *j*, which are summarised over each trimester based on exact date of conception and date of birth *t*. Our coefficients of interest are b_1^k.

In specifying the fixed effects, we follow prior studies in using fixed effects for the month-year of conception, γ_{my}, and for the mother's place of residence, interacted with

[31] We identify location based on the mother's suburb/community of residence, which we have geocoded, and grouped into locations of maximum 50-mile distance between points using cluster analysis.

the month of conception and the baby's sex, θ_{jmi} (Cil and Kim 2022; Isen, Rossin-Slater, and Walker 2017). Interacting location fixed effects with month and sex allows for the possibility that there may be different seasonal patterns in birth outcome across locations – this may be the case in the NT given, for example, that some communities experience regular flooding in the wet season, and hence their experience of the wet season may be different from other communities that do not flood. Furthermore, we know that different types of exposures may affect male and female fetuses differently (DiPietro and Voegtline 2017). Interacting these location- and season-specific effects with sex allows for this possibility. That said, Appendix C shows that our findings are similar regardless of the specifics of how we define these fixed effects.

3.4.4 Our approach to selecting a preferred heat metric

Part of our goal in this analysis is descriptive: to run the same model sequentially with each heat metric, and learn how choice of heat metric affects our conclusions around the impacts of heat exposure on preterm birth. To do this, we present regression coefficients for the same model, using each of our five heat metrics.

However, we also wish to select a preferred heat metric. Out of our five candidate metrics, we want to know which one best captures the impact of heat exposure on preterm birth in the NT. In doing so, we face a challenge: because each regression uses different heat exposure metrics, the magnitudes and statistical significance of the resulting coefficients cannot be directly compared with each other.

We therefore take a two-stage approach to selecting a preferred metric. First, we present an F-test of joint significance of the regressors included in each heat exposure metric. If the regressors are not jointly statistically significant, we exclude the metric from further comparison. Second, we compare goodness of fit, for the remaining metrics.

We use two measures of goodness of fit. First, the R-squared. This tells us how much variation in the outcome is explained by the regressors. A standard R-squared increases when additional regressors are added. To account for different numbers of regressors in each metric, we use an adjusted R-squared: this penalises additional regressors. A higher adjusted R-squared, therefore, indicates that the model fits the data better, when comparing among models with different numbers of regressors. Second, we present the

Akaike Information Criterion (AIC). The AIC helps us to compare several alternative, non-nested models, on a single measure: it is calculated as the likelihood of the model (estimated using maximum likelihood), penalised for the number of regressors. It therefore weighs up model fit with model complexity, allowing us to compare multiple working hypotheses against each other (Portet 2020). In their review on variable selection, Heinze, Wallisch and Dunkler (2018) recommend using AIC in cases such as ours, where theory supports a relatively small set of competing models, and we want to select between them. In comparing AIC scores, a lower score indicates better fit.

3.5 Estimation results

3.5.1 Goodness of fit comparison

Table 2 presents a comparison of how well each heat metric fits the data.

Across our five metrics, we can first exclude 'Trimester average' and 'Heatwave count' from comparison: in an F-test of joint significance at the 5% level of significance, we would fail to reject the hypothesis that the coefficients in these metrics are jointly equal to zero, hence that they do not help to explain variation in preterm birth rates.

	P-value: F test for joint significance of heat exposure terms	Adj R^2	AIC
Benchmark	0.00	0.036	13214
Trimester average	0.09	0.035	13251
Heatwave count	0.12	0.035	13267
Max and min	0.00	0.037	13174
Wet bulb	0.00	0.035	13233

Table 2: Measures of model fit for each heat metric

Note: Each row of this table represents results from a regression using the indicated heat metric. The first column in this table shows p-values for an F-test of joint statistical significance of all terms in the heat exposure metric. The second column, Adj R^2, represents the adjusted R-squared, for which a higher value represents better fit. Note that here adjusted R-squared is 'within' the community × sex × month fixed effects (which are absorbed using 'areg' in Stata). The Third column, AIC, is the Akaike Information Criteria, for which lower value indicates better fit. Sample size: 34,178.

We then have three remaining metrics. Across all metrics, the adjusted R-squared is low and varies little, ranging between 0.035 to 0.037. However, the metrics produce very different AIC scores. On both R-squared and AIC measures, the 'Max and min' metric performs better: it has a slightly higher adjusted R-squared and substantially lower AIC.

We separately consider whether a combination of metrics provides a better fit – for instance, a combination of the 'Max and min' metric, the 'Wet bulb' metric and 'Heatwave count' metric together. However, while some of the additional coefficients are statistically significant, they do not meaningfully improve model fit (see Appendix Table D8), and interpretation of individual coefficients becomes very difficult.

We therefore select the 'Max and min' metric as our preferred heat exposure metric.

3.5.2 Coefficient estimates

Having identified the 'Max and min' metric as our preferred one, we now turn to a presentation of the coefficients for the regressors in each heat metric. This tells us both the estimates from our preferred model, and the potential consequences, in terms of inferences we may draw, from selecting a non-preferred metric.

We present estimates graphically, showing coefficients and their 95% confidence intervals. Tables with these coefficients and their standard errors can be found in Appendix D.

Benchmark heat metric

Figure 4 presents estimates from our benchmark metric. They suggest that heat exposure in all trimesters affects the risk of preterm birth, but in different ways.

We can see that exposure to high temperatures between 30 and 40 degrees in the first trimester slightly increases the risk of preterm birth. However, in the second and third trimesters, it is exposure to cooler temperatures (below our omitted category of 25-30 degrees) that contribute to lower risk of preterm birth; we see no additional impact of heat exposure above 30 degrees. This suggests that it may be predominantly heat strain (i.e. prolonged exposure to moderate heat), instead of heat stress (i.e. exposure to extreme heat) affecting pregnancies in the second and third trimesters.

If it is the case that maternal heat strain affects the baby through reduced flow of nutrients to the placenta (Samuels et al. 2022), this could explain why we see positive effects of cooler temperatures in the second and third trimesters, but not the first trimester: in the first trimester, the fetus receives nourishment from the yolk sack and not the placenta, and is therefore less reliant on maternal blood flow for nutrients. When the placenta takes over around the beginning of the 2nd trimester, this is when we might expect an impact of heat strain.

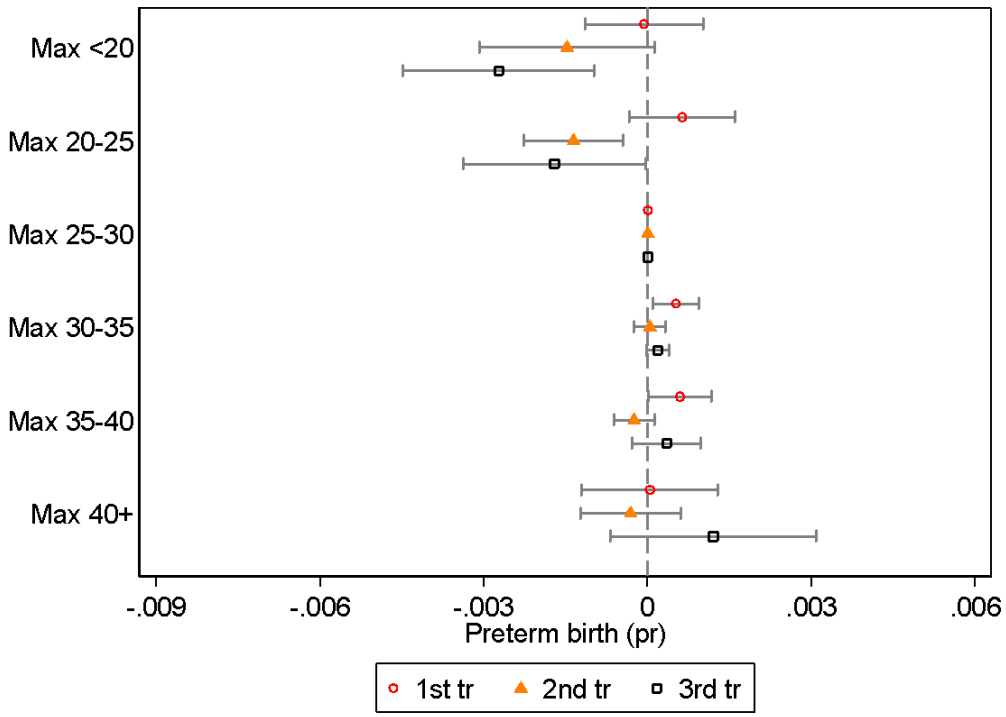


Figure 4: benchmark heat metric
This figure shows the regression coefficients and 95% confidence intervals from equation 1, where the heat metric is a count of the number of days within each trimester that the daily maximum temperature fell into the following ranges: <20, 20–25, 25–30, 30–35, 35–40, 40+. The omitted category is 25–30 degrees. Sample size: 34,178.

Average temperatures

Figure 5 shows estimates from the 'average temperature' metric. Note the difference in scale from Figure 4 due to differences in units: a 1 degree increase in average temperatures represents more heat exposure than a single day of extreme heat. For example, in a trimester with average temperatures of 30 degrees, a 1 degree increase in the average would represent over one week of temperatures of 40 or above.

Figure 5 shows that a 1 degree increase in average temperatures in the third trimester leads to a 0.007 percentage point increase in the probability of preterm birth. Although we saw some statistically significant impacts of heat exposure in the first and second trimesters of the benchmark model, these are not evident here.

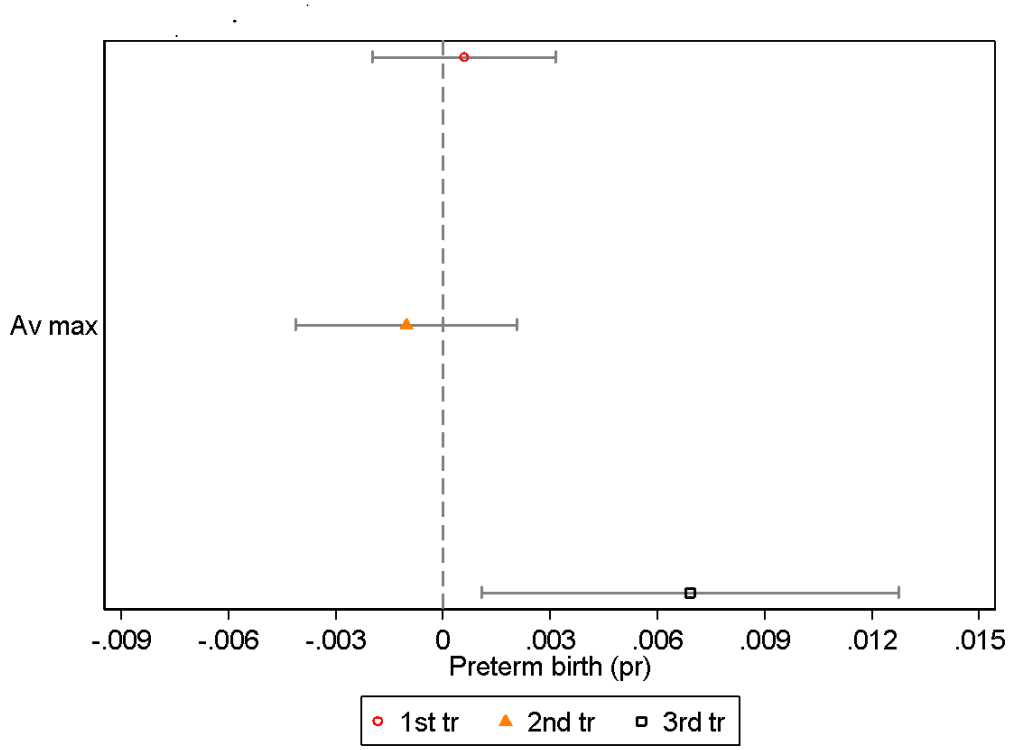


Figure 5: trimester average
This figure shows the regression coefficients and 95% confidence intervals from equation 1, where the heat metric the trimester-average of daily maximum temperatures. Sample size: 34,178.

Heatwave count

Figure 6 shows estimates from our 'heatwave count' metric. We find that while the coefficients are positive (indicating that experiencing an additional heatwave in utero increases the risk of preterm birth), these effects are not statistically significant.[32] This makes sense given that, in our benchmark model, much of the effect came through improved outcomes from exposure to cooler temperatures (below 25-30 degrees), and not from extreme heat conditions that would contribute to a heatwave.

[32] However, there are significant coefficients on average birthweight with each additional heatwave in the second trimester associated to a 5.7 gram reduction in birthweight (Appendix Table D3).

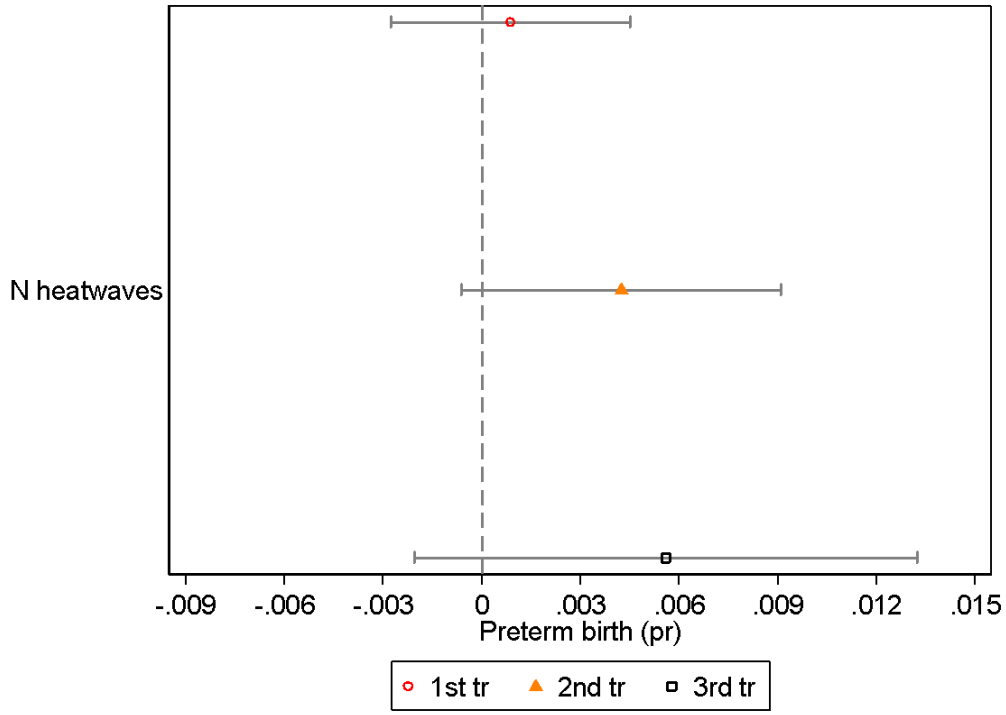


Figure 6: heatwaves
This figure shows the regression coefficients and 95% confidence intervals from equation 1, where the heat metric is a count of the number of heatwaves. Heatwaves are defined as 3 or more consecutive days of unusually hot weather relative to usual conditions within a location – using on the Australian Bureau of Meteorology's method (Nairn and Fawcett 2015). Sample size: 34,178.

Max and min (preferred metric)

The 'Max and min' metric – our preferred metric, based on goodness of fit – adds minimum temperature exposure bands to our benchmark metric. We see similar patterns to the benchmark metric, though the addition of minimum temperatures reveals further effects of prolonged heat exposure – i.e. on days when minimum temperatures do not fall below 20 degrees.

In the first trimester, high minimum temperatures of 20 or above increase the risk of preterm birth, while low minimums (below 10 degrees) reduce the risk; we no longer see a statistically significant impact of high maximum temperatures.

In the second trimester, we continue to see a protective impact of low maximum temperatures (below 25 degrees), and these estimates are larger in magnitude here than

in the benchmark metric. We also find that high minimum temperatures (25 degrees or above) increase the risk of preterm birth.

In the third trimester, we now see that high maximum temperatures (between 30-39.99 degrees) increase the risk of preterm birth, though with no statistically significant impact of minimum temperatures. Some of these changes in estimates may be due to multicollinearity, given that low values of maximum and minimum temperatures each have similar coefficients in the third trimester when modelled separately (Appendix D).

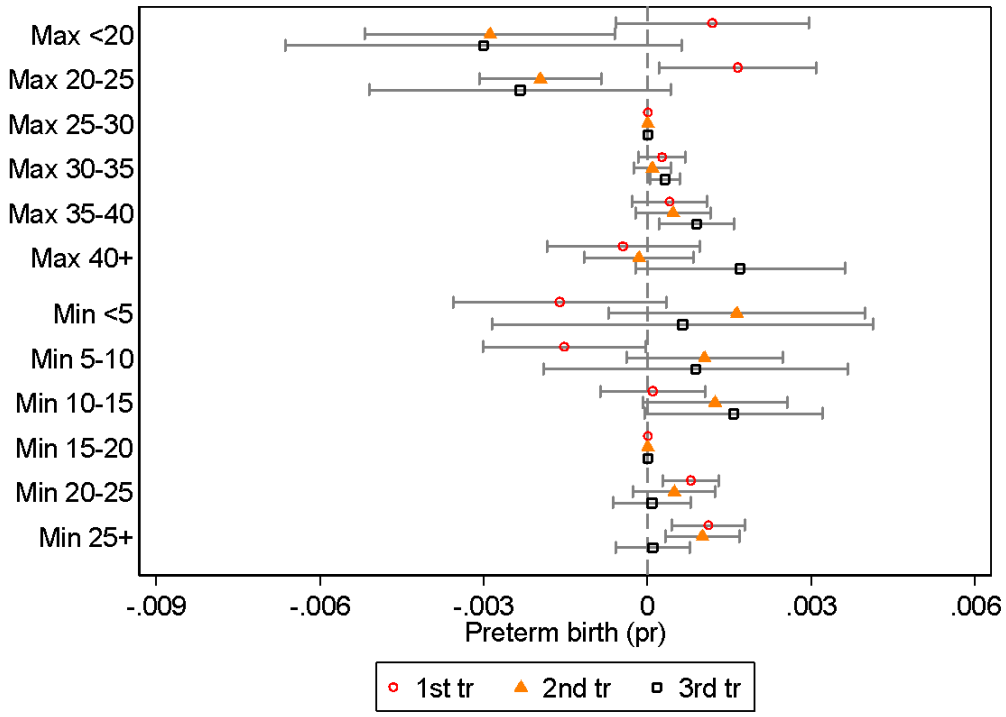


Figure 7: max + min
This figure shows the regression coefficients and 95% confidence intervals from equation 1, where the heat metric is a count of the number of days within each trimester that the daily maximum temperature fell into the following ranges: <20, 20-25, 25-30, 30-35, 35-40, 40+, and a count of the number of days within each trimester that daily minimum temperatures fell into the following ranges:<5, 5-10, 10-15, 15-20, 20-25, 25+. The omitted categories are 25-30 degrees (max) and 15-20 (min). Sample size: 34,178.

As this is our preferred metric, we ran our analysis separately for Aboriginal and non-Aboriginal babies, to assess whether the effects of heat exposure are different amongst these different populations (see Appendix Table E1). We find estimates are similar in both groups, though among Aboriginal children, the only statistically significant estimates are for the third trimester: exposure to cooler maximum temperatures reduces

the risk of preterm birth, and exposure to temperatures of 35-40 degrees increases the risk. In contrast, the effects of higher maximum and minimum temperatures in the first and second trimesters are statistically significant among non-Aboriginal children, but there is little effect of third trimester exposure.

Wet bulb

Figure 8 shows coefficients for our 'Wet bulb' metric – which is analogous to the 'Max and min' metric, but constructed using wet bulb temperatures instead of air temperature. We see little statistically significant relationship between wet bulb temperatures and preterm birth: the only statistically significant coefficient suggests that high wet bulb temperatures in the third trimester (of 25 to 29.99 degrees) decrease the risk of preterm birth – that is, an effect in the opposite direction to that suggested by all other metrics.

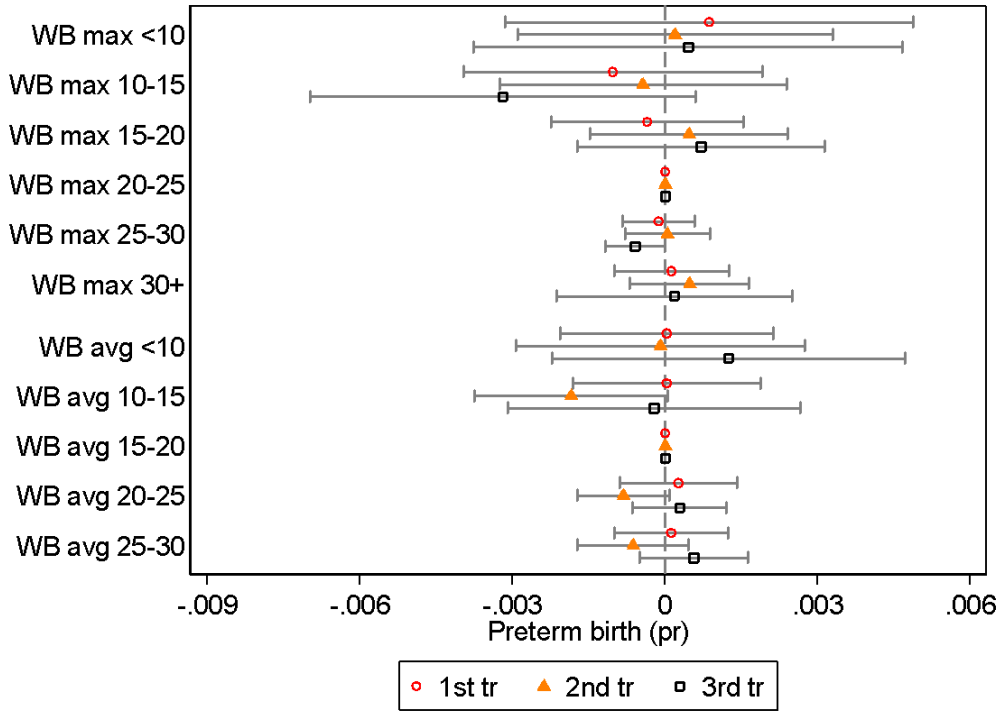


Figure 8: wet bulb

This figure shows the regression coefficients and 95% confidence intervals from equation 1, where the heat metric is a count of the number of days within each trimester that the daily maximum wet bulb temperature fell into the following ranges: <10, 10–15, 15–20, 20–25, 25–30, 30+, and a count of the number of days within each trimester that daily average wet bulb temperature fell into the following ranges: <10, 10–15, 15–20, 20–25, 25+. The omitted categories are 20-25 (max) and 15-20 (avg). Sample size: 34,178.

Figure 8 presents both average and maximum daily wet bulb temperatures in a single metric. But sensitivity testing shows that our conclusions would be similar if we were to construct our metric using maximum or average values separately (see Appendix Tables D6 and D7).[33]

3.5.3 Do these findings differ by climate zone?

As described in the Introduction, the relationship between various measures of heat exposure – and hence the extent to which selection of measures matters – varies across climate zones. We therefore repeat our analysis, but running models separately for the tropical climate zone in the north of the NT, and the arid climate zone in central NT.

We see some differences in coefficients between the climate zones (Appendix Table E2). For instance, while high minimum temperatures in both the first and second trimesters significantly increase the risk of preterm birth in the tropical climate zone, we do not find this in the arid climate zones. In addition, while high maximum temperatures increase the risk of preterm birth in both climate zones, the magnitudes of these coefficients are at least twice as large in arid climate zones.

We may have expected the metric based on wet bulb measures to be a better fit in the tropical climate zone. However, we do not find this. Across both climate zones, the 'Max and min' metric provides the best fit (Appendix Tables E4 and E5).

3.6 Discussion

At a high level, it is clear from Figures 4–8 that choice of heat metric can have a substantial impact on our conclusions around how much heat exposure affects the risk of preterm birth. Our preferred metric, 'max and min', shows that exposure to both extreme heat (defined as additional days with maximum temperatures of 35–39.99 degrees) and prolonged exposure to more moderate heat (defined as additional days with minimum temperatures of 20 degrees or higher) both increase the risk of preterm birth. This general pattern is evident in both our preferred 'max and min' metric, and in the benchmark

[33] We see small statistically significant coefficients on low values of daily maximum wet bulb temperatures in trimesters 1 and 3, but this model has low predictive power.

metric, which is more commonly used in prior research. However, our preferred metric fits the data better. This finding is consistent with other research showing daily minimum temperatures may affect health independently of maximum temperatures (Kim et al. 2023).

However, perhaps more importantly, we find that in the NT, and likely in other contexts as well, the choice of heat metric can determine whether we conclude that prenatal heat exposure affects birth outcomes at all. In particular, we find no statistically significant impact of the number of heatwaves during pregnancy on the risk of preterm birth. In addition, while our preferred metric finds some impact of higher temperatures exposure in all trimesters, these effects are not linear, and hence are not evident in the first and second trimesters in our 'Trimester average' metric, which imposes linearity.

Somewhat surprisingly, we also find that our 'Wet bulb' metric, although it allows for the nonlinearities we find in our preferred metric, does not capture any clear effect of heat exposure. We discuss possible reasons for this below.

Having established a preferred heat metric, we now turn to a discussion of the practical implications of using non-preferred metrics, and our recommendations for metric selection in future research.

3.6.1 Practical implications of using non-preferred heat metrics

The analysis we have presented above demonstrates the statistical significance and model fit of alternative metrics. But the varied scales of the coefficients make it difficult to compare the practical significance of these different estimates. To quantify the implications of metric choice, we use our regression estimates to generate predictions of the impact of normal within-year differences in heat exposure on preterm birth rates.

We generate two sets of predictions: first, the hypothetical case where all babies in our analysis data were exposed to heat typical of the hottest 9 months of the year in utero (born in May), and second, the hypothetical case where all babies were exposed to heat typical of the coolest 9 months of the year (born in November). We then compare the difference between these two predictions.

Figure 9 shows these predictions. In our preferred 'Max and min' heat metric, we estimate that if all babies were in utero during the coolest 9 months of the year, preterm birth rates would be 9 percent, and if all babies were in utero during the hottest 9 months of the year, preterm birth rates would be 13.5 percent. This is a 4.5 percentage point difference.

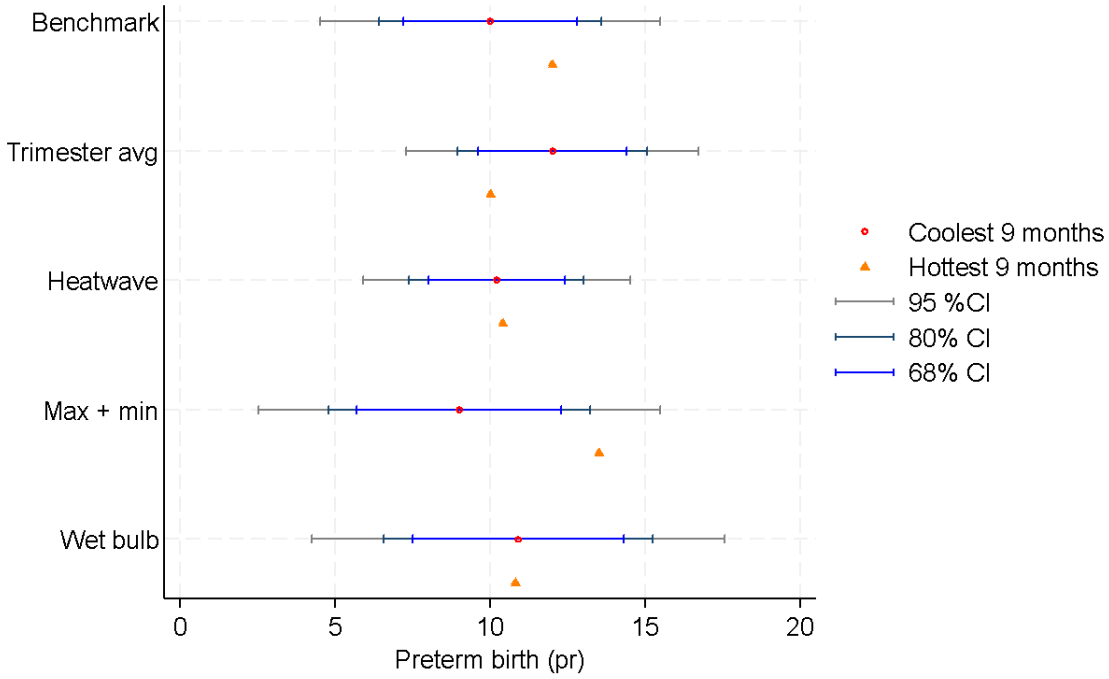


Figure 9: Model predictions with 95%, 80% and 68% confidence intervals.
Predictions shown for all heat metrics, and represent the models' predicted rate of preterm birth for a population exposed to typical conditions in the coolest 9 months of the year in utero, compared with a population exposed typical conditions in the hottest 9 months of the year in utero. The confidence intervals are shown around the estimate for the coolest 9 months for each metric.

In contrast, the alternative metrics imply effects of less than half this size: we estimate a 2 percentage point difference using the benchmark metric, and smaller differences from the other metrics. Therefore, choosing an alternative means not only that we may mis-understand the relationship between heat exposure and preterm birth, but also that we may severely under-estimate heat's implications for population health.

Of course, there is uncertainty around these predictions: the difference between these two predictions with our preferred metric is not statistically significant at the 95% level.

However, based on the large body of research and our own analysis, we know that it is more likely than not that hot weather causes a substantial worsening in health at birth, and a 4.5 percentage point effect is our best estimate of the impact of normal seasonal variation in the NT. To put this in context, an effect of this magnitude is equivalent to intensive rates of smoking during pregnancy (Liu et al. 2020).

3.6.2 Recommendations for future analysis

In our analysis, we have compared five different approaches to defining and measuring heat exposure – we call these 'heat metrics'. We have found that a metric based on air temperature which allows for non-linear effects of increased temperatures, and includes both daily minimum and maximum temperatures, provides the best fit. Our findings lead us to two general recommendations for future research.

First, it is very unlikely that the true relationship between heat exposure and birth outcomes is the one implied by either the 'Trimester average' or the 'Heatwave count' metric. As outlined in our Conceptual framework, we have good reason to believe that both moderate and extreme heat during pregnancy can affect birth outcomes, possibly in different ways. Therefore, at a minimum, our models should allow for these nonlinear effects. While some studies have found that these metrics nevertheless produce statistically significant estimates, we have found that they will not necessarily do so. Furthermore, as shown in Figure 9, even among alternative metrics where the regressors are statistically significant, the implied impact of heat exposure from a more limited metric may differ from a more flexible metric by orders of magnitude.

Second, we find that metrics based on air temperature, and not on wet bulb temperatures, better capture the effects of heat exposure – even in a humid, tropical climate zone. While surprising, our findings are not unique; many epidemiological studies have also found little impact of humidity on health outcomes (Baldwin et al. 2023). In understanding this finding, we cannot rule out issues with data quality and choice of measure. Our wet bulb data are derived from satellite images, and not directly measured at a weather station. This is the only metric incorporating humidity that is available within our region – weather station data do not cover the whole of the NT, meaning we do not have the underlying data series needed to construct alternative heat indices. It may be that, if they

were available, alternative heat indices that combine heat and humidity could have more explanatory power than wet bulb measures.

However, our findings may also reflect real, complex relationships between heat, humidity and health in a tropical climate.[34] It may be that these relationships are not well captured by a single data series, whether it is wet bulb or another heat index. In Darwin, for instance, days with higher humidity and rainfall typically do not reach particularly high maximum temperatures; in fact, humidity and air temperatures are negatively correlated with each other (see Appendix G). But because humidity is a contributor to wet bulb measures, wet bulb temperatures are mechanically higher on cooler, rainy days. Therefore, it is possible that hot, humid days with higher wet bulb temperatures can, in theory, contribute to greater heat strain. But the same wet bulb temperatures can also be the result of very humid, cooler-than-average and rainy days, which provide some relief from high temperatures. This may explain why our estimates suggest that higher wet bulb temperatures in the third trimester can actually reduce preterm birth rates. It may also be that the same level of humidity feels different on a day with intense rainfall (i.e. the wet season) compared with days with no rainfall (i.e. the build up to the wet season). Future research could analyse air temperature, humidity and rainfall, and the interaction between the three as separate regressors, to test whether this yields more informative estimates.

Our takeaway is that in the NT, as well as in many other contexts where the only data available are satellite or reanalysis data, heat exposure metrics based on air temperature may well provide more explanatory power than metrics based on wet bulb temperatures. However, future research could compare a range of heat indices in a location where data are known to be of high quality and where humidity and extreme heat co-occur: this would help to more definitively assess whether the lack of consistent effects of humidity that we and others find is due to data quality, or to a mis-specification of the relationship between heat, humidity and health.

[34] This may explain, for instance, why a recent study from Western Australia finds clear effects from a metric based on a heat index (Nyadanu et al. 2024). While some parts of Western Australia have a tropical climate, the vast majority of the population lives in Perth, which has a dry, Mediterranean climate, and hence humidity will play little role in determining variation in the index.

3.7 Conclusions

We have found that in the Northern Territory of Australia, the way that we define and measure heat matters a lot for our conclusions on how prenatal exposure affects health at birth. Choice of metric can mean the difference between finding no effect, finding a large, practically significant effect or finding a statistically significant effect, but which is practically insignificant given normal fluctuations in heat exposure.

In our context, we find that a heat metric allowing for non-linear effects of both minimum and maximum daily air temperature provides the best fit. Such a metric may be particularly valuable in tropical climates, where correlations between daily maximum and minimum temperatures are low (Figure 1).

Recent methodological discussions by Brimicombe et al. (2024) and Leung (2023) have recommended constructing heat metrics based on both air temperature and humidity, moving away from use of air temperature alone. In contrast, we find that such a metric provides a worse fit, and as a result, under-estimates the impact of typical variation in heat exposure. We do not know whether this is because of data quality or the underlying relationship between heat, humidity and health outcomes. However, the implication is that, at least with data that are currently widely available, researchers should be cautious when applying such metrics. Furthermore, even in climates where these measures are highly correlated with air temperature, meaning they can be used interchangeably, researchers' decisions around how to allow for non-linear and within-day temperature variation may be more important than their choice of underlying data series.

Overall, our estimates show that both moderate heat and more extreme heat contribute to poorer health at birth. The relative contributions of different heat levels and the relevant timing of exposure are not stable across different heat metrics or climate zones, and therefore we do not draw any strong conclusions on what levels and timing of heat exposure matter most. A range of unobserved factors are likely to influence these thresholds in any given context. For instance, medical advice, culture, work norms for pregnant women, and adoption of air conditioning and other adaptations will affect the size of this relationship and the relative contributions of heat stress versus heat strain-related conditions. Consequently, while our analysis is informative, it should be replicated in diverse contexts to test the generalisability of our findings. To understand these

mechanisms further, it would also be helpful to learn more about how pregnant people perceive and respond to heat exposure risks. For instance, Kc et al. (2024) conduct a survey assessing perceived risks of heat exposure among pregnant people in Nepal – the same survey could be carried out in other contexts to better understand behavioural responses to heat exposure.

While uncertain, our central estimate is that normal levels of heat experienced in the hottest months of the year could be contributing to a 4.5 percentage point higher risk of preterm birth in the NT – rates 45% higher than we would predict in babies in utero during the coolest 9 months of the year. The scale of these fluctuations, and their longer-term implications for children's development, may be even greater in lower-resource settings with similar climates, where healthcare and remedial neonatal care are less readily available. Furthermore, as climate change leads to additional exposure for populations previously unexposed to extreme heat or prolonged periods of moderate heat, pregnant people who do not have the habits and adaptations that are already present in regions such as the NT may experience more acute effects.

We identify two important directions for further research. First, analysing the long-term health impacts of heat exposure, looking beyond measures of health at birth – which are themselves imperfect indicators of fetal health and development (Goldenberg et al. 2008). As noted in our conceptual framework, outcomes like preterm birth are informative, but not all children who are born preterm will experience poorer health later in life, and similarly, not all children who experience poorer health later in life as a result of in utero heat exposure will have been born preterm. Second, analysing dose-response relationships between heat exposure and birth outcomes, and investigating whether there may be cumulative effects of continued exposure and/or periods during pregnancy where heat exposure poses the greatest risk.

Appendix A: Correlations between heat measures by city

Temperate climates

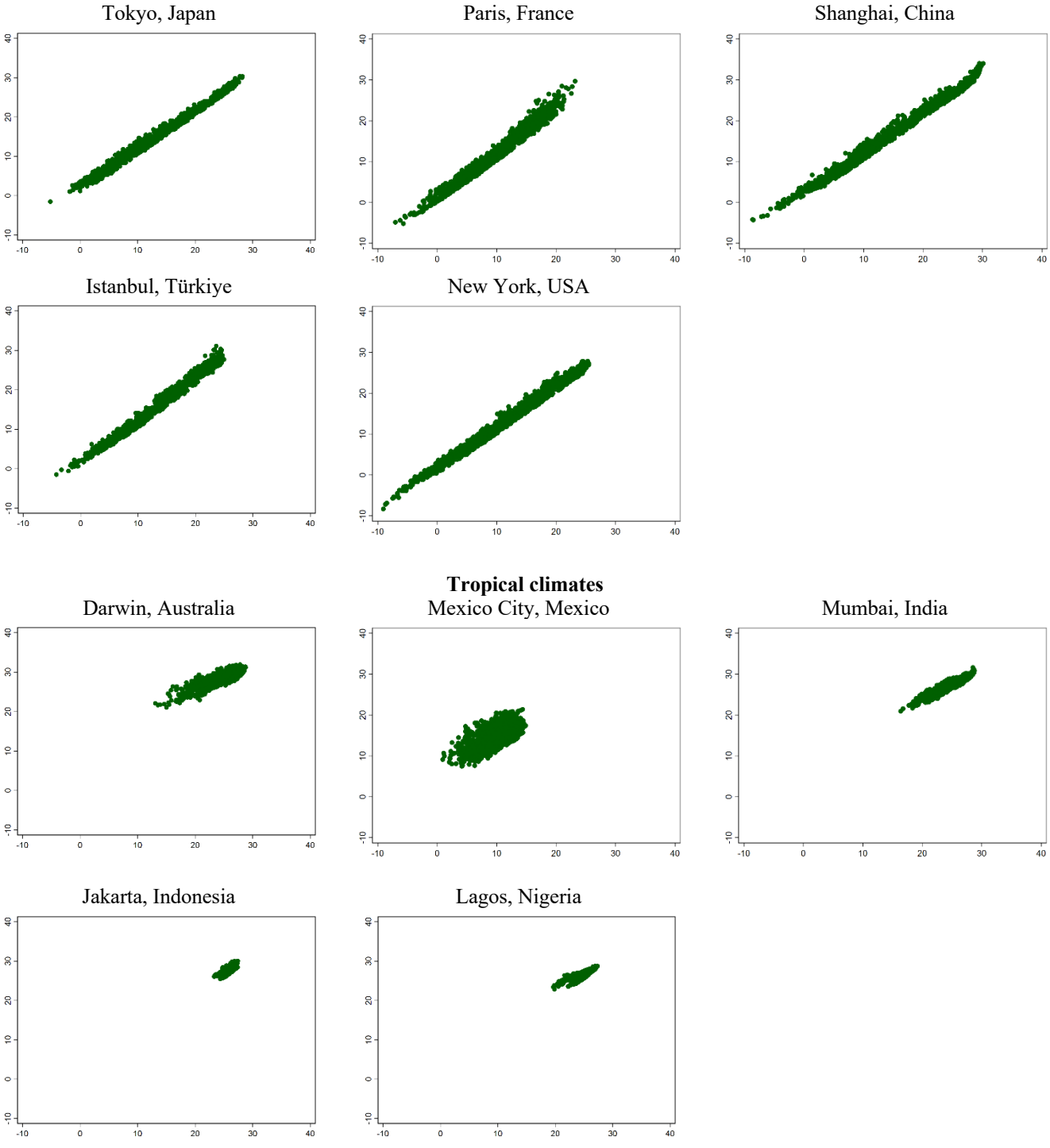


Figure A1: Scatter plot of daily average air temperatures and daily average wet bulb temperatures, 2020-2023

Source: NASA reanalysis data

Appendix B: NASA satellite vs observational weather data

The weather data we use in this chapter comes from NASA satellite data, and are estimated using NASA's MERRA-2 model.

We also have observed weather data from the Australian Bureau of Meteorology for some, but not all, parts of the NT. As a test of how well these two series relate to each other, below we present graphs plotting the NASA modelled data (blue lines) and the Bureau of Meteorology observed data (red lines) together, in Alice Springs. As is evident here, while the estimates are not exactly the same, the two series line up well. However, there are some periods without complete observational data (e.g. Figure B1.B), in which the NASA data are preferrable. This issue is more acute in more remote areas, which is why we use the NASA data in our analysis.

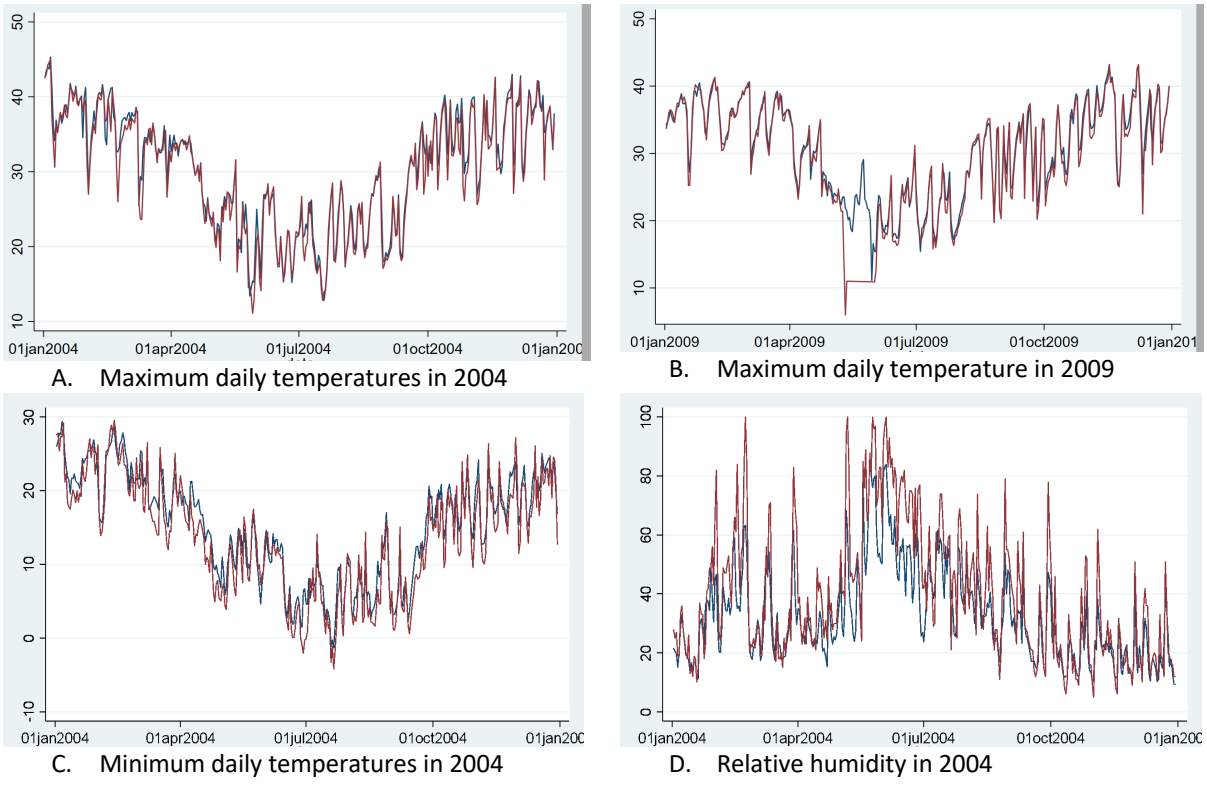


Figure B1: Observational and NASA data in Alice Springs
The above figures plot NASA data (blue lines) against Australian Bureau of Meteorology data (red lines) for the same series in Alice Springs.
Source: NASA reanalysis data and Australian Bureau of Meteorology

Appendix C: Alternative specifications of fixed effects

		(1)	(2)	(3)	(4)	(5)
	<20	0.000	0.000	0.000	-0.000	-0.000
		(0.001)	(0.001)	(0)	(0)	(0)
	20-25	0.001	0.001	-0.000	0.000	0.000
		(0)	(0.001)	(0.001)	(0.001)	(0.001)
1st trimester	30-35	0.001**	0.001***	0.000**	0.000	0.000***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	35-40	0.001**	0.001**	0.001**	0.000	0.001**
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	40+	0.000	0.000	0.001*	0.000	0.000
		(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
	<20	-0.001*	-0.002**	-0.001	0.000	-0.001
		(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
	20-25	-0.001***	-0.001***	0.000	0.000	0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
2nd trimester	30-35	0.000	0.000	0.000	0.000	0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	35-40	-0.000	-0.000	-0.000	-0.000	-0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	40+	-0.000	-0.000	0.000	0.000	0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	<20	-0.003***	-0.003***	-0.000	-0.000	-0.000
		(0.001)	(0.001)	(0)	(0)	(0)
	20-25	-0.002**	-0.001	-0.001*	-0.001	-0.001
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
3rd trimester	30-35	0.000*	0.000*	-0.000	-0.000	-0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	35-40	0.000	0.000	0.000	0.000	0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	40+	0.001	0.001	0.001	0.001	0.001
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
N		34258	34258	34258	34258	34258
R-sq		0.086	0.058	0.026	0.023	0.023
Fixed effects						
Month					x	x
Year					x	
Month-year		x	x	x		
Cluster				x	x	x
Cluster-month			x			
Cluster-month-sex		x				

Table C1: Regression estimates with alternative fixed effects specifications

Note: This table shows regression estimates using the model specified in equation (1), with our benchmark heat metric. Each column presents estimates from a different approach to specifying the time and location fixed effects. Cluster-robust standard errors in parentheses.

Appendix D: Regression tables

		Preterm birth	Birthweight (grams)	Small for Gestational Age	Apgar 5 score	Admitted to Special Care Nursery
	Max	-0.000	0.493	-0.001	-0.003	0.002***
	<20	(0.001)	(1.217)	(0.000)	(0.002)	(0.001)
	Max	0.001	-0.507	-0.000	0.001	-0.000
1st	20-25	(0.000)	(1.098)	(0.001)	(0.002)	(0.001)
trime	Max	0.001**	-0.813***	0.000	0.000	0.001***
ster	30-35	(0.000)	(0.212)	(0.000)	(0.001)	(0.000)
	Max	0.001**	-0.341	0.001**	-0.000	0.001***
	35-40	(0.000)	(0.269)	(0.000)	(0.001)	(0.000)
	Max	0.000	-1.160	0.002***	0.003	-0.000
	40+	(0.001)	(0.855)	(0.000)	(0.004)	(0.001)
	Max	-0.001*	2.911*	-0.001	0.000	-0.001
	<20	(0.001)	(1.518)	(0.001)	(0.003)	(0.001)
	Max	-0.001***	0.781	0.000	0.000	-0.001
	20-25	(0.000)	(0.975)	(0.001)	(0.002)	(0.001)
2nd	Max	0.000	-0.535**	0.000*	0.000	0.000
trime	30-35	(0.000)	(0.229)	(0.000)	(0.001)	(0.000)
ster	Max	-0.000	0.144	0.000	-0.000	0.000
	35-40	(0.000)	(0.207)	(0.000)	(0.001)	(0.000)
	Max	-0.000	0.159	-0.001*	0.002	0.000
	40+	(0.000)	(0.695)	(0.000)	(0.002)	(0.001)
	Max	-0.003***	3.477***	-0.000	0.008**	0.001
	<20	(0.001)	(0.906)	(0.001)	(0.003)	(0.001)
	Max	-0.002**	1.176	0.000	-0.001	-0.002*
	20-25	(0.001)	(1.151)	(0.001)	(0.003)	(0.001)
3rd	Max	0.000*	-0.600**	-0.000	-0.001	0.000***
trime	30-35	(0.000)	(0.247)	(0.000)	(0.001)	(0.000)
ster	Max	0.000	-0.669	-0.000	-0.001	0.000**
	35-40	(0.000)	(0.452)	(0.000)	(0.001)	(0.000)
	Max	0.001	-1.079	-0.000	-0.005	0.001
	40+	(0.001)	(1.178)	(0.001)	(0.004)	(0.001)
Constant		0.040**	3,404.696***	0.079***	8.972***	0.036*
		(0.018)	(40.737)	(0.018)	(0.092)	(0.021)
N		34,258	34,258	34,258	34,258	34,258
R-sq		0.086	0.125	0.087	0.080	0.088

Table D1: benchmark model

Note: this table shows the regression coefficients and cluster-robust standard errors in parentheses from the model specified in equation (1) using the benchmark heat metric. Estimates are shown for preterm birth and four other measures of health at birth. As specified in equation (1), the regressions also include covariates (mother's age, Aboriginal status, whether mother's first pregnancy), month-year fixed effects and location-month-sex fixed effects (these are absorbed using the Stata 'areg', which affects the intercept but not the coefficients).

	Preterm birth	Birthweight (grams)	Small for Gestational Age	Apgar 5 score	Special Care Nursery
Max temp – 1st trimester avg	0.001	1.120	0.005***	0.005	0.000
	(0.001)	(2.215)	(0.001)	(0.006)	(0.002)
Max temp – 2nd trimester avg	-0.001	0.726	-0.000	0.002	0.002
	(0.002)	(2.881)	(0.001)	(0.006)	(0.002)
Max temp – 3rd trimester avg	0.007**	-8.761**	-0.002	-0.019	0.005**
	(0.003)	(3.492)	(0.001)	(0.012)	(0.002)
Constant	-0.135	3,541.611***	0.032	9.333***	-0.107
	(0.126)	(201.286)	(0.045)	(0.383)	(0.078)
N	34,258	34,258	34,258	34,258	34,258
R-sq	0.085	0.124	0.087	0.080	0.087

Table D2: average max temperature specification

Note: this table shows the regression coefficients and cluster-robust standard errors in parentheses from the model specified in equation (1) using the trimester-average daily maximum heat metric. Estimates are shown for preterm birth and four other measures of health at birth. As specified in equation (1), the regressions also include covariates (mother's age, Aboriginal status, whether mother's first pregnancy), month-year fixed effects and location-month-sex fixed effects (these are absorbed using the Stata 'areg', which affects the intercept but not the coefficients).

	Preterm birth	Birthweight (grams)	Small for Gestational Age	Apgar 5 score	Special Care Nursery
Heatwave in 1st tri	0.001	1.096	0.000	-0.003	0.002
	(0.002)	(3.195)	(0.002)	(0.011)	(0.003)
Heatwave in 2nd tri	0.004*	-5.665**	-0.000	0.004	0.003
	(0.002)	(2.525)	(0.002)	(0.010)	(0.003)
Heatwave in 3rd tri	0.006	-6.456	-0.002	-0.003	0.005
	(0.004)	(4.850)	(0.002)	(0.014)	(0.003)
Constant	0.061***	3,326.564***	0.112***	8.955***	0.094***
	(0.013)	(18.451)	(0.011)	(0.045)	(0.024)
N	34,258	34,258	34,258	34,258	34,258
R-sq	0.084	0.124	0.087	0.079	0.087

Table D3: Heatwave specification

Note: this table shows the regression coefficients and cluster-robust standard errors in parentheses from the model specified in equation (1) using the heatwave count heat metric. Heatwaves are defined based on the Australian Bureau of Meteorology's methodology. Estimates are shown for preterm birth and four other measures of health at birth. As specified in equation (1), the regressions also include covariates (mother's age, Aboriginal status, whether mother's first pregnancy), month-year fixed effects and location-month-sex fixed effects (these are absorbed using the Stata 'areg', which affects the intercept but not the coefficients).

		Preterm birth	Birthweight (grams)	Small for Gestational Age	Apgar 5 score	Special Care Nursery
1st trimester	WB max <10	0.001	6.253	-0.000	0.011	-0.000
		(0.002)	(4.215)	(0.003)	(0.008)	(0.002)
	WB max 10-15	-0.001	3.073	0.003	0.007	-0.001
		(0.002)	(2.472)	(0.002)	(0.006)	(0.002)
	WB max 15-20	-0.000	1.750	0.002	0.008	-0.001
		(0.001)	(1.532)	(0.001)	(0.005)	(0.001)
	WB max 25-30	-0.000	0.218	0.000	0.000	-0.000
		(0.000)	(0.422)	(0.000)	(0.001)	(0.000)
	WB max 30+	0.000	0.289	0.000	0.001	0.000
		(0.001)	(1.126)	(0.001)	(0.003)	(0.001)
	WB avg <10	0.000	-1.802	-0.003*	-0.005	-0.000
		(0.001)	(2.249)	(0.001)	(0.005)	(0.001)
	WB avg 10-15	0.000	-1.497	-0.001	-0.002	-0.000
		(0.001)	(1.045)	(0.001)	(0.005)	(0.001)
	WB avg 20-25	0.000	-0.904	-0.000	0.000	0.000
		(0.001)	(0.621)	(0.000)	(0.001)	(0.001)
	WB avg 25+	0.000	-0.237	-0.001	0.002	0.001
		(0.001)	(0.654)	(0.000)	(0.002)	(0.001)
2nd trimester	WB max <10	0.000	-1.603	-0.001	-0.015**	-0.004*
		(0.002)	(3.655)	(0.002)	(0.007)	(0.002)
	WB max 10-15	-0.000	1.829	-0.002	0.002	-0.003
		(0.001)	(2.574)	(0.002)	(0.007)	(0.002)
	WB max 15-20	0.000	-0.687	-0.001	-0.005	-0.001
		(0.001)	(1.826)	(0.001)	(0.004)	(0.001)
	WB max 25-30	0.000	-0.690	0.000	-0.001	0.000
		(0.000)	(0.699)	(0.000)	(0.001)	(0.000)
	WB max 30+	0.000	-0.217	0.001**	0.001	0.002***
		(0.001)	(0.884)	(0.000)	(0.003)	(0.001)
	WB avg <10	-0.000	-0.151	0.001	0.002	0.002
		(0.001)	(2.392)	(0.002)	(0.006)	(0.002)
	WB avg 10-15	-0.002*	0.894	-0.000	0.005	-0.000
		(0.001)	(1.623)	(0.001)	(0.003)	(0.001)
	WB avg 20-25	-0.001*	-0.362	-0.000	0.002	-0.001
		(0.000)	(1.071)	(0.000)	(0.001)	(0.001)
	WB avg 25+	-0.001	-0.164	-0.000	0.004**	-0.001**
		(0.001)	(1.100)	(0.000)	(0.002)	(0.000)
3rd trimester	WB max <10	0.000	-1.572	0.000	-0.000	0.001
		(0.002)	(3.346)	(0.002)	(0.008)	(0.002)
	WB max 10-15	-0.003	2.795	0.002*	0.011*	-0.001
		(0.002)	(3.254)	(0.001)	(0.006)	(0.003)
	WB max 15-20	0.001	0.001	0.000	-0.003	-0.001
		(0.001)	(2.447)	(0.001)	(0.005)	(0.001)
	WB max 25-30	-0.001**	0.455	0.000	0.003**	-0.001**
		(0.000)	(0.406)	(0.000)	(0.001)	(0.000)
	WB max 30+	0.000	-0.870	-0.000	0.002	0.000
		(0.001)	(2.142)	(0.001)	(0.002)	(0.001)

WB avg <10	0.001	-0.276	-0.001	-0.006	-0.000
	(0.002)	(2.989)	(0.001)	(0.006)	(0.002)
WB avg 10-15	-0.000	-1.068	0.001	0.001	0.001
	(0.001)	(2.916)	(0.001)	(0.004)	(0.001)
WB avg 20-25	0.000	-0.427	0.000	-0.001	0.000
	(0.000)	(0.714)	(0.000)	(0.002)	(0.001)
Wet bulb avg 25+	0.001	-0.818	0.000	0.000	0.001**
	(0.001)	(0.788)	(0.000)	(0.003)	(0.001)
Constant	0.132**	3,393.610***	0.096*	8.522***	0.123
	(0.063)	(108.642)	(0.057)	(0.341)	(0.096)
N	34,258	34,258	34,258	34,258	34,258
R-sq	0.085	0.125	0.087	0.081	0.088

Table D4: wet bulb metric

Note: this table shows the regression coefficients and cluster-robust standard errors in parentheses from the model specified in equation (1) using the wet bulb heat metric. Estimates are shown for preterm birth and four other measures of health at birth. As specified in equation (1), the regressions also include covariates (mother's age, Aboriginal status, whether mother's first pregnancy), month-year fixed effects and location-month-sex fixed effects (these are absorbed using the Stata 'areg', which affects the intercept but not the coefficients).

		Preterm birth	Birthweight (grams)	Small for Gestational Age	Apgar 5 score	Special Care Nursery
1st trimester	Max <20	0.001	-3.114*	0.000	-0.002	0.003***
		(0.001)	(1.637)	(0.001)	(0.004)	(0.001)
	Max 20-25	0.002**	0.011	-0.000	0.005*	0.001
		(0.001)	(1.375)	(0.001)	(0.003)	(0.001)
	Max 30-35	0.000	-0.841***	0.000	0.001	0.000***
		(0.000)	(0.317)	(0.000)	(0.001)	(0.000)
	Max 35-40	0.000	-0.375	0.001***	0.002	0.001***
		(0.000)	(0.403)	(0.000)	(0.001)	(0.000)
	Max 40+	-0.000	-1.307	0.002***	0.006	-0.001
		(0.001)	(1.096)	(0.001)	(0.004)	(0.001)
	Min <5	-0.002	5.537***	-0.001	0.001	-0.001
		(0.001)	(1.388)	(0.001)	(0.005)	(0.001)
	Min 5-10	-0.002**	-0.695	0.001	-0.006**	-0.001
		(0.001)	(1.360)	(0.001)	(0.003)	(0.001)
	Min 10-15	0.000	-0.073	0.000	0.004*	0.000
		(0.000)	(0.824)	(0.001)	(0.002)	(0.000)
	Min 20-25	0.001***	-0.347	-0.000	-0.000	0.001
		(0.000)	(0.431)	(0.000)	(0.002)	(0.000)
	Min 25+	0.001***	-0.111	-0.001*	-0.001	0.001***
		(0.000)	(0.706)	(0.000)	(0.002)	(0.000)
2nd trimester	Max <20	-0.003**	1.731	-0.001	0.004	-0.002
		(0.001)	(1.861)	(0.001)	(0.005)	(0.002)
	Max 20-25	-0.002***	0.362	-0.000	0.002	-0.001
		(0.001)	(1.422)	(0.001)	(0.003)	(0.001)
	Max 30-35	0.000	-0.396*	0.000	0.001	0.000*
		(0.000)	(0.226)	(0.000)	(0.001)	(0.000)
	Max 35-40	0.000	0.023	-0.000	-0.001	0.001**
		(0.000)	(0.405)	(0.000)	(0.002)	(0.000)
	Max 40+	-0.000	0.734	-0.001*	0.003	0.001
		(0.001)	(0.727)	(0.000)	(0.002)	(0.001)
	Min <5	0.002	1.312	0.002	-0.006	0.001
		(0.001)	(1.910)	(0.002)	(0.005)	(0.002)
	Min 5-10	0.001	1.469	-0.000	-0.003	0.000
		(0.001)	(1.184)	(0.001)	(0.003)	(0.001)
	Min 10-15	0.001*	-1.556	0.000	-0.003	0.001
		(0.001)	(0.970)	(0.001)	(0.002)	(0.001)
	Min 20-25	0.000	-1.697***	0.000	-0.003*	0.000
		(0.000)	(0.515)	(0.000)	(0.002)	(0.000)
	Min 25+	0.001***	-2.217***	0.000	-0.004**	-0.000
		(0.000)	(0.404)	(0.000)	(0.002)	(0.000)
3rd trimester	Max <20	-0.003	-0.385	0.003***	0.007	0.001
		(0.002)	(2.815)	(0.001)	(0.008)	(0.001)
	Max 20-25	-0.002*	1.369	0.001*	-0.002	-0.001
		(0.001)	(2.481)	(0.001)	(0.003)	(0.001)
	Max 30-35	0.000**	-0.758**	-0.000	-0.001	0.000*
		(0.000)	(0.295)	(0.000)	(0.001)	(0.000)
	Max 35-40	0.001**	-1.631***	-0.000	-0.004**	0.001

	(0.000)	(0.416)	(0.000)	(0.002)	(0.000)
Max 40+	0.002*	-2.243*	-0.001	-0.009**	0.000
	(0.001)	(1.187)	(0.001)	(0.004)	(0.001)
Min <5	0.001	3.421	-0.003***	-0.003	-0.001
	(0.002)	(2.263)	(0.001)	(0.005)	(0.001)
Min 5-10	0.001	-0.853	-0.001	0.001	-0.000
	(0.001)	(2.661)	(0.001)	(0.003)	(0.001)
Min 10-15	0.002*	-2.612**	0.001**	-0.008**	0.000
	(0.001)	(1.049)	(0.000)	(0.004)	(0.001)
Min 20-25	0.000	-0.563	0.000	0.001	0.000
	(0.000)	(0.644)	(0.000)	(0.002)	(0.000)
Min 25+	0.000	-0.555	0.001***	0.002	0.000
	(0.000)	(0.491)	(0.000)	(0.002)	(0.000)
Cons t.	-0.112***	3,647.727***	0.052	9.280***	-0.054
	(0.037)	(59.625)	(0.041)	(0.201)	(0.054)
N	34,258	34,258	34,258	34,258	34,258
R-sq	0.087	0.126	0.088	0.081	0.088

Table D5: Maximum and minimum piecewise linear specification

Note: this table shows the regression coefficients and cluster-robust standard errors in parentheses from the model specified in equation (1) using the Max + min heat metric. Estimates are shown for preterm birth and four other measures of health at birth. As specified in equation (1), the regressions also include covariates (mother's age, Aboriginal status, whether mother's first pregnancy), month-year fixed effects and location-month-sex fixed effects (these are absorbed using the Stata 'areg', which affects the intercept but not the coefficients).

		Preterm birth	Birthweight (grams)	Small for Gestational Age	Apgar 5 score	Special Care Nursery
1st tri	WB max <10	0.001	4.252	-0.002	0.007	-0.000
		(0.002)	(3.658)	(0.002)	(0.006)	(0.001)
	WB max 10-15	-0.001*	1.709*	-0.000	0.002	-0.002*
		(0.001)	(0.917)	(0.001)	(0.003)	(0.001)
	WB max 15-20	-0.000	0.971	-0.000	0.005***	-0.001**
		(0.001)	(0.628)	(0.001)	(0.001)	(0.000)
	WB max 25-30	-0.000	0.324	-0.000	0.001	0.000
		(0.000)	(0.341)	(0.000)	(0.001)	(0.000)
	WB max 30+	-0.000	0.682	0.000	0.001	-0.000
		(0.000)	(0.829)	(0.001)	(0.002)	(0.001)
2nd tri	WB max <10	-0.000	-1.020	-0.000	-0.013**	-0.003***
		(0.001)	(2.534)	(0.001)	(0.005)	(0.001)
	WB max 10-15	0.000	1.534	0.000	0.003	-0.001
		(0.001)	(0.930)	(0.001)	(0.003)	(0.001)
	WB max 15-20	-0.000	0.259	-0.000	-0.001	-0.000
		(0.000)	(0.707)	(0.000)	(0.002)	(0.000)
	WB max 25-30	-0.000	-0.661**	0.000	0.000	-0.000
		(0.000)	(0.333)	(0.000)	(0.001)	(0.000)
	WB max 30+	0.000	-0.124	0.001***	0.000	0.001**
		(0.000)	(0.512)	(0.000)	(0.002)	(0.000)
3rd tri	WB max <10	0.002	-1.774	-0.000	-0.004	0.002*
		(0.001)	(1.598)	(0.002)	(0.006)	(0.001)
	WB max 10-15	-0.002***	2.683***	-0.000	0.004	-0.001*
		(0.001)	(0.873)	(0.000)	(0.003)	(0.001)
	WB max 15-20	0.001	-0.736	0.000	-0.004	-0.000
		(0.001)	(0.714)	(0.000)	(0.003)	(0.000)
	WB max 25-30	-0.000	-0.092	0.000*	0.003*	0.000
		(0.000)	(0.359)	(0.000)	(0.001)	(0.000)
	WB max 30+	0.001	-1.243	0.000	0.000	0.001
		(0.001)	(1.871)	(0.000)	(0.003)	(0.001)
Constant		0.100***	3,331.013***	0.090***	8.717***	0.117***
		(0.027)	(44.384)	(0.026)	(0.132)	(0.027)
N		34,258	34,258	34,258	34,258	34,258
R-sq		0.085	0.125	0.087	0.080	0.087

Table D6: maximum wet bulb (WB) piecewise specification

Note: this table shows the regression coefficients and cluster-robust standard errors in parentheses from the model specified in equation (1) using the counts of the number of days with maximum wet bulb temperatures under 10, 10-15, 15-20, 25-30, and 30+. We estimate maximum wet bulb temperatures using Stull's (2011) equation based on maximum air temperature and humidity. Estimates are shown for preterm birth and four other measures of health at birth. As specified in equation (1), the regressions also include covariates (mother's age, Aboriginal status, whether mother's first pregnancy), month-year fixed effects and location-month-sex fixed effects (these are absorbed using the Stata 'areg', which affects the intercept but not the coefficients).

		Preterm birth	Birthweight (grams)	Small for Gestational Age	Apgar 5 score	Special Care Nursery
1st tri	WB avg <10	-0.000	0.745	-0.001	0.001	-0.001*
		(0.000)	(0.978)	(0.001)	(0.002)	(0.001)
	WB avg 10-15	0.000	-0.495	0.000	0.004	-0.000
		(0.001)	(0.909)	(0.001)	(0.002)	(0.001)
	WB avg 20-25	0.000	-0.633	-0.000	0.000	-0.000
		(0.001)	(0.556)	(0.000)	(0.001)	(0.001)
	WB avg 25+	-0.000	0.139	-0.001**	0.002	-0.000
		(0.000)	(0.459)	(0.000)	(0.001)	(0.000)
2nd tri	WB avg <10	0.000	0.694	0.000	0.001	-0.001
		(0.000)	(0.500)	(0.001)	(0.002)	(0.001)
	WB avg 10-15	-0.001*	0.283	-0.000	0.001	-0.001
		(0.001)	(1.312)	(0.001)	(0.003)	(0.001)
	WB avg 20-25	-0.001**	-0.388	-0.000	0.001	-0.001
		(0.000)	(0.885)	(0.000)	(0.001)	(0.001)
	WB avg 25+	-0.000	-0.567	-0.000	0.001	-0.001*
		(0.000)	(0.661)	(0.000)	(0.002)	(0.000)
3rd tri	WB avg <10	-0.001	1.519**	-0.000	0.001	-0.001
		(0.001)	(0.672)	(0.000)	(0.003)	(0.000)
	WB avg 10-15	0.001	-1.161	0.001	-0.003	0.000
		(0.001)	(1.578)	(0.001)	(0.005)	(0.001)
	WB avg 20-25	0.000	-0.198	0.000	0.001	0.000
		(0.000)	(0.737)	(0.000)	(0.002)	(0.001)
	WB avg 25+	0.000	-0.409	0.000	0.003	0.000
		(0.001)	(0.761)	(0.000)	(0.003)	(0.001)
	Constant	0.110*	3,368.704***	0.122***	8.644***	0.156
		(0.058)	(120.327)	(0.043)	(0.279)	(0.102)
	N	34,258	34,258	34,258	34,258	34,258
	R-sq	0.084	0.124	0.087	0.080	0.087

Table D7: average wet bulb piecewise linear

Note: this table shows the regression coefficients and cluster-robust standard errors in parentheses from the model specified in equation (1) using counts of the number of days with average wet bulb temperature under 10, 10-15, 20-25, and 25+. Estimates are shown for preterm birth and four other measures of health at birth. As specified in equation (1), the regressions also include covariates (mother's age, Aboriginal status, whether mother's first pregnancy), month-year fixed effects and location-month-sex fixed effects (these are absorbed using the Stata 'areg', which affects the intercept but not the coefficients).

	(1)	(2)	(3)	(4)
Statistical significance: F-test of joint significance p-value)				
Max temp	0.000	0.000	0.000	0.000
Min temp		0.000	0.009	0.013
Wet bulb max			0.003	0.010
Heatwaves				0.736
Model fit				
RMSE	0.301	0.300	0.300	0.300
Adjusted R-squared	0.036	0.037	0.037	0.037

Table D8: Model build

Appendix E: Heterogeneity analysis – by Aboriginal status and climate zone

		Aboriginal	Non-Aboriginal
1st tri	Max <20	0.002	0.000
		(0.001)	(0.001)
	Max 20-25	0.002**	0.002***
		(0.001)	(0.001)
	Max 30-35	0.000	0.000
		(0.000)	(0.000)
	Max 35-40	0.000	0.001***
		(0.001)	(0.000)
	Max 40+	-0.000	-0.002**
		(0.001)	(0.001)
	Min <5	-0.001	-0.002**
		(0.002)	(0.001)
	Min 5-10	-0.002	-0.001*
		(0.002)	(0.001)
	Min 10-15	-0.000	0.000
		(0.001)	(0.001)
	Min 20-25	0.001	0.001***
		(0.001)	(0.000)
	Min 25+	0.001	0.001***
		(0.001)	(0.000)
2nd tri	Max <20	-0.003	-0.002
		(0.002)	(0.001)
	Max 20-25	-0.002	-0.002**
		(0.001)	(0.001)
	Max 30-35	-0.000	0.000**
		(0.000)	(0.000)
	Max 35-40	0.000	0.001***
		(0.001)	(0.000)
	Max 40+	-0.001	0.000
		(0.001)	(0.000)
	Min <5	0.000	0.002
		(0.002)	(0.001)
	Min 5-10	0.000	0.002**
		(0.002)	(0.001)
	Min 10-15	0.001	0.001*
		(0.001)	(0.001)
	Min 20-25	0.000	0.000
		(0.001)	(0.001)
	Min 25+	0.001	0.001
		(0.001)	(0.001)
3rd tri	Max <20	-0.007***	-0.000
		(0.003)	(0.001)
	Max 20-25	-0.004**	0.002
		(0.002)	(0.002)
	Max 30-35	0.000	0.000**

	(0.000)	(0.000)
Max 35-40	0.002***	0.000
	(0.001)	(0.001)
Max 40+	0.001	0.002
	(0.001)	(0.001)
Min <5	0.003	-0.002***
	(0.002)	(0.000)
Min 5-10	0.002	-0.001
	(0.002)	(0.001)
Min 10-15	0.001	0.002***
	(0.001)	(0.001)
Min 20-25	-0.000	0.000
	(0.001)	(0.000)
Min 25+	-0.000	0.000**
	(0.001)	(0.000)
Constant	0.024	-0.153***
	(0.063)	(0.047)
N	14,646	19,612
R-sq	0.132	0.042

Table E1: Regression estimates for probability of preterm birth, preferred metric

Note: this table shows the regression coefficients and cluster-robust standard errors in parentheses from the model specified in equation (1) using the Max + Min heat metric, with the sample split by whether the child is Aboriginal or not. The outcome is preterm birth. As specified in equation (1), the regressions also include covariates (mother's age, Aboriginal status, whether mother's first pregnancy), month-year fixed effects and location-month-sex fixed effects (these are absorbed using the Stata 'areg', which affects the intercept but not the coefficients).

		Tropical	Arid	Tropical	Arid
1st tri	Max <20	0.014	0.002	0.021	0.003**
		(0.027)	(0.002)	(0.023)	(0.002)
	Max 20-25	0.001**	0.002**	0.002**	0.004**
		(0.001)	(0.001)	(0.001)	(0.002)
	Max 30-35	0.000	0.001	0.000	0.001
		(0.000)	(0.001)	(0.000)	(0.001)
	Max 35-40	0.000	0.003***	0.000	0.003*
		(0.000)	(0.001)	(0.000)	(0.002)
	Max 40+	-0.000	0.004***	-0.000	0.002
		(0.001)	(0.001)	(0.001)	(0.002)
	Min <5			-0.007*	0.001
				(0.004)	(0.002)
	Min 5-10			-0.004*	-0.001
				(0.002)	(0.002)
	Min 10-15			0.000	0.001
				(0.001)	(0.002)
	Min 20-25			0.001***	0.001
				(0.000)	(0.001)
	Min 25+			0.001***	0.003
				(0.000)	(0.002)
2nd tri	Max <20	0.015	-71.278***	0.013	-0.002
		(0.023)	(18.577)	(0.030)	(0.002)
	Max 20-25	-0.002***	3.097*	-0.003***	-0.001
		(0.001)	(1.618)	(0.001)	(0.002)
	Max 30-35	-0.000	-0.596*	0.000	0.002
		(0.000)	(0.325)	(0.000)	(0.002)
	Max 35-40	-0.000	0.168	0.001	0.001
		(0.000)	(0.296)	(0.000)	(0.002)
	Max 40+	-0.000	-1.283	-0.000	0.002
		(0.001)	(1.155)	(0.001)	(0.003)
	Min <5			-0.090*	0.004*
				(0.047)	(0.003)
	Min 5-10			0.000	0.003
				(0.001)	(0.003)
	Min 10-15			0.000	0.005**
				(0.001)	(0.002)
	Min 20-25			0.001**	0.003
				(0.000)	(0.003)
	Min 25+			0.001***	0.002
				(0.000)	(0.003)
3rd tri	Max <20	-0.008	43.241	-0.027	-0.004
		(0.011)	(36.817)	(0.018)	(0.006)
	Max 20-25	-0.002	2.920	-0.003	-0.001
		(0.002)	(2.822)	(0.002)	(0.003)
	Max 30-35	0.000**	-0.469	0.000***	0.004***
		(0.000)	(0.316)	(0.000)	(0.001)
	Max 35-40	0.000	-0.401	0.001***	0.007***
		(0.000)	(0.474)	(0.000)	(0.002)
	Max 40+	0.003***	-2.746*	0.003***	0.008***
		(0.001)	(1.634)	(0.001)	(0.002)
	Min <5			0.019	0.004
				(0.020)	(0.005)
	Min 5-10			0.003	0.003
				(0.003)	(0.003)
	Min 10-15			0.000	0.006***
				(0.001)	(0.002)
	Min 20-25			0.000	-0.002**
				(0.000)	(0.001)
	Min 25+			0.000	-0.003*

			(0.000)	(0.002)
Constant	0.076***	0.070***	0.025	-0.168***
	(0.020)	(0.189)	(0.051)	(0.253)
N	26,521	7,379	26,521	7,379
R-sq	0.069	0.170	0.070	0.179

Table E2: estimates by climate – air temperature

Note: this table shows the regression coefficients and cluster-robust standard errors in parentheses from the model specified in equation (1) using both benchmark heat metric and the Max + Min heat metric, with the sample split by climate zone Estimates are shown for preterm birth and four other measures of health at birth. As specified in equation (1), the regressions also include covariates (mother's age, Aboriginal status, whether mother's first pregnancy), month-year fixed effects and location-month-sex fixed effects (these are absorbed using the Stata 'areg', which affects the intercept but not the coefficients).

		tropical	arid	tropical	arid
1st trimester	WB max <10	0.010	0.002		
		(0.008)	(0.003)		
	WB max 10-15	-0.001	-0.003		
		(0.003)	(0.004)		
	WB max 15-20	0.002	-0.004		
		(0.001)	(0.003)		
	WB max 25-30	-0.000	-0.002		
		(0.000)	(0.001)		
	WB max 30+	-0.000	0.005		
		(0.001)	(0.007)		
	WB avg <10	-0.001	0.002	-0.001	-0.001
		(0.002)	(0.004)	(0.001)	(0.001)
	WB avg 10-15	0.000	0.002	0.002	-0.002*
		(0.001)	(0.003)	(0.001)	(0.001)
	WB avg 20-25	0.001	0.001	0.000	-0.001
		(0.001)	(0.002)	(0.001)	(0.002)
	WB avg 25+	0.001	0.001	0.000	-0.000
		(0.001)	(0.003)	(0.000)	(0.002)
2nd trimester	WB max <10		-0.003		
			(0.003)		
	WB max 10-15	-0.002	-0.001		
		(0.002)	(0.004)		
	WB max 15-20	-0.001	-0.002		
		(0.002)	(0.003)		
	WB max 25-30	0.001	0.001		
		(0.000)	(0.003)		
	WB max 30+	0.001	-0.006		
		(0.001)	(0.012)		
	WB avg <10	0.002	0.000	0.001	-0.001
		(0.002)	(0.004)	(0.001)	(0.001)
	WB avg 10-15	-0.003***	0.001	-0.003***	-0.001
		(0.001)	(0.003)	(0.001)	(0.001)
	WB avg 20-25	-0.001	-0.002	-0.001**	-0.001
		(0.001)	(0.002)	(0.000)	(0.001)
	WB avg 25+	-0.001**	-0.005	-0.001**	-0.004**
		(0.001)	(0.004)	(0.000)	(0.002)
3rd trimester	WB max <10		-0.007		
			(0.007)		
	WB max 10-15	-0.004	-0.005		
		(0.005)	(0.006)		
	WB max 15-20	-0.000	-0.002		
		(0.002)	(0.005)		
	WB max 25-30	-0.000	-0.000		
		(0.000)	(0.002)		
	WB max 30+	0.001	-0.017***		
		(0.001)	(0.004)		
	WB avg <10	0.002	0.003	-0.000	-0.001
		(0.004)	(0.006)	(0.001)	(0.001)
	WB avg 10-15	-0.002	0.003	-0.001	0.002
		(0.002)	(0.005)	(0.001)	(0.002)
	WB avg 20-25	0.000	-0.000	-0.000	-0.001
		(0.000)	(0.002)	(0.001)	(0.001)
	WB avg 25+	0.000	-0.001	-0.000	-0.001
		(0.001)	(0.002)	(0.001)	(0.001)
Constant		0.025	0.091	0.266*	0.153
		(0.110)	(0.136)	(0.094)	(0.115)
N		26,521	7,379	26,521	7,379
R-sq		0.068	0.171	0.068	0.168

Table E3: estimates by climate -wet bulb temperature

Note: this table shows the regression coefficients and cluster-robust standard errors in parentheses from the model specified in equation (1) using both the Wet bulb heat metric, and a simpler metric based on average wet bulb temperatures alone, with the sample split by climate zone Estimates are shown for preterm birth and four other measures of health at birth. As specified in equation (1), the regressions also include covariates (mother's age, Aboriginal status, whether mother's first pregnancy), month-year fixed effects and location-month-sex fixed effects (these are absorbed using the Stata 'areg', which affects the intercept but not the coefficients)

.

	P-val: joint significance of heat exposure terms	Adj R^2	AIC
Benchmark	0.00	0.031	10240
Trimester average	0.00	0.031	10265
Heatwave count	0.00	0.031	10261
Max and min	**0.00**	**0.032**	**10207**
Wet bulb	0.00	0.030	10244

Table E4: Measures of model fit for each heat metric – tropical climate zone

Note: Each row of this table represents results from a regression using the indicated heat metric, for the sample in tropical climate zones only. The first column in this table shows p-values for an F-test of joint statistical significance of all terms in the heat exposure metric. We present adjusted R-squared as this penalises additional coefficients, and some metrics (Max and min and Wet bulb) have many more coefficients than the other metrics. The Akaike Information Criteria (AIC) is a measure of model fit suitable for comparing across models with different coefficients but the same outcome measure. A lower value indicates better fit. Sample size: 26,521.

	P-val: joint significance of heat exposure terms	Adj R^2	AIC
Benchmark	0.00	0.058	2477
Trimester average	0.00	0.055	2509
Heatwave count	0.00	0.051	2540
Max and min	**0.00**	**0.065**	**2401**
Wet bulb	0.00	0.056	2472

Table E5: Measures of model fit for each heat metric – arid climate zone

Note: Each row of this table represents results from a regression using the indicated heat metric, for the sample in arid climate zones only. The first column in this table shows p-values for an F-test of joint statistical significance of all terms in the heat exposure metric. We present adjusted R-squared as this penalises additional coefficients, and some metrics (Max and min and Wet bulb) have many more coefficients than the other metrics. The Akaike Information Criteria (AIC) is a measure of model fit suitable for comparing across models with different coefficients but the same outcome measure. A lower value indicates better fit. Sample size: 7,379.

Appendix F: Prediction methodology

We use predictions to estimate the likely rate of preterm birth if all babies were in utero during the hottest 9 months of the year, and compare this to the likely rate of preterm birth if all babies were in utero during the coolest 9 months of the year. By generating these predictions for each heat metric, we are able to compare the implications of each heat metric like-for-like.

To do this, we generate predictions for all babies in our sample, with the point estimate reported in the main text representing the average across the full sample. However, we replace the heat conditions that each baby actually experienced with heat conditions typical of the hottest (coolest) 9 months of the year in the mother's place of residence. These typical conditions are calculated based on location-specific averages for 1996-2014 in May (for the hottest 9 months) and November (for the coolest 9 months).

These predicted values are given by:

$$\widehat{preterm}_{\iota tJ} = b_0 + \sum_{k}^{3} b_1^k heat_j^k + b_2 Aboriginalstatus + b_3 motherage + b_4 firstpregnancy + \gamma_{my} + \theta_{jmi}$$

where we substitute the estimated regression coefficients for the 'b's and fixed effects, substitute the child's actual characteristics for the covariates, and insert location-average values of heat exposure as 'heat'.

As an example, consider a female baby, it is her mother's first pregnancy, she is Aboriginal, aged between 25–29, and lives in Darwin. She was actually conceived in January 2001 and born in October 2001. Her predicted risk of preterm birth if she were instead born in May with the 'Trimester average' heat metric would be given by the following parameters:

variable	Coefficient	Value – May birth	Value – Nov birth
b_0	-0.323		
Heat_tri1	0.000585	32.22	30.04
Heat_tri2	-0.00103	30.66	29.59
Heat_tri3	0.00691	30.01	32.11
Aboriginal status	0.0567392	Yes	Yes
Mother's age=25-29	-0.005715	Yes	Yes
First pregnancy	0.0047182	Yes	Yes
Conception month-year (FE) = Jan 2001	0.209		
Location-month-sex (FE) = Darwin, January, female	0.002		
Predicted risk of preterm birth		0.138	0.153

In this example, a baby born in November actually has a higher risk of preterm birth than a baby born in May, because this heat metric estimates that it is third trimester heat exposure which has the largest impact – and average third trimester temperatures are higher for babies born in November than those born in May.

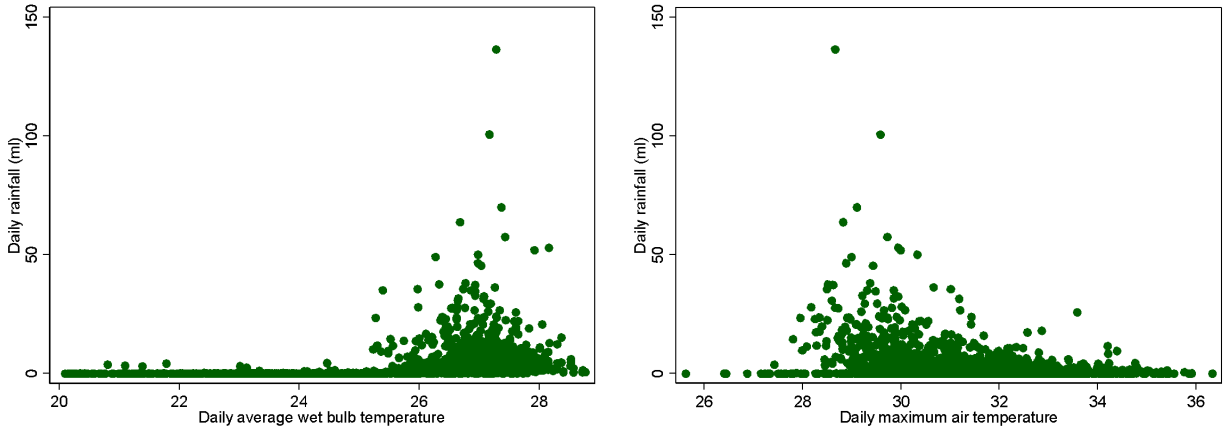
Note also that the intercept reported here is different from the one in our main results table – this is because our main estimates are estimated using the 'areg' command in Stata, which does not generate estimates of each fixed effect. For the purposes of this example, we have re-estimated our model entering the fixed effects terms manually. The main coefficients are the same in both cases, but the intercept is different.

We generate these predictions for every individual in the sample, then take an average across all individuals to generate our aggregate predictions. The standard errors we report are similarly the average of all individual standard prediction errors ('predict, stdp' in Stata).

Appendix G: Correlations among various weather metrics in Darwin

	Rainfall (ml)	Average wet bulb temperature	Maximum air temperature	Relative humidity (%)
Rainfall (ml)	1.0000			
Average wet bulb temperature	0.3180	1.0000		
Maximum air temperature	–0.3268	0.2185	1.0000	
Relative humidity (%)	0.4782	0.8059	–0.3531	1.0000

Note: Daily NASA reanalysis data from Darwin, January 2020 to December 2023



Source: NASA January 2020 to December 2023

Chapter 4: Womb to Classroom: the lasting effects of prenatal heat exposure on health and cognitive outcomes

Abstract

Much evidence finds that prenatal exposure to hot weather causes poorer health at birth, but we know little about longer-term effects on children's development. In this study, we investigate the effects of prenatal heat exposure on health and education outcomes in Australia's Northern Territory for children aged 0-12, using linked administrative data. We find that prenatal exposure to unusually warm conditions leads to lower school test scores at ages 8, 10, and 12, and increases the likelihood of hospital admissions in early childhood. Importantly, we find that these impacts persist even when controlling for health at birth, and even among children with no health concerns at birth. Our results have important implications for understanding the consequences of climate change on human capital formation.

4.1 Introduction

A large body of evidence shows that exposure to hot weather in utero affects children's health at birth (Chersich et al. 2020). A smaller, but growing, body of research suggests that the effects of heat exposure in utero can persist into adulthood (Brink et al. 2024). There is evidence from China, for example, that heat exposure in utero can affect physical growth, resulting in shorter adult height (Hu and Li 2019). There is also evidence from China and the USA of impacts on cognitive and economic outcomes, with prenatal heat exposure resulting in lower adult education attainment, and lower earnings (Hu and Li 2019; Isen, Rossin-Slater, and Walker 2017).

Previous research shows that anthropogenetic climate change will make hot and very hot weather more frequent, both in locations which currently have hot climates, and in locations with milder climates (e.g. see the latest IPCC climate change synthesis report (Calvin et al. 2023)). If hot weather experienced in utero leads to poorer health and cognitive development, and the cohorts of children born in the coming decades will be more exposed to these conditions, then, without action to prevent these effects, we may expect population-level reductions in both wellbeing and productivity.

In order to prevent such effects, we must understand more about how they come about. Two questions are particularly important: first, are the effects of prenatal heat exposure, beyond its impact on health at birth, evident in childhood? Second, what is the link between the impact of heat exposure on birth outcomes and on outcomes measured later in childhood? In this chapter we focus on the first question, and provide suggestive evidence on the second.

To do this, we analyse the impacts of in utero heat exposure on childhood health and education outcomes, for children growing up in the Northern Territory (NT) of Australia. We use longitudinal linked administrative data, which includes hospital admissions and school test scores for children from birth through to age 12. We isolate the element of heat exposure in pregnancy that is exogenous: that is, variation within a given location and time of year. The strengths of this approach to causal identification is discussed by Dell, Jones and Olken (2014) in their review of research on the impacts or weather on economic outcomes.

The NT is an ideal context in which to answer our questions because of its unique combination of climate and data availability. The NT spans tropical and arid climate zones, with high average temperatures, and extreme heat in the summer/wet season. This means that, unlike studies from other, milder climates, where the highest temperatures analysed are of 32 degrees Celsius (Hu and Li 2019; Isen, Rossin-Slater, and Walker 2017), we can separately estimate the effects of extreme heat conditions that are most likely to cause heat stress (e.g. temperatures of 35-40 and 40+).

Analysis of the impacts of heat conditions typical of tropical and arid climates is also inherently important, as these climate zones are home to around half of the world's population (Doyle 2023). Unlike other regions with similar climates, the NT has comprehensive linked administrative data, covering birth outcomes, hospital admissions and standardised school test scores. This allows us to conduct detailed analysis on multiple outcomes – and the links between them – which is not possible in most other regions with similar climates.

We find that in utero heat exposure worsens both health and education outcomes in childhood. We find suggestive evidence that children exposed to unusually warm conditions in utero are more likely to be admitted to hospital in early childhood, and strong evidence that they score lower on standardised tests at ages 8, 10 and 12. These effects are small relative to other risks, but they are meaningful at the population level, given that heat exposure is ubiquitous.

We do not find any impact of prenatal heat exposure on total hospital admissions to age 12, but in our exploratory analysis, we find an increase in admissions in early childhood, which fades after age 3. Effects are largest for exposure to extreme heat – above 40 degrees – in the third trimester of pregnancy. These effects are not explained by any single condition or diagnosis. Instead, we see an increase in the average number of diagnoses per admission. We interpret this to mean that health conditions caused by heat exposure are typically not acute enough to require hospital admission, but may accentuate the risks children face from other illnesses – making hospital admission more likely.

However, while the impacts on hospital admission fade over time, the impacts on cognitive outcomes are persistent. We see statistically significant effects on scores for

reading, writing, grammar, spelling and numeracy, all of similar magnitude, at ages 8, 10 and 12.

Furthermore, our analysis reveals the interactions – or lack thereof – between health and educational outcomes. We find that it is not the case that children follow a linear pathway from prenatal heat exposure to poorer health at birth, to poorer health in childhood, to lower school test scores. Instead, we see impacts of heat exposure on hospital admissions and school test scores even for children with no health concerns at birth.

This chapter makes two main contributions to our understanding of how prenatal heat exposure affects children's development.

First, we add to the evidence base on how children are affected by prenatal heat exposure. There is a lot of evidence on the impacts of heat exposure on health at birth, and a small body of evidence on adult outcomes. But there is very little evidence on whether and how effects may manifest in early and middle childhood. We find effects on health in early childhood, and on cognitive outcomes through to age 12. Furthermore, we find these effects in a region with a hot climate; this suggests that standard adaptations like air conditioning – available in most households in the NT – do not fully attenuate the effects of heat exposure.[35]

Second, our analysis helps shed light on the mechanisms through which heat exposure in utero affects child development. Importantly, we find that the effects of heat exposure in utero on childhood health and cognitive outcomes are not explained by its impacts on health at birth. This lends support to the theories put forward recently, suggesting that there are many mechanisms through which heat exposure affects fetal development (Brink et al. 2024; HIGH Horizons and World Health Organization 2024): our findings suggest that the mechanisms that lead to poorer health at birth appear to be distinct from those that health and cognitive outcomes later in childhood.

Furthermore, our analysis has practical policy implications for the NT. We find that current levels of heat exposure during pregnancy are affecting health in early childhood,

[35] Data on air conditioning units are based on a survey which does not cover remote Aboriginal communities (Australian Bureau of Statistics 2014) – availability and use of air conditioning is substantially lower in Aboriginal communities.

and school performance. This not only affects children's wellbeing and life chances, but also leads to costs to the public schooling and healthcare systems. The effects are small on average for any individual, but exposure to at least some level of extreme heat in pregnancy is ubiquitous. Interventions that help to shield pregnant people from heat exposure, or that provide remedial support in early childhood, could prevent effects of heat exposure on population health and productivity in current and future birth cohorts.

The remainder of the chapter is structured as follows. In Section 2 we review prior research on the impacts of heat exposure during pregnancy, and set out potential causal pathways that may explain how this can affect children's development. Section 3 provides some context on the NT, describes the administrative data we use, and defines our measures and outcomes. In Section 4 we outline the empirical strategy that we use to estimate the causal effect of heat exposure. Section 5 presents the main results, while Section 6 presents further analysis on which pathways to long-term impact are best supported by the data. In Section 7 we discuss the implications of our findings and conclude.

4.2 Prior evidence and conceptual framework

4.2.1 Evidence on the impacts of heat exposure in pregnancy

Impacts at birth

A large body of evidence tells us that exposure to heat in utero can affect fetal development (Beltran, Wu, and Laurent 2014; Yunquan Zhang, Yu, and Wang 2017; Chersich et al. 2020). This results in higher risk of stillbirth and preterm birth, and lower average birthweight. Heat exposure has also been associated with certain pregnancy complications like preeclampsia and high blood pressure (Algert et al. 2010; Shashar et al. 2020; Beltran, Wu, and Laurent 2014).

From a separate body of research, we know that children with poorer health at birth have, on average, poorer health in childhood, and they perform worse at school than their peers – though the strength of this association has decreased over time (Goisis, Özcan, and Myrskylä 2017; Bharadwaj, Lundborg, and Rooth 2018).

But does this mean that children exposed to hotter temperatures in utero will face the same risks as children who have poor health at birth for other reasons? Not necessarily. Standard measures of health at birth like preterm birth and low birthweight can result from a wide range of underlying conditions (Romero, Dey, and Fisher 2014). The mechanisms linking heat exposure to birth outcomes may be different from those linking other risks to birth outcomes, and this may mean that the long-term consequences will be different.

Long-term impacts

Brink et al.'s (2024) recent systematic review shows that, of the small number of studies analysing the long-term effects of prenatal heat exposure, most point to adverse effects. Studies have analysed impacts on physical health, mental health, and economic outcomes. For instance, Isen et al. (2017) find that in the USA, each additional hot day in utero or in the first year of life reduced adult earnings by 0.1%; in Ecuador, Fishman et al. (2019) find each 1 degree increase in average prenatal temperatures leads to a 0.7%-1.1% reduction in earnings. In China, Hu and Li (2019) find that an additional day of temperatures above 29 degrees in pregnancy led to 0.02 fewer years of completed schooling.

These findings suggest that prenatal heat exposure can have substantial impacts on people's lives and wellbeing. This is concerning because, in many parts of the world, at least some exposure to these levels of heat during the 9 months of pregnancy is unavoidable. Therefore, while the effects are small on average for any individual, the implied effects on population health and labour productivity are considerable.

To better understand these risks and how to mitigate them, we need to know how these effects come about: what are the mechanisms linking prenatal heat exposure to economic outcomes in adulthood? To answer this, it is helpful to understand whether and how the effects of prenatal heat exposure manifest in childhood.

The main evidence we have on outcomes measured in childhood is stunting: prenatal heat exposure has been shown to increase the risk of stunting in China (Ai and Tan 2024; Hu and Li 2019), Ethiopia (Randell, Gray, and Grace 2020), and across Sub-Saharan Africa (Bratti, Frimpong, and Russo 2021). Ai and Tan (2023) also show that prenatal heat

exposure worsens mental health in adolescence. In their review, Brink et al. (2024) highlight the need for further research into other outcomes that may be affected by prenatal heat exposure, and the causal pathways leading to these outcomes. In this study, we help to fill this gap by analysing additional outcomes – childhood hospital admissions and school test scores – and assessing the evidence for various high-level causal pathways.

4.2.2 Heat exposure and the fetal origins hypothesis

How could heat exposure in utero affect education attainment and earnings decades later? At a high level, this evidence fits within the broader body of research showing that the conditions that children are exposed to before birth can have long-term impacts. This is referred to as the 'fetal origins hypothesis' (as proposed by Barker (1998)), and evidence shows that a broad range of factors experienced in utero, including pollution, parents' economic resources, and even the seasonal flu, can have long-term impacts on children's development (Almond 2006; Almond, Currie, and Duque 2018; Schwandt 2019; Martin-Bassols et al. 2024). The consensus is that these occur through epigenetic effects – that is, certain shocks, when experienced during a critical period in development, affect gene expression – and this 'programs' long-term impacts on children's development. Experimental animal studies confirm that prenatal heat exposure can indeed affect gene expression (e.g. Skibiel et al. (2018) for cattle).

While the fetal origins hypothesis provides a helpful framework, alone it tells us little about how a given shock – like heat exposure – can cause lower earnings years later. We face two particular challenges in understanding these mechanisms.

Challenge 1: what does heat exposure do to fetal development?

A first challenge is in articulating what it is about heat exposure in pregnancy that affects the fetus. There are likely a range of mechanisms at play (Brink et al. 2024), and their relative importance may vary across contexts. It may be, for instance, that exposure to extreme temperatures leads to maternal heat stress. This may cause congenital anomalies (Auger et al. 2017), or could affect fetal development in more subtle ways. It may be that exposure to moderate heat leads to maternal heat strain (i.e. excess sweating and redirection of maternal blood flow away from the placenta), which, if experienced over a

prolonged period, may reduce the flow of nutrients to the fetus (Samuels et al. 2022). Alternatively, there may be indirect effects: heat exposure may change maternal behaviours like eating, sleeping, and exercise patterns during pregnancy, which may lead to epigenetic effects.

We cannot disentangle these channels in our analysis. But understanding that there may be multiple mechanisms at play, and that they are not necessarily mutually exclusive, helps to contextualise our analysis.

Challenge 2: what are the longer-term causal pathways?

A second challenge is in understanding the pathway from an early life shock to long-term outcomes. We consider two distinct sets of outcomes: hospital admissions and school test scores. The causal pathways linking heat exposure or one or both of these outcomes are complex, because children's outcomes are affected not only by the biological effects of the prenatal exposure itself, but by their parents' reactions to any early indicators of these effects (Almond, Currie, and Duque 2018).

Much of the research on the 'early origins' of adult human capital focuses on epigenetic effects, as described above (Almond, Currie, and Duque 2018). Such theories were first advanced to explain the link between poor nutrition in utero and chronic disease observed in adulthood (as proposed by Barker (1998)). However, more recent research has shown that many outcomes, in addition to chronic disease, can be affected (Almond, Currie, and Duque 2018).

An alternative explanation for how prenatal conditions could cause outcomes in childhood and adulthood is to consider a chain of events. It may be, for instance, that prenatal heat exposure affects health at birth, and health at birth affects health in infancy, and health in infancy affects children's ability to engage in activities that advance their cognitive development, and so on. Such patterns are sometimes referred to as 'developmental cascades' (Masten and Cicchetti 2010). In the economics literature, we may similarly think of a prenatal shock affecting the child's health 'stock' at birth and in early childhood, and that the impact of this initial setback may be self-reinforcing, fitting with Cunha and Heckman's (2007) theory of 'self-productivity'.

These two explanations are not necessarily mutually exclusive: it may be that both fetal programming and developmental cascades operate together (Masten and Cicchetti 2010). In this chapter, we cannot conclusively establish which of these pathways explains the impacts of heat exposure, though we can provide suggestive evidence. Because we have linked administrative data including birth records, we can analyse, for instance, whether any longer-term effects of heat exposure on cognitive outcomes are explained by birth outcomes, or by health in childhood (consistent with ideas of self-productivity or developmental cascades). If instead the effects on health and education are independent of each other, this would be more consistent with epigenetic effects as the primary causal pathway. We present exploratory analysis on this in Section 6.

4.3 Data

4.3.1 The Northern Territory

The Northern Territory (NT) is one of Australia's eight states and territories. It is a large and sparsely populated area, spanning tropical and arid climate zones. Around 60% of residents in the NT live in the capital city, Darwin, in the tropical north. A further 11% live in Alice Springs, which is the largest city in the arid southern region of the NT. Around one-third of NT residents identify as Aboriginal. Most Aboriginal people in the NT live outside of Darwin, often in remote and very remote communities.

Baseline heat exposure is high in the NT. Most of the time, maximum daily temperatures are between 25 and 35 degrees, and temperatures about 40 degrees are not the norm, but are not uncommon (Table 1).

Because of the NT's hot climate, use of air conditioning is high. In 1999 (just before the first children in our study cohort were born), 84 percent of households in the NT used air conditioning in their home (Australian Bureau of Statistics 2008), compared with 35 percent in Australia overall. By 2014, this had increased to 94 percent in the NT (Australian Bureau of Statistics 2014).[36]

[36] However, this survey covers mostly urban areas, and therefore is not representative of residents in remote Aboriginal communities.

4.3.2 The linked data

This study uses de-identified administrative data covering all children born in the NT between 2000 and 2017, linked with small-area daily weather data. These birth records come from the NT Perinatal Registry, and are linked with public hospital inpatient admissions and schools' standardised test scores.[37]

Hospital admissions

We have data on children's inpatient hospital admission in the NT from 2000 to mid-2017. We use this to construct our main health outcome measure: counts of the number times each child is admitted to hospital.[38] We also count the number of days each child spends in hospital, and construct an indicator for whether the child was ever admitted to hospital in the NT –70% of our cohort were admitted at least once by age 12, with an average of 2 admissions per child and 8.4 days spent in hospital (Table 1).

School test scores

In Australia, the National Assessment Program – Literacy and Numeracy (NAPLAN) test is administered each year to all children in Years 3 (age 8), 5 (age 10) and 7 (age 12).[39]

NAPLAN scores are used to benchmark children's progress against their cohort, and to benchmark schools' performance. NAPLAN tests skills that are considered essential to progressing in school, across 5 domains: reading, writing, grammar, spelling and

[37] There is single private hospital in the NT for which we do not hold linked admission data. However there are few, low acuity, admissions of children to the private hospital. All admissions to the NT's 5 public hospitals are included. About 70% of school children attend government schools, for which we have linked enrolment and attendance data. Test scores data are also linked for most non-government schools, including Catholic schools – meaning they cover over 90% of children.

[38] We exclude the birth admission from our count, as this is covered in a separate indicator: admission to a special care nursery. We also exclude 'statistical' discharges (e.g. when an individual is discharged from hospital and immediately re-admitted) to avoid double counting a single episode of care. In our count variables, we top-code number of admissions and days in hospital for children above the 97.5th percentile, to avoid the influence of extreme outliers.

[39] While the test is also administered at age 14 (Year 9), we do not analyse these results as only a small number of children in our cohort reach age 14 within our period of analysis. Within our analysis data, a small number of children had multiple records for a single test – e.g. a child who appears in the test records for the age 8 test in two consecutive years. In most cases, the child withdrew from the test or was absent in one of the instances, and therefore we take the results from the year they sat the test. Of the small number of children who sat the test on multiple occasions, we take results from the second sitting.

numeracy. Raw scores are converted into 'proficiency bands', on a scale of 1 to 10.[40] In our main results, we report a simple average of the band scores across all five domains, though our detailed results show impacts on each domain.

Weather data

We geo-link the mother's place or residence from the administrative birth records to NASA's daily weather reanalysis data, measured at intervals of 0.5×0.625 degrees of latitude and longitude (roughly 50×55km).

As our main measure of heat exposure, we construct counts of the number of days with maximum temperatures under 20 degrees, 20 to 29.99 degrees, 30 to 34.99 degrees, and 35 degrees or above. We measure heat exposure in this way because, as Dell, Jones and Olken (2014) explain, this approach allows for non-linear effects of increased temperatures. This is particularly important given the possibility of multiple mechanisms linking heat exposure to development in utero: for instance, maternal heat strain, heat stress, and behavioural changes in response to heat, may be caused by different levels of heat exposure (see Chapter 3.4).

We construct these categories for each trimester of pregnancy, as is standard in the literature on prenatal exposures. We know from prior research that heat exposure's effects may differ based on the timing of exposure; for instance, the placenta develops in the first trimester, and heat exposure in the first trimester may increase the risk of abnormal development of the placenta leading to complications like pre-eclampsia (Part et al. 2022). In the third trimester, heat exposure has been shown to bring forward labour (Barreca and Schaller 2019).

Our categories of heat exposure are coarse. We use them to summarise the impacts of heat exposure across multiple outcome measures. In our detailed results, we analyse more granular categories of maximum temperature exposure (under 20, 20–24.99, 25–29.99, 30–34.99, 35–39.99, 40 and above) as well as daily minimum temperatures (under 5, 5–9.99, 10–14.99, 15–19.99, 20–24.99, 25+).

[40] These same bands apply across all year levels, but the 'cutoff' point for meeting the minimum standard increases with age. In year 3, children scoring 3 or above on the 10 point scale are meeting the national minimum standard. In year 5, it is 5 or above, in year 7, it is 6 or above, and in year 9, 7 or above.

	Mean	SD	Min	Max	N
Health outcomes					
Birth weight [a]	3272.3	525.8	2160.0	4180.0	23756
Hospital admissions ages 0–2	1.1	1.7	0.0	8.0	23756
Hospital admissions ages 3–7	0.0	0.1	0.0	2.0	23756
Hospital admissions ages 8–12	0.0	0.3	0.0	7.0	11194
Hospital admissions ages 0–12	2.0	2.5	0.0	12.0	11194
Days in hospital ages 0–2	5.6	11.6	0.0	67.0	23756
Days in hospital ages 0–12	8.4	14.9	0.0	83.0	11194
Any admissions ages 0–2	0.5	0.5	0.0	1.0	23756
Any admissions ages 0–12	0.7	0.5	0.0	1.0	11194
Education outcomes – NAPLAN test					
Average score – year 3 (age 8)	2.8	1.4	1.0	6.0	11681
Average score – year 5 (age 10)	4.5	1.3	3.0	8.0	8013
Average score – year 7 (age 12)	5.5	1.3	4.0	9.0	4068
Heat exposure – 1st trimester					
Max under 20°	1.7	6.4	0.0	58.0	23756
Max 20–25°	3.7	9.3	0.0	69.0	23756
Max 25–30°	26.2	22.9	0.0	84.0	23756
Max 30–35°	38.5	23.6	0.0	84.0	23756
Max 35–40°	11.3	17.8	0.0	81.0	23756
Max 40°+	2.6	7.3	0.0	67.0	23756
Heat exposure – 2nd trimester					
Max under 20°	1.5	6.1	0.0	55.0	23756
Max 20–25°	3.6	9.4	0.0	70.0	23756
Max 25–30°	29.2	25.0	0.0	98.0	23756
Max 30–35°	45.8	25.8	0.0	98.0	23756
Max 35–40°	14.4	20.7	0.0	92.0	23756
Max 40°+	3.5	8.7	0.0	70.0	23756
Heat exposure – 3rd trimester					
Max under 20°	1.5	5.9	0.0	53.0	23756
Max 20–25°	3.3	9.1	0.0	69.0	23756
Max 25–30°	28.1	23.6	0.0	91.0	23756
Max 30–35°	42.6	24.1	0.0	91.0	23756
Max 35–40°	12.4	18.9	0.0	85.0	23756
Max 40°+	3.1	8.0	0.0	69.0	23756

Table 1: Summary statistics

Source: NASA and Analysis dataset, average outcomes and heat exposure for all babies conceived in the NT from March 2000 to September 2009.

a: variable is top- and bottom-coded at the 2.5th and 97.5th percentiles, to reduce the influence of extreme outliers

Beyond ambient temperature, the heat a person experiences is also determined by humidity, as well as other environmental factors. However, in Chapter 3 we find that empirically in the NT, heat exposure metrics that include humidity have lower explanatory power. This is consistent with findings from other contexts (Baldwin et al. 2023). We therefore focus on temperature alone as our measure of heat exposure in this research.

4.3.3 Sample selection

We limit our analysis to children conceived from March 2000 to September 2009.[41] This gives us a cohort for whom all children will be age 7, and will therefore be in school, by mid-2017 (when our hospital admission data ends). There were 34,259 children born in this period.

To ensure that our findings are not affected by attrition due to migration out of the NT, we conduct our analysis only on children who appear at least once in the school records, meaning that they were present in the NT at birth, and (at least) at school starting age. We explain the reasons for this in more detail in Appendix A. This reduces our analysis sample to 25,416. In our analysis on hospital admissions at older ages (ages 8-12), this falls to 11,194, as only the older children in our cohort reach age 12 by 2017, when our admissions data ends.

For our analysis of school test scores, our cohort is reduced further, both due to narrower ranges of data availability and incomplete participation rates. Standardised testing began in 2008, and therefore the earliest-born children in our cohort were too old to sit the age 8 test by 2008. Data are only available through to 2016, at which point not all children in our birth cohort will have reached age 8, and even fewer will have reached ages 10 and 12. In total, 11,881 children in our cohort sat the test at age 8, 8,013 sat the test at age 10 and 4,068 sat the test at age 12.

[41] We define our cohort based on date of conception instead of date of birth because definitions based on date of birth will systematically exclude children born preterm at the beginning of the sample, and exclude those born late at the end of the sample period – this could lead to bias, especially when exposures contributing to gestational age are seasonal. We determine date of conception by subtracting gestational age (in weeks) from birthdate.

The implications of these various windows of data availability and rates of attrition is that our analysis for different outcomes is based on slightly different populations. We retain the largest possible cohort for each outcome to maximise our statistical power. An alternative approach would be to base our analysis on a narrower birth cohort who appear in all outcomes: we do this as a sensitivity test for the 2,862 children who appear in all outcomes through to age 12 (Appendix Table A1), and on hospital records for the 11,194 children for whom we have hospital data to age 12 (Appendix Table A2). We find estimates of similar magnitude to our main results, though with larger standard errors.

4.4 Methods

4.4.1 Causal identification strategy

Our goal in this analysis is to estimate the causal effect of in utero heat exposure on health and education outcomes. We may find correlations between heat exposure and later outcomes, but to interpret these correlations as causal effects, we must be confident that they are not driven by reverse causality or omitted variable bias.

Within our context, we see no risk of reverse causality: it is not plausible, for example, that school test scores at age 12 could affect a child's heat exposure in utero. However, there is a risk of omitted variable bias, as heat exposure may correlate with other determinants of health and education outcomes. This could come from two sources.

First, it may be due to location. Children in communities that experience higher temperatures may also face other risks to their health and development. If not controlled for in our analysis, this may lead us to believe that it is heat exposure affecting children's outcomes, when in fact it may be some other factor, such as access to healthcare and fresh food – or it may be socioeconomic factors leading more disadvantaged families to live in places with warmer climates. This is particularly important to consider in the NT, with the majority of Aboriginal families living in remote or very remote communities – many in central NT (where extreme temperatures are more common), and most non-Aboriginal residents living in the capital city, Darwin.

Second, it may be due to the time of year that a child is in utero. Children conceived in the hottest months of the year will experience the warmest weather in their first trimester, while those conceived at the end of the summer/wet season could avoid exposure to the hottest months of the year in utero entirely. At face value, these are precisely the types of differences we want to measure. However, timing of conception is not entirely random. As Currie and Schwandt (2013) find in the USA, parents from different demographic groups may choose to conceive at different times of year. In addition, there are other seasonal risks that affect the prenatal environment, like economic conditions and the spread of seasonal diseases (Doyle 2023). For our purposes, if children exposed to hot weather in utero experience poorer health and education outcomes, we want to isolate the portion of this association that is because of the heat exposure itself, and not due to other factors that are seasonally correlated with heat exposure.

To avoid these two sources of omitted variable bias, our causal identification strategy is to isolate the element of heat exposure that is exogenous to location, seasonal conception and other seasonal risks: we therefore estimate the impact of weather 'shocks'. These are deviations from usual conditions, within a given community and time of year. This approach, and its strengths in causal identification, is discussed in detail by Dell et al. (2014).

The basic idea of this identification strategy is that locations each have their own climates, which determine the distribution of weather conditions experienced in each season. But temperatures experienced day-to-day within a given month are effectively random draws from that climate's distribution. Therefore, an estimation approach that exploits these day-to-day deviations can estimate the impact of those weather conditions, abstracting from location and seasonal effects.

4.4.2 Regression models

Health outcomes

We operationalise our estimation strategy with the following regression model:

$$H_{ijmy} = \alpha + \sum_{k}^{3}(\sigma_1^k maxunder20 + \sigma_2^k max30to35 + \sigma_3^k max35plus) + \beta X_i + \gamma_{my} + \theta_{mij} + \varepsilon_{itj} \quad (1)$$

Here H_{ijmy} represents health outcomes for child *i,* born to a mother living in location *j*, and conceived in month *m*, year *y*. In our main analysis, the outcome we focus on is number of hospital admissions, though our detailed analysis also presents estimates on days spent in hospital, an indicator of whether ever admitted to hospital, and the number of admissions for different diagnoses. We also present estimates on birthweight, as a means of comparing between heat exposure's impacts at birth and later in childhood.

The heat exposure variables are counts of the number of days within each of the *k*=3 trimesters, when maximum daily temperatures were under 20 degrees Celsius, between 30 and 34.99 degrees, or 35 degrees and over. The reference category is days with maximum temperatures between 20 and 29.99 degrees. The coefficients on these variables, σ_1^k, $\sigma_2^k$and σ_3^k are our main parameters of interest.

In our main results table, we adjust for multiple hypothesis testing. We do this because we are estimating coefficients on 3 categories of heat exposure (under 20, 30–34.99 and 35+) across three trimesters. This means we are effectively testing 9 hypotheses within each regression. We are also analysing the same outcomes at different ages. The large number of hypotheses we are testing makes it more likely that we may, by chance, find estimates that appear statistically significant, but which are not true effects. To account for this, we indicate statistical significance in our main results table (Table 2) using Benjamini, Krieger, and Yekutieli's (2006) sharpened Q-values, estimated using Stata code provided by Anderson (2008). These sharpened-Q values account for the false discovery rate, and can be interpreted as adjusted p-values. We report unadjusted p-values in our additional analyses because these analyses test distinct hypotheses across varying subsamples, specifications and outcome measures, making joint false discovery rate adjustment inappropriate.

X is a vector of individual-level controls including the mother's age at birth, an indicator for whether it is the mother's first pregnancy, and whether either the child or the mother is Aboriginal.

Our fixed effects γ_{my} and θ_{mij} control for the month-year of conception, and the baby's month of conception, sex and location (interacted), respectively. Location is based on the mother's place of residence at the time of birth.[42] Our use of interacted fixed effects follows prior studies on this topic, for instance by Cil and Kim (2022), and Isen, Rossin-Slater and Walker (2017). We interact location fixed effects with month and sex to allow for different seasonal patterns in birth outcome across locations. This is a particular concern in the NT because some communities experience regular flooding in the wet season, which cuts off road access and hence access to fresh food and other supplies: their experience of the wet season may be very different from towns or communities that do not flood. Furthermore, as DiPietro and Voegtline (2017) show in their review of sex differences in early development, different types of exposures may affect male and female fetuses differently. Flexibly interacting these location- and season-specific unobserved effects with sex allows for this possibility. We also interact month and year to allow both for NT-wide month-to-month seasonal differences, and for time trends over the 9 ½ years of birth cohorts.

Standard errors are clustered by location.

Education outcomes

In analysing educational outcomes, we specify our model slightly differently, to allow for the disadvantages faced by children who are the youngest in their class (e.g. see Peña (2017)). Our model already includes fixed effects for month of conception, which should absorb most of the differences in performance based on relative age within the class. However, we want to ensure that any effects we find are not driven by remaining correlations between heat exposure and relative age.

[42] We have used cluster analysis to create groups from the suburb, community or locality that the mother reported in her hospital records. We group these together with maximum distance between each point within a cluster of 50 miles. We know that families in the NT have relatively high rates of intrastate mobility and therefore children may move throughout childhood (Su, Guthridge, and He 2023), but this does not impact our measure of exposure, which is based only on location during pregnancy.

In the NT, the cutoff for entry into the school year is 1 July; therefore, children born in April, May and June will be the youngest in their year. They will also have been in utero during the summer/wet season, when temperatures above 35 degrees are more common.

To address this, we create an indicator 'Youngest in year', for children born in April, May or June, and interact this with heat exposure of 35 degrees or more; the model is otherwise the same as in Equation (1).

$$
\begin{aligned}
Y_{ijmy} = \alpha + \sum_{k}^{3} \big(& \sigma_1^k maxunder20 + \sigma_2^k max30to35 \\
& + \sigma_3^k max35plus \\
& + \sigma_4^k max35plus \times YoungestinYear \big) \\
& + \eta YoungestinYear_i + \beta X_i + \gamma_{my} + \theta_{mj} + \varepsilon_{itj}
\end{aligned}
\tag{2}
$$

In this specification, as in Equation (1), our main coefficients of interest are σ_1^k, σ_2^k and σ_3^k. We report σ_4^k (the interaction terms) in Appendix Table B5, but not in our main reporting of results.

4.5 Results

Table 2 summarises our estimates of the impact of prenatal heat exposure on birthweight, hospital admissions in childhood, and school test scores. The table shows estimates and standard errors, and the asterisks indicate statistical significance after adjusting for multiple hypothesis testing (see Appendix Table B1 for the sharpened-Q values).

Our main finding is that prenatal heat exposure negatively impacts health and education, though there are differences in the level and timing of heat exposure that is important for each outcome. The most precisely estimated effects are for school test scores. In our main analysis we find no impact on total hospital admissions at ages 0 to 12, though in our additional analysis, prenatal heat exposure appears to increase admissions in early childhood, but these effects fade over time.

	Birth weight (grams) (1)	N. hospital admissions (age 0-12) (2)	NAPLAN scores Year 3 (age 8) (3)	Year 5 (age 10) (4)	Year 7 (age 12) (5)
1st trimester (reference category: 20-30 degrees)					
Max <20°	-0.688	-0.004	0.004	0.002	-0.002
	(1.453)	(0.007)	(0.003)	(0.002)	(0.003)
Max 30°-35°	-0.665**	0.002	-0.002***	-0.003***	-0.001
	(0.230)	(0.002)	(0.001)	(0.001)	(0.001)
Max 35°+	-0.428	0.001	-0.003*	-0.005***	-0.000
	(0.280)	(0.004)	(0.001)	(0.001)	(0.002)
2nd trimester (reference category: 20-30 degrees)					
Max <20°	1.608	-0.008	-0.003	0.002	0.002
	(1.837)	(0.014)	(0.003)	(0.005)	(0.007)
Max 30°-35°	-0.213	-0.000	-0.002	-0.001	-0.001
	(0.250)	(0.002)	(0.001)	(0.001)	(0.002)
Max 35°+	0.507	-0.001	-0.003	-0.003**	-0.004***
	(0.280)	(0.003)	(0.002)	(0.001)	(0.001)
3rd trimester (reference category: 20-30 degrees)					
Max <20°	4.800***	-0.009	-0.001	0.007**	-0.003
	(1.258)	(0.012)	(0.004)	(0.003)	(0.006)
Max 30°-35°	-0.665	0.004	-0.001**	-0.002***	-0.002**
	(0.335)	(0.002)	(0.000)	(0.001)	(0.001)
Max 35°+	-0.943	0.008	-0.004***	-0.003***	-0.005***
	(0.542)	(0.004)	(0.001)	(0.001)	(0.001)
Constant	3,363.215	1.385	3.150	4.907	5.864
	(48.046)	(0.270)	(0.110)	(0.092)	(0.083)
Demographic controls	Yes	Yes	Yes	Yes	Yes
Month-year FE	Yes	Yes	Yes	Yes	Yes
Month-loc-sex FE	Yes	Yes	Yes	Yes	Yes
Control for youngest in class	No	No	Yes	Yes	Yes
Observations	23,756	11,194	11,681	8,013	4,068
R-sq	0.137	0.369	0.522	0.540	0.552

Table 2: Main regression estimates

Notes: This table shows regression estimates and cluster-robust standard errors in parentheses. *, ** and *** denote Q-statistics for statistical significance at the 1%, 5% and 10% levels, after adjusting for false detection rate with multiple hypothesis testing. Adjustment is done separately for each set of outcomes (birthweight, hospital admissions and school test scores). Each column represents a different regression, with the outcome specified in the first row. Regressions with birth and hospital admission outcomes are estimated with the model specified in equation (1) while the regressions for the NAPLAN test scores outcomes are estimated with the model specified in equation (2). As specified in equations (1) and (2), the regressions also include covariates (mother's age, Aboriginal status, whether mother's first pregnancy), month-year fixed effects and location-month-sex fixed effects (these are absorbed using the Stata 'areg', which affects the intercept but not the coefficients of interest).

Through a series of robustness tests, we find that our conclusions are robust to different ways of specifying fixed effects, different ways of controlling for relative age in class, and are not driven by a specific cohort or event. These are detailed in Appendix B.

4.5.1 Birthweight

Column 1 of Table 2 shows the impacts of prenatal heat exposure on birthweight. After adjusting for multiple hypothesis testing, we find that exposure to temperatures of 30–34.99 degrees in the first trimester (relative to 20–29.99 degrees) reduces birthweight by around 0.7 grams per day of exposure, on average. In addition, third trimester exposure to cooler temperatures (days with maximum temperatures under 20 degrees) is beneficial, increasing birthweight by around 5 grams per day of exposure. There is no statistically significant impact of any level of exposure in the second trimester.

4.5.2 Hospital admissions

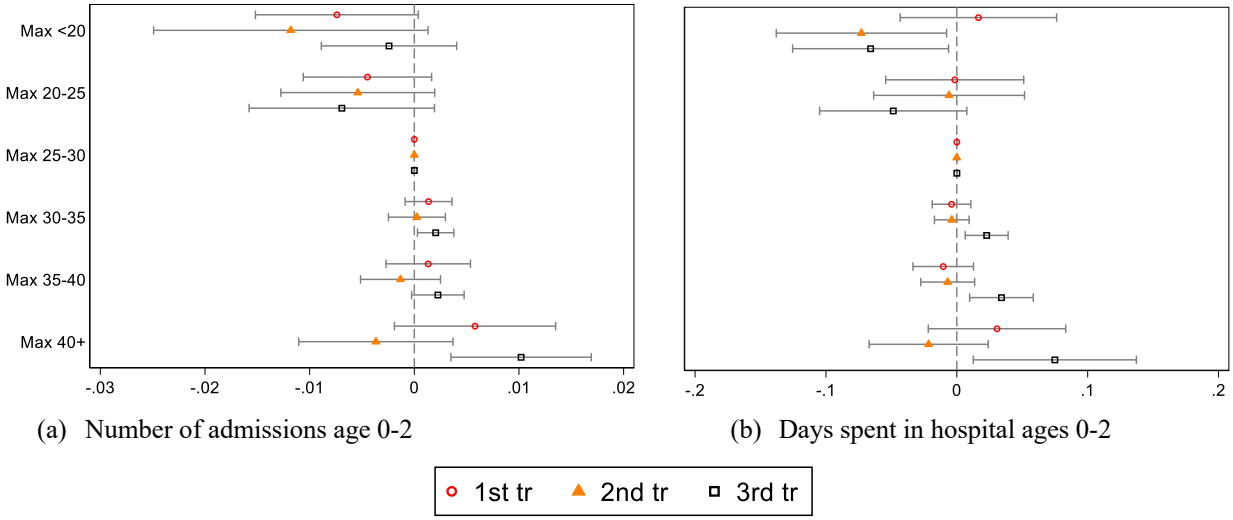
Column 2 of Table 2 show the impact of prenatal heat exposure on total hospital admissions from age 0 through to age 12. We estimate that children have 0.004 and 0.008 additional admissions from each additional day of hot (30-34.99 degrees) and very hot (35+) temperatures in the third trimester, respectively. However, these estimates are not statistically significant after adjusting for multiple hypothesis testing. The estimates of the impact of exposure in the first and second trimesters are substantially smaller and are also not statistically significant.

4.5.3 By age and level of exposure

While we do not see any significant impact of heat exposure on total admissions to age 12, we do see some effects early in childhood. Appendix Tables C1 and C2 show estimates based on more granular bands of heat exposure (daily maximums of under 20, 20–24.99, 25–29.99, 30–34.99, 35–39.99, 40+) and split by age (ages 0–2, 3–7 and 8–12). We see larger estimates of extreme heat exposure on hospital admissions for young children (aged 0–2). An additional day with maximum temperatures of 40 degrees or over in the third trimester increases average admissions at ages 0–2 by 0.01 (Appendix Table C2). We see no impact of heat exposure in the first or second trimesters, and no impact of heat exposure on admissions among older children. Therefore, it appears that extreme

heat increases the risk of hospital admission in early childhood, but these effects fade over time.[43]

Figure 2 presents our estimates for admissions at ages 0–2 with the more granular bands of heat exposure. It shows a clear pattern: the estimated effect of heat increases with the level of exposure, with largest estimates for exposure to temperatures of 40 degrees or above. This pattern comes out more clearly for exposure in the third trimester, and for days spent in hospital (Figure 2(b)) than on number of admissions (Figure 2(a)). In addition, exposure to cooler temperatures in the second and third trimesters – of under 20 degrees – reduces days spent in hospital, suggesting that even relatively mild heat (25–29.99 degrees, our reference category here) may have small adverse effects.



(a) Number of admissions age 0-2

(b) Days spent in hospital ages 0-2

Figure 2: Estimated impact of heat exposure on hospital admissions – regression coefficients

Note: This figure shows estimates from regressions of the impact of additional days during each trimester of pregnancy of temperatures in the indicated range, on number of admissions and number of days spent in hospital, at ages 0 to 2 (up until the child's third birthday). This is based on equation (1) above. Appendix Table C2 presents these estimates.

By diagnosis

We see clearer effects in Figure 2 on days spent in hospital than on number of admissions. In addition, we do not find any impact of heat exposure on a child's probability of ever being admitted to hospital (Appendix Table C2). This may imply that if heat exposure

[43] The estimates of each age group do not sum to the 0-12 total because we use a narrower birth cohort for older age groups, to maximise the sample size. However, we show in Appendix Tables A1 and A2 that the same patterns hold, albeit with less precise estimates, when we limit our analysis to a common sample.

does affect admissions, it does so through accentuating other conditions – leading to additional hospital treatment or observation – more so than causing new illnesses require a child who is otherwise healthy to seek acute care. To explore this further and test whether there are specific conditions caused by heat exposure, we analyse admissions by primary diagnosis.

We do not find any single diagnosis type that explains the overall effect on admissions (see Appendix Table D1). There are statistically significant effects of third trimester heat exposure on some diagnoses (accidents and 'factors influencing health status'),[44] but these effect sizes are very small. However, we do find that the number of diagnoses per admission is higher for children exposed to higher temperatures in the third trimester.

Together, these estimates suggest that while heat exposure does not lead to any specific condition requiring hospitalisation, it may accentuate symptoms from other conditions – some of which are not themselves caused by heat exposure, and can generally be treated at home – to become severe enough to require hospital admission.

4.5.3 School test scores

Columns 3 to 5 of Table 2 show the impact of heat exposure on school test scores at ages 8, 10 and 12. We see that children exposed to additional days of 30 degrees and over in utero have lower scores across all three tests, with larger effects for exposure of 35 degrees and over. Unlike our findings for hospital admissions, we see no additional impact of exposure to temperatures above 40 degrees, compared with temperatures of 35 or over (see Appendix E).

The estimates shown in Table 2 are averages across 5 domains of the test: reading, writing, grammar, spelling and numeracy. The effects of heat exposure are statistically significant, and of similar magnitude, across all five domains (Appendix F).

[44] 'Chapter Z of ICD-10, titled 'Factors influencing health status and contact with health services' contains admissions for observation, examination and screening, including for suspected concerns that were ruled out. Some of these admissions may be endogenous – i.e. warm weather around the time of birth may lead doctors to admit newborns as a precaution when there is no specific diagnosis, and this would be recorded in Chapter Z. However, we find no impact on 'Z03.7 Observation and evaluation of newborn for suspected condition', suggesting that this is not a driver of admissions.

In contrast to the pattern of results across hospital admissions, these effects do not fade over time: the impacts of heat exposure in the third trimester remain of similar magnitude and statistically significant at the 1% level across all three tests, after adjusting for multiple hypothesis testing. There are some differences in the timing of exposure (first trimester exposure has statistically significant impacts on scores at ages 8 and 10, but not age 12; second trimester exposure affects scores at ages 10 and 12, but not age 8). However, we do not read anything into these differences, because the coefficients are not significantly different from each other, in an F-test of equivalence of coefficients across all trimesters (see Appendix Table B8).

We may be concerned that heat exposure in utero would affect participation in the test. While all children are expected to sit the test, some children either withdraw, or are absent – e.g. due to illness. In the NT, around 25 percent of children do not sit the test. If it is the children most affected by heat exposure who do not attend – and if they did attend they would score lower than average – then this may mean our results are biased towards zero, and the true impact of heat exposure may be larger. However, this is not the case: Appendix G shows estimates of the impact of heat exposure on completing the full test. Our estimates are close to zero across all three tests.

4.5.4 Heterogeneous effects

The NT is a large geographical area, spanning tropical and arid climate zones, and with a large Aboriginal population. We next test whether the aggregate effects we find are evident across both climate zones, and among Aboriginal and non-Aboriginal children.

Table 3 shows that the impact of heat exposure is greater for non-Aboriginal children than for Aboriginal children, and for children living in arid climate zones, compared with those living in the tropical north of the NT (Table 3).

At face value, the smaller relative impact of heat exposure for Aboriginal children may suggest that Aboriginal children are less affected by heat exposure than non-Aboriginal children. This fits with findings from Quilty et al. (2023), which suggest that Aboriginal peoples in the NT maintain social and cultural practices that make them more resilient to heat exposure. But there could be two additional explanations. First, there is greater diversity among Aboriginal peoples, in terms of heat exposure, adaptations and

community characteristics. This may mean that heat exposure affects different populations in different ways, and we do not find a clear aggregate effect across all Aboriginal children. Second, there are more complex drivers of poor outcomes for Aboriginal children, including the additional burdens of food insecurity, poor housing, poverty, and lower levels of health and education literacy. Heat is one of many challenges contributing to these poor outcomes, whereas for non-Aboriginal children, heat may have a more isolated and singular influence. Note that this second explanation stands in contrast with the idea, outlined for instance by Aquino, Brand and Torche (2022), that children already facing socioeconomic disadvantage are less resilient to disruptive events, and hence may be expected to experience larger effects; we cannot disentangle these competing explanations in our analysis, though doing so would be a valuable avenue for further research.

Turning to differences by climate, we find that the effect of heat exposure is greater for children born in arid climate zones. For hospital admissions at ages 0-2, it appears that the effect of third trimester exposure that we find in the full population (as shown in Table 2) is almost entirely driven by children in arid climate zones; we see very little impact of third trimester heat exposure for children born in tropical climates. We do, however, see some impacts of milder heat exposure in the tropical climate zone in the first and second trimesters. For school test scores, the impacts of heat exposure are more precisely estimated for children born in tropical climate zones, but are larger for children born in arid climate zones.

In sum, we see some differences across groups in terms of the level and timing of heat exposure that affects them, and the magnitudes of these effects. But all groups – Aboriginal or not, living in tropical or arid climates – appear to experience some effects of prenatal heat exposure on both health and cognitive outcomes. These heterogeneous effects highlight the difficulty in pinpointing a specific period at which heat exposure is most risky. They make us more cautious in our interpretation of the apparently larger impact of third trimester exposure in the full population: exposure in the second and third trimesters may be equally, or more, important in some groups.

	Number of admissions ages 0-2				School test scores age 8			
	Tropical	Arid	Aboriginal	Non-Aboriginal	Tropical	Arid	Aboriginal	Non-Aboriginal
1st trimester (reference category: 20-30 degrees)								
	0.112	-0.003	-0.002	-0.010***	-0.001	-0.000	0.005	0.003
Max <20°	(0.093)	(0.007)	(0.007)	(0.001)	(0.069)	(0.009)	(0.004)	(0.002)
	0.002**	-0.001	0.001	0.002**	-0.002***	-0.007	-0.002	-0.003***
Max 30°-35°	(0.001)	(0.006)	(0.002)	(0.001)	(0.001)	(0.005)	(0.001)	(0.001)
	0.001	0.020***	0.001	0.002**	-0.004***	-0.010*	-0.002	-0.002
Max 35°+	(0.001)	(0.006)	(0.005)	(0.001)	(0.001)	(0.005)	(0.001)	(0.002)
2nd trimester (reference category: 20-30 degrees)								
	-0.246*	0.007	-0.009	-0.012***	-0.090	-0.006	-0.008	0.005
Max <20°	(0.128)	(0.009)	(0.010)	(0.003)	(0.064)	(0.010)	(0.005)	(0.005)
	0.002	0.001	-0.002	0.001	-0.001	-0.008*	-0.000	-0.004*
Max 30°-35°	(0.001)	(0.007)	(0.002)	(0.001)	(0.002)	(0.005)	(0.001)	(0.002)
	-0.000	0.000	-0.004	-0.000	-0.003*	0.002	0.001	-0.006***
Max 35°+	(0.001)	(0.005)	(0.002)	(0.001)	(0.002)	(0.004)	(0.001)	(0.002)
3rd trimester (reference category: 20-30 degrees)								
	0.101	-0.007	0.001	-0.010***	0.071	0.003	0.000	0.004*
Max <20°	(0.094)	(0.008)	(0.005)	(0.001)	(0.100)	(0.006)	(0.003)	(0.003)
	0.001	0.001	0.002	0.001*	-0.000	-0.001	-0.002**	-0.001*
Max 30°-35°	(0.001)	(0.006)	(0.002)	(0.001)	(0.000)	(0.004)	(0.001)	(0.000)
	0.001	0.010**	0.005*	0.001*	-0.003***	-0.005*	-0.003***	-0.003**
Max 35°+	(0.001)	(0.005)	(0.003)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
Const.	0.647***	0.621	1.835***	0.640***	3.025***	3.943***	1.783***	3.308***
	(0.083)	(0.487)	(0.179)	(0.191)	(0.063)	(0.542)	(0.077)	(0.086)
N	18,285	5,593	12,035	11,721	9,266	2,366	5,677	6,004
R-sq	0.210	0.317	0.243	0.076	0.498	0.638	0.423	0.166

Table 3 – Regression estimates – heterogeneity analysis

Notes: This table shows regression estimates and cluster-robust standard errors in parentheses, separately for Aboriginal and non-Aboriginal children, and children born in tropical vs arid climates. Regressions are estimated with the model specified in equation (1) for health outcomes and equation (2) for NAPLAN outcomes. As specified in equations (1) and (2), the regressions also include covariates (mother's age, Aboriginal status, whether mother's first pregnancy), month-year fixed effects and location-month-sex fixed effects (these are absorbed using the Stata 'areg', which affects the intercept but not the coefficients of interest).

4.6 Causal pathways

As noted in Section 2, if the impacts of heat exposure on health and cognitive outcomes are the result of epigenetic effects (often referred to as fetal programming), then it may be that heat exposure has direct effects on brain development and hence may affect cognitive outcomes independently of any impacts on health. An alternative explanation is that there are 'developmental cascades' through which prenatal heat exposure causes poor health in early life, and this poor health is part of the causal chain affecting cognitive development.

If the latter is true, we would expect health at birth to be either an intermediary cause or an early symptom of long-term impacts of heat exposure. This would mean that when we re-estimate our regressions controlling for indicators of perinatal health, the treatment effects should fall dramatically. And similarly, if poorer health in childhood hampers cognitive development, then our estimated effect of heat exposure on school test scores should be much smaller when we control for hospital admissions in childhood.

To test these pathways, we rerun our regressions controlling for these early indicators of health (Appendix Tables H2 and H3). On hospital admissions, we find that when we control for indicators of health at birth (gestational age, birthweight and Apgar scores) our estimates are little changed, though separately, birth outcomes are highly statistically significant in predicting admissions. For instance, the estimated impact of additional days of 35+ degrees in the third trimester reduces from 0.004 to 0.003. Similarly, we find very little change in our estimated treatment effects if instead of (or as well as) health at birth, we control for obstetric and labour complications and admission to a special care nursery.

On school test scores, our estimates are unchanged when we control health at birth, obstetric and labour complications. In addition, while hospital admissions in childhood are strongly predictive of school test scores, controlling for them does little to attenuate the impact of heat exposure on test scores.

An alternative approach is to analyse who is affected by heat exposure: are effects of heat exposure experienced by children with poor health at birth, by children who are otherwise healthy at birth, or both groups? To do this, we rerun our analysis only on children with no health concerns at birth. We find that, across both sets of outcomes, there are some

groups for whom effects of prenatal heat exposure are smaller and not statistically significant. But these are different groups for each outcome: children with no health concerns at birth experience smaller impacts on hospital admissions (Table H4) but heat exposure still impacts their school test scores (Table H5); conversely, children with no obstetric or labour complications experience no statistically significant effects of heat exposure on school test scores (Tabel H5), but heat exposure still affects their hospital admissions (Table H4). Our interpretation of these estimates is that poor early life health may accentuate the effects of heat exposure on cognitive development (i.e. act as an effect moderator), but this is not the sole pathway. Children with and without poor health at birth can be affected.

An additional possibility is that there are specific health conditions caused by heat exposure which themselves set off an unfavourable 'developmental cascade'. Auger et al. (2017) find for instance, that heat exposure is associated with some congenital conditions. As discussed above, we did not find any specific primary diagnosis driving hospital admissions. In Appendix Table H1, we further test for a link between prenatal heat exposure and diagnosis for any congenital condition. We find precise zero effects. While this is not conclusive evidence that heat exposure does not lead to congenital conditions (as these conditions may not require hospital admission), it confirms that the effects we find on hospital admissions in the NT are not the result of congenital conditions.

Taken together, this additional analysis suggests that it is unlikely heat exposure affects health in childhood through its impacts on birth outcomes, nor that heat exposure affects school test scores through its impacts on acute health conditions in childhood. The fact that treatment effects are smaller for groups with better health at birth suggests that developmental cascades may play a role. But even children with no apparent health conditions at birth experience some effects of heat exposure on their health, their cognitive outcome, or both. It is likely that heat exposure also has epigenetic effect, directly affecting cognitive development.

4.7 Discussion and conclusion

We have estimated the impact of heat exposure in utero on children's health and education outcomes from birth through to age 12. We first confirm that in our population, as in many other studies, we find impacts of prenatal heat exposure on birth outcomes: additional hot days in pregnancy (maximum above 30 degrees) reduce average birthweight, and additional cooler days (maximum below 20 degrees) increase birthweight. We then turn to outcomes in childhood: we find that additional days of exposure to temperatures above 30 degrees Celsius appear to impair cognitive development, reflected in lower school test scores at ages 8, 10 and 12, across all domains tested (reading, writing, spelling, grammar and numeracy). We also find that while prenatal heat exposure does not affect total hospital admissions at ages 0-12, it does increase admissions at ages 0-2, with effects largest for prenatal exposure to extreme heat of 40 degrees or higher.

The timing of heat exposure that matters depends on the outcome: while third trimester exposure appears to have the greatest effect on hospital admissions, exposure in all trimesters affects school test scores. These patterns vary somewhat within subsamples, so we cannot draw any strong conclusions around the timing of exposure that has the greatest impact.

We do not find any specific diagnoses that explain the impact of prenatal heat exposure on hospital admissions. We interpret this to mean that the effect of heat exposure is not driven by any specific condition that requires acute treatment in hospital. Instead, it may be that heat exposure leads to more minor health concerns which do not usually require hospitalisation, but which may act as complicating factors when children fall ill for other reasons, making them more likely to be hospitalised. For children aged 3 and above, for whom hospitalisation is a rare event (see Table 1), we see no clear impact of heat exposure on admissions. Our findings do not necessarily mean that the impacts of heat exposure on health disappear from age 3, just that they do not show up in acute conditions requiring hospitalisation. Further research, in settings where primary care data are available, may be able to detect or rule out impacts on less acute health conditions.

The pattern of effects we see sheds new light on how heat exposure in utero affects children's development. Many studies find that prenatal heat exposure leads to poorer

birth outcomes and speculate that this implies higher risks to long-term health and social outcomes. We find that while heat exposure in utero does affect longer-term health and education outcomes, this is not *because of* poorer health at birth. Health at birth explains, at best, a small share of the impacts of heat exposure on childhood health and education outcomes. Furthermore, even children who appear healthy at birth, whose mothers experienced no complications during pregnancy or labour, and who did not have health concerns at birth leading to admission to a special care nursery, are affected in childhood by prenatal heat exposure. It therefore appears that the impacts of prenatal heat exposure on birth outcomes, health in early childhood and cognitive development through to age 12 are, at least to some extent, independent of each other.

This finding is important for the way we interpret future research. Clearly birth outcomes are a more immediate, and more easily measured, outcome than cognitive development at age 12. If birth outcomes were a reliable early indicator of the later life impacts of in utero heat exposure, this would be very helpful. But our analysis suggests this is not the case. This has two implications: first, if we observe no impacts of prenatal heat exposure on birth outcomes in a given context, this does not mean that there will be no longer-term impacts on children's development. Second, children with poor health at birth may be at greater risk of poorer cognitive outcomes due to heat exposure than other children, but other children are also at risk.

The fact that we find precise and persistent effects of heat exposure in the NT is interesting, because the NT has a warm baseline climate and widespread availability of air-conditioning; indeed, we see the largest effects among non-Aboriginal population, who are most likely to live in urban areas where access to air-conditioning is almost universal (Australian Bureau of Statistics 2014). Our findings stand in contrast with suggestions in past research by Isen et al. (2017) and Ai & Tan (2023) that the effects of in utero heat exposure may disappear with the increased uptake of air conditioning. This is a puzzle: it may be that many households in the NT do not regularly use air conditioning, even though it is available, or that they use it, but not as frequently as would be required to shield them from the adverse effects of heat exposure. The fact that we find impacts of heat exposure even among this population suggests that the impacts of heat exposure in other contexts where air-conditioning is less readily available, may be even larger.

The magnitudes of our estimates are small. However, the unit of treatment is also small – it is just one unusually hot day within 9 months of pregnancy, in a region where it is already hot most of the time. On average, babies in the NT experience 15 days above 35 degrees in the third trimester, with a standard deviation of 24 days. A standard deviation increase in the number of days with maximum temperatures of 35+ days would therefore imply 0.1 additional hospital admission at ages 0–2 – a 9.4% increase on baseline – and a 0.01 point reduction in school test scores at age 8 – a 3.5% decrease on baseline.

As we are unaware of any other studies that have analysed the impacts of prenatal heat exposure on school test scores in childhood, it is difficult to put these estimates into context. However, we make two broad comparisons.

First, we may consider the lasting impacts of prenatal heat exposure relative to other early life interventions. For example, Doyle (2024) estimates that a parenting intervention targeted at disadvantaged families in early childhood increased children's school test scores at age 9 by between 0.3 and 0.54 SDs. Chyn et al. (2021) find that intensive hospital care for very low birthweight children increased their school test scores to middle school by 0.34 SDs. Our estimates are orders of magnitude smaller: the impact we estimate of an additional day with maximum temperatures of 30-34.99 and 35+ degrees in the third trimester corresponds to a 0.0016 and 0.004 SD reduction in school test scores at age 12, respectively. This makes sense: these are intensive interventions targeted towards specific at-risk groups. In contrast, we estimate the impact of a single day of additional heat exposure within the 270 days of pregnancy.

Second, we may compare the impact of heat exposure in utero to heat exposure later in life (see Appendix I for comparison table). Our estimates are broadly similar to studies of how heat exposure in the lead up to the test day affects school test scores, which reflect learning loss due to heat exposure during class or study time – these are reviewed by Venegas Marin, Schwarz and Sabarwal (2024). For instance, our estimates are close to Cho's (2017) from South Korea, on the impact of additional hot days in the prior summer, on college entrance test scores a few months later (0.0064 to 0.0042 SDs). Our estimates are 2 to 6 times larger than estimates found by Park et al. (2020; 2021) of the impact of additional hot days in the prior year on high school test scores in the USA. They are,

however, substantially smaller than Park's (2022) estimates of the impact of temperatures above 32 degrees during exam periods (0.13SDs).

In sum, the impacts we find on school test scores are of similar magnitude to the impacts of learning loss due to heat exposure within the year of the test. But importantly, in hot climates like that of the NT, they likely occur in addition to contemporaneous impacts of heat exposure on learning. Another difference is that, although students may be able to catch up on lost learning at school, our estimates suggest persistent effects of prenatal heat exposure from ages 8 to 12. While these effects are small, the exposure affects everyone, and in a hot climate, most children are exposed to substantially more than one hot day in pregnancy. Our estimates imply that a standard deviation in exposure to maximum temperatures of 35+ in the third trimester (24 days) could reduce age 12 school test scores by 0.096 SDs – roughly 30% of the effect sizes of the more intensive interventions described above.

Our analysis raises new questions for further research. First, our findings raise questions around strategies to prevent long-term effects of heat exposure: given that we find persistent effects of heat exposure at ages 8, 10 and 12, it is possible that these effects would be evident earlier. Our analysis of these age groups reflects the way cognitive skills are measured institutionally in Australia. But there may be early predictors of cognitive development, at or before school starting age, with which affected children could be identified and targeted for early intervention. Additional data would be required to analyse this. Second, our analysis raises questions on the biological mechanisms behind these impacts. Are there different types of heat exposure, or different mechanisms – in terms of heat stress, heat strain or maternal behavioural responses to heat – that are driving these different channels? And are there increasing or decreasing marginal effects of additional days of extreme heat? Learning more about how these effects come about will be important to predict which outcomes are likely in other contexts, and which preventative measures would be most effective. Third, on whether there is any interaction between the impacts of prenatal heat exposure and heat exposure later in life: is it the case that prenatal heat exposure helps children 'adapt' to their warmer environment, making them less affected by heat exposure in childhood – i.e. a version of the 'thrifty phenotype' hypothesis (Barker 1998)? Or are children exposed to heat in utero more affected by heat

exposure on or around the test day? Future research could test for these effects, jointly analysing prenatal heat exposure and heat exposure around the time of the test.

Heat exposure, and in particular exposure to extreme levels of heat, is likely to continue to increase due to climate change. This will be the case in hot climates like the NT – where there will be even more days with extreme temperatures (CSIRO and Bureau of Meteorology 2022). It will also mean that extreme temperatures will become more frequent in locations where they are currently very rare. It will be increasingly important to both minimise the risk of adverse long-term outcomes by protecting pregnant people from heat exposure, and to find ways to identify which children are at highest risk of poor outcomes, and design interventions to help them catch up to their peers. Policies and actions to limit anthropogenic climate change would, of course, prevent some of these impacts on population health and productivity.

Appendix A: Sample selection and sensitivity

A key issue with using the presence or absence of interactions with the healthcare system as an outcome measure is dealing with attrition. There are three possible reasons that children may not appear in the admissions data: first, because they are healthy, and never required a hospital admission – this is what we want to measure. Second, because the child has moved out of the NT. Third, because they have died. We do not observe which children have moved interstate, or died, so we cannot distinguish between these three possibilities.

It is possible that some children may die as a result of heat exposure, leading to differential attrition in our population. However, even in remote locations where child mortality is higher than the national average, the all-cause child mortality rate is 25 deaths per 100,000 children (Australian Institute of Health and Welfare 2022). Therefore, any bias is likely to be very small.

We would not necessarily expect migration to cause differential attrition based on heat exposure in utero: it is difficult to imagine how marginally more or fewer hot days during pregnancy would lead families to leave the NT, relative to others experiencing the same weather conditions before or after pregnancy. We are therefore not concerned about differential attrition. Nonetheless, attrition due to migration is a concern because we analyse hospital admissions over several years. We would expect the number of families who have migrated out of a given birth cohort to increase as time passes, meaning that our statistical power to detect an effect, if one exists, decreases as our cohort ages. Outward migration rates from the NT are high – for instance, in 2001, around 8 percent of the resident population moved out of the NT, with slightly higher rates for families with young children (Australian Bureau of Statistics 2021). If our estimate of the impact of heat exposure on the number of hospital admissions loses statistical significance as children age, we want to know that this is because the causal effects of heat exposure fade over time, and not the result of a loss of statistical power due to migration.

To ensure that our findings are not affected by this attrition, we conduct our analysis only on children who appear at least once in the school records, meaning that they were present in the NT at birth, and (at least) at school starting age. This reduces our analysis sample to 25,416. In our analysis on hospital admissions at older ages (ages 8-12), this falls to

11,194, as only the older children in our cohort reach age 12 by 2017, when our admissions data ends.

Outward migration during childhood means that the sample we study is not fully representative of the NT population. We may expect a 'healthy worker effect', wherein children in families who migrate out of the NT after a child is born (and hence fall out of our sample) and children in families who migrate into the NT after the child is born (and hence were never in our sample) may be healthier. In addition, because migration rates are higher among non-Aboriginal than Aboriginal families, Aboriginal children will be over-represented in our sample.[45] This does not bias our estimates; we show in Appendix E that there is no evidence that prenatal heat exposure affects selection into our analysis sample. However, it does mean that we do not have a full picture of the impacts of heat exposure on children in migrating families, for whom the impacts of heat exposure may be different.

[45] Aboriginal children make up around 40% of births in the NT, but 50% of our analysis cohort.

		Number of hospital admissions				NAPLAN scores		
	Birth weight (g)	Age 0 to 2	Ages 3 to 7	Ages 8 to 12	Ages 0 to 12	Year 3 (age 8)	Year 5 (age 10)	Year 7 (age 12)
1st trimester (reference category: 20-30 degrees)								
Max <20°	-0.246	0.006	0.008***	-0.005***	0.007	-0.006**	-0.003	-0.002
	(2.927)	(0.013)	(0.003)	(0.001)	(0.019)	(0.003)	(0.005)	(0.003)
Max 30°-35°	-3.845***	0.003	-0.004**	0.000	0.001	0.001	-0.001	-0.002
	(1.412)	(0.004)	(0.002)	(0.000)	(0.005)	(0.001)	(0.001)	(0.002)
Max 35°+	-2.061	0.004	-0.005***	0.000	0.003	-0.001	-0.004	0.000
	(1.828)	(0.005)	(0.002)	(0.000)	(0.007)	(0.004)	(0.004)	(0.003)
2nd trimester (reference category: 20-30 degrees)								
Max <20°	20.994** *	-0.008	0.007***	-0.010***	-0.051	-0.022***	0.002	-0.010**
	(7.862)	(0.012)	(0.002)	(0.003)	(0.053)	(0.005)	(0.005)	(0.004)
Max 30°-35°	1.036	-0.001	0.001	-0.001	-0.002	-0.004***	-0.004***	-0.003**
	(1.470)	(0.003)	(0.001)	(0.001)	(0.005)	(0.001)	(0.001)	(0.001)
Max 35°+	2.296*	-0.004	0.004**	-0.001	0.000	-0.007**	-0.006***	-0.007***
	(1.206)	(0.006)	(0.002)	(0.002)	(0.007)	(0.003)	(0.002)	(0.002)
3rd trimester (reference category: 20-30 degrees)								
Max <20°	21.022** *	-0.007	0.018***	-0.002	0.026	0.006	-0.001	-0.009
	(3.237)	(0.011)	(0.002)	(0.001)	(0.022)	(0.009)	(0.010)	(0.007)
Max 30°-35°	-0.992**	0.002	-0.001	-0.000	0.002	-0.000	-0.001	-0.002
	(0.382)	(0.003)	(0.001)	(0.000)	(0.003)	(0.001)	(0.001)	(0.001)
Max 35°+	-3.073*	0.006	-0.001	0.000	0.007	-0.005***	-0.004**	-0.005***
	(1.609)	(0.007)	(0.001)	(0.001)	(0.009)	(0.002)	(0.002)	(0.002)
Constant	3,370.526***	0.646**	0.229**	0.066	1.282**	2.996***	4.853***	5.995***
	(118.511)	(0.277)	(0.102)	(0.101)	(0.488)	(0.120)	(0.111)	(0.109)
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month-community-sex FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Interact 35+ with indicator for youngest in class	No	No	No	No	No	Yes	Yes	Yes
N	2,862	2,862	2,862	2,862	2,862	2,862	2,862	2,862
R-sq	0.263	0.467	0.209	0.187	0.457	0.495	0.523	0.520

Table A1: Regression estimates for full common cohort

	Number of admissions - cohort to age 7			Number of admissions - cohort to age 12		
	Ages 0-2	Ages 3-7	Ages 0-7	Ages 0-2	Ages 3-7	Ages 0-7
1st trimester (reference category: 20-30 degrees)						
Max <20°	-0.006	0.002***	-0.007	-0.005	-0.000	-0.006
	(0.004)	(0.001)	(0.006)	(0.005)	(0.001)	(0.007)
Max 30°-35°	0.001	-0.001*	-0.000	0.001	-0.001	0.001
	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.002)
Max 35°+	0.001	-0.001*	0.001	0.001	-0.001*	0.000
	(0.002)	(0.000)	(0.003)	(0.002)	(0.001)	(0.003)
2nd trimester (reference category: 20-30 degrees)						
Max <20°	-0.011*	-0.002	-0.010	-0.018	0.008**	-0.012
	(0.007)	(0.001)	(0.009)	(0.011)	(0.003)	(0.013)
Max 30°-35°	-0.001	0.001	-0.001	0.000	0.001***	0.000
	(0.001)	(0.000)	(0.002)	(0.001)	(0.000)	(0.002)
Max 35°+	-0.002	0.001**	-0.001	-0.002	0.002***	-0.001
	(0.002)	(0.000)	(0.003)	(0.002)	(0.001)	(0.003)
3rd trimester (reference category: 20-30 degrees)						
Max <20°	-0.003	0.001	0.001	-0.016*	0.004	-0.015
	(0.003)	(0.001)	(0.005)	(0.008)	(0.002)	(0.010)
Max 30°-35°	0.002**	-0.000**	0.003***	0.003**	-0.001*	0.003*
	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.002)
Max 35°+	0.004**	-0.001**	0.006***	0.006*	-0.002***	0.007*
	(0.002)	(0.000)	(0.002)	(0.003)	(0.001)	(0.004)
Constant	0.815***	0.143***	1.170***	0.820***	0.179***	1.153***
	(0.143)	(0.022)	(0.245)	(0.128)	(0.024)	(0.162)
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Month-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Month-community-sex FE	Yes	Yes	Yes	Yes	Yes	Yes
Interact 35+ with indicator for youngest in class	No	No	No	No	No	No
Observations	23,756	23,756	23,756	11,194	11,194	11,194
R-sq	0.282	0.064	0.291	0.354	0.116	0.355

Table A2: Regression estimates for common admissions cohort

Appendix B: Robustness tests

We conduct four sets of robustness tests.

First, Appendix Table B1 presents unadjusted p-values alongside the sharpened-Q values (our adjustment for multiple hypothesis testing), which we use to indicate statistical significance in our main results (Table 2).

Second, we test the impact of changes to the way that we specify fixed effects (Appendix Tables B2 and B3). Our benchmark model uses fixed effects for the month-year that the child is conceived, and their month of conception, community and sex (interacted). If we instead specify the location fixed effects as month-community, or simply as community (not interacted with month or sex), our estimate are little changed. In the latter, our estimate of the impact of temperatures of 35 degrees on admissions becomes smaller and loses statistical significance – though this is only the case in our simple reporting of heat exposure – in our more detailed breakdown with a category for 40+ degrees, our estimates match our benchmark model (Appendix Table B4). We see very little impact of these different fixed effects specifications on our estimates on school test scores (Appendix Table B3).

Third, we test alternative approaches to accounting for relative age in class for our school test score outcomes. The first column in Appendix Table B5 shows our benchmark model, in which we interact exposure to temperatures above 35 degrees with an indicator for whether the child is in the youngest quarter of their year group. The interaction term is statistically significant for exposure in the first trimester, showing that the youngest children in the class experience a particularly large effect of exposure to temperatures of 35 and over in the first trimester. Columns 2 and 3 shows our estimates if we do not control for relative age in class at all, or when we include the dummy variable 'youngest in class' as a covariate but do not interact it with heat exposure. In both cases, we see a larger estimated effect of exposure to temperatures of 35+ in the first trimester, but no other differences. Column 4 shows estimates where we control for month of birth (given that birthday affects position in year more directly than month of conception, which is in our baseline model), and Column 5 omits the 'youngest in year' children from analysis entirely. In all cases, the magnitudes of our estimates vary, but they remain statistically significant at the 5% level or higher.

Fourth, to ensure that our estimates are not unduly influenced by a single test year, birth cohort, or unrelated events, we rerun our benchmark model sequentially omitting each year of data, for both hospital admissions and school test scores at age 8. Appendix Tables B6 and B7 shows that our estimates are similar across all subsamples, which suggests our conclusions are not driven by a single event or cohort.

In addition to these robustness tests, Appendix Table B8 reports results on a test of the equivalence of each level of heat exposure across all three trimesters of pregnancy.

		N. hospital	NAPLAN scores		
	Birth weight (grams)	admissions age 0-12	Year 3 (age 8)	Year 5 (age 10)	Year 7 (age 12)
	(1)	(2)	(3)	(4)	(5)
1st trimester (reference category: 20-30 degrees)					
Max <20°	-0.688	-0.004	0.004	0.002	-0.002
P-value	0.638	0.539	0.192	0.195	0.589
Sharpened-Q	0.271	1	0.209	0.209	0.498
Max 30°-35°	-0.665**	0.002	-0.002***	-0.003***	-0.001
P-value	0.005	0.479	<0.001	<0.001	0.329
Sharpened-Q	0.02	1	0.001	0.001	0.315
Max 35°+	-0.428	0.001	-0.003*	-0.005***	-0.000
P-value	0.129	0.774	0.0427	<0.001	0.740
Sharpened-Q	0.178	1	0.06	0.001	0.498
2nd trimester (reference category: 20-30 degrees)					
Max <20°	1.608	-0.008	-0.003	0.002	0.002
P-value	0.379	0.571	0.342	0.720	0.746
Sharpened-Q	0.24	1	0.315	0.498	0.498
Max 30°-35°	-0.213	-0.000	-0.002	-0.001	-0.001
P-value	0.401	0.929	0.229	0.306	0.409
Sharpened-Q	0.24	1	0.233	0.313	0.375
Max 35°+	0.507	-0.001	-0.003	-0.003**	-0.004***
P-value	0.073	0.866	0.149	0.027	<0.001
Sharpened-Q	0.137	1	0.176	0.042	0.005
3rd trimester (reference category: 20-30 degrees)					
Max <20°	4.800***	-0.009	-0.001	0.007**	-0.003
P-value	<0.001	0.445	0.810	0.025	0.587
Sharpened-Q	0.003	1	0.498	0.042	0.498
Max 30°-35°	-0.665	0.004	-0.001**	-0.002***	-0.002**
P-value	0.0505	0.418	0.015	0.003	0.010
Sharpened-Q	0.134	1	0.029	0.009	0.023
Max 35°+	-0.943	0.008	-0.004***	-0.003***	-0.005***
P-value	0.0857	0.066	<0.001	0.001	<0.001
Sharpened-Q	0.137	1	0.001	0.005	0.001
Constant	3,363.215	1.385	3.150	4.907	5.864

Table B1: Regression estimates, p-values and sharpened-Q values for main results table

	Admission ages 0-2				
	(1)	(2)	(3)	(4)	(5)
1st trimester (reference category: 20-30 degrees)					
Max <20°	-0.005	-0.007*	-0.004	-0.007*	-0.007
	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)
Max 30°-35°	0.001	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Max 35°+	0.001	0.002	0.002	0.002	0.002
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
2nd trimester (reference category: 20-30 degrees)					
Max <20°	-0.011*	-0.011*	-0.009**	-0.011*	-0.012*
	(0.007)	(0.006)	(0.004)	(0.006)	(0.007)
Max 30°-35°	-0.001	-0.001	0.000	-0.000	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Max 35°+	-0.002	-0.002	0.000	-0.002	-0.002
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
3rd trimester (reference category: 20-30 degrees)					
Max <20°	-0.003	-0.004	-0.004	-0.003	-0.002
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Max 30°-35°	0.002**	0.002**	0.002**	0.002**	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Max 35°+	0.004**	0.004**	0.002	0.003**	0.003***
	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)
Constant	0.804***	0.693***	0.734***	0.677***	1.919
	(0.148)	(0.123)	(0.188)	(0.145)	(1.573)
Observations	23,756	23,756	23,756	23,757	23,757
R-sq	0.282	0.250	0.212	0.278	0.281
Demographic controls	Y	Y	Y	Y	Y
Month-year conceived FE	Y	Y	Y		
Quarter-year conceived FE				Y	
Month-year born FE					Y
Month-community-sex FE	Y			Y	Y
Month-community FE		Y			
Community FE			Y		

Table B2: Regression estimates with alternative FE – hospital admissions

	NAPLAN average score, year 3				
	(1)	(2)	(3)	(4)	(5)
1st trimester (reference category: 20-30 degrees)					
Max <20°	0.004	0.003	-0.000	0.005	0.004
	(0.003)	(0.003)	(0.002)	(0.004)	(0.004)
Max 30°-35°	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
Max 35°+	-0.003**	-0.003**	-0.004***	-0.001	-0.003**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
2nd trimester (reference category: 20-30 degrees)					
Max <20	-0.003	-0.004	-0.002	-0.003	-0.004
	(0.003)	(0.004)	(0.002)	(0.003)	(0.003)
Max 30°-35°	-0.002	-0.001	-0.001*	-0.002*	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Max 35°+	-0.003	-0.002	-0.002**	-0.004**	-0.002
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
3rd trimester (reference category: 20-30 degrees)					
Max <20°	-0.001	0.001	0.004	0.001	-0.001
	(0.004)	(0.004)	(0.003)	(0.003)	(0.004)
Max 30°-35°	-0.001**	-0.001**	-0.001*	-0.001**	-0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Max 35°+	-0.004***	-0.004***	-0.003***	-0.003***	-0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Constant	3.112***	3.265***	3.212***	3.754***	5.276***
	(0.116)	(0.084)	(0.122)	(0.070)	(0.057)
Observations	11,681	11,681	11,681	11,682	11,682
R-sq	0.523	0.510	0.493	0.503	0.507
Demographic controls	Y	Y	Y	Y	Y
Month-year conceived FE	Y	Y	Y		
Quarter-year conceived FE				Y	
Month-year born FE					Y
Month-community-sex FE	Y			Y	Y
Month-community FE		Y			
Community FE			Y		

Table B3: Regression estimates with alternative FE – NAPLAN 3

	Admissions ages 0-2				
	(1)	(2)	(3)	(4)	(5)
1st trimester (reference category: 25-30 degrees)					
Max <20°	-0.005	-0.007*	-0.002	-0.006*	-0.005
	(0.004)	(0.004)	(0.002)	(0.004)	(0.004)
Max 20°-25°	-0.004	-0.004	-0.004**	-0.005**	-0.004
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Max 30°-35°	0.002**	0.002**	0.002**	0.002**	0.002*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Max 35°-40°	0.002	0.002	0.001	0.002	0.002
	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
Max 40°+	0.003	0.004	0.006**	0.003	0.003
	(0.004)	(0.004)	(0.002)	(0.003)	(0.004)
2nd trimester (reference category: 25-30 degrees)					
Max <20°	-0.010*	-0.010**	-0.007**	-0.009*	-0.009*
	(0.005)	(0.005)	(0.003)	(0.005)	(0.005)
Max 20°-25°	-0.007**	-0.007**	-0.002	-0.006**	-0.007**
	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)
Max 30°-35°	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Max 35°-40°	-0.001	-0.002	-0.001	-0.001	-0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Max 40°+	-0.001	-0.002	0.002	-0.000	-0.002
	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)
3rd trimester (reference category: 25-30 degrees)					
Max <20	-0.001	-0.001	-0.005	-0.000	0.001
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
Max 20°-25°	-0.005	-0.004	-0.000	-0.005	-0.004
	(0.004)	(0.003)	(0.003)	(0.004)	(0.004)
Max 30°-35°	0.002**	0.002**	0.001***	0.001**	0.002*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Max 35°-40°	0.002**	0.002**	0.002**	0.002**	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Max 40°+	0.006**	0.007**	0.004**	0.006**	0.006**
	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)
Constant	0.692***	0.671***	0.732***	0.594***	0.377***
	(0.113)	(0.105)	(0.117)	(0.085)	(0.064)
Observations	34,258	34,258	34,258	34,259	34,259
R-sq	0.269	0.240	0.209	0.265	0.267
Demographic controls	Y	Y	Y	Y	Y
Month-year conceived FE	Y	Y	Y		
Quarter-year conceived FE				Y	
Month-year born FE					Y
Month-community-sex FE	Y			Y	Y
Month-community FE		Y			
Community FE			Y		

Table B4: Regression estimates with alternative FE – hospital admissions with additional exposure categories

	NAPLAN average score year 3				
	(1) (benchmark)	(2)	(3)	(4)	(5)
1st trimester (reference category: 20-30 degrees)					
Max <20°	0.004	0.004	0.004	0.004	0.003
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
Max 30°-35°	-0.002***	-0.002***	-0.002***	-0.002***	-0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Max 35°+	-0.003**	-0.006***	-0.006***	-0.007***	-0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
2nd trimester (reference category: 20-30 degrees)					
Max <20°	-0.003	-0.002	-0.003	-0.003	-0.005
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Max 30°-35°	-0.002	-0.002	-0.002	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Max 35°+	-0.003	-0.002	-0.002	-0.002	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
3rd trimester (reference category: 20-30 degrees)					
Max <20°	-0.001	0.002	0.001	-0.001	-0.012**
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
Max 30°-35°	-0.001**	-0.001**	-0.001**	-0.001**	0.000
	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)
Max 35°+	-0.004***	-0.004***	-0.004***	-0.004***	-0.003**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Youngest in class (born Apr-Jun)	-0.033		-0.099***		
	(0.051)		(0.026)		
Youngest in class x days max 35°+ in 1st trimester	-0.005***				
	(0.001)				
Youngest in class x days max 35°+ in 2nd trimester	0.001				
	(0.002)				
Youngest in class x days max 35°+ in 3rd trimester	-0.000				
	(0.002)				
	3.185***	3.143***	3.142***	3.191***	3.024***
	(0.112)	(0.125)	(0.128)	(0.127)	(0.118)
Youngest in class control method	Interacted	No control	Covariate only	No control	Omitted
Month control	Conception	Conception	Conception	Birth	Conception
	11,681	11,681	11,681	11,681	8,532
	0.523	0.522	0.522	0.524	0.531

Table B5: Regression estimates with alternative ways of controlling for relative age in class – NAPLAN

	Admissions ages 0 to 2									
Year of conception omitted:	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
1st trimester (reference category: 20-30 degrees)										
Max <20°	-0.006	-0.006	-0.004	-0.007*	-0.004	-0.003	-0.004	-0.009*	-0.006	-0.006
	(0.006)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.005)	(0.004)	(0.004)
Max 30°-35°	0.001	0.001	0.001	0.001	0.001	0.000	0.001	0.002	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Max 35°+	0.002	0.001	0.001	0.002	0.002	0.001	0.001	0.002	0.002	0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
2nd trimester (reference category: 20-30 degrees)										
Max <20°	-0.012	-0.014	-0.005	-0.012	-0.006	-0.012**	-0.014**	-0.011**	-0.013**	-0.007
	(0.009)	(0.011)	(0.005)	(0.008)	(0.006)	(0.005)	(0.007)	(0.005)	(0.006)	(0.006)
Max 30°-35°	-0.000	-0.000	-0.001	-0.000	-0.002	0.000	-0.001	-0.001	-0.000	-0.001
	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
Max 35°+	-0.001	-0.001	-0.002	-0.001	-0.003	-0.001	-0.002	-0.003	-0.002	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
3rd trimester (reference category: 20-30 degrees)										
Max <20°	-0.001	-0.008**	-0.001	-0.001	-0.004	-0.005	-0.002	-0.002	-0.003	-0.003
	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.004)	(0.003)	(0.004)	(0.003)
Max 30°-35°	0.002*	0.002**	0.002**	0.001	0.003**	0.001	0.002**	0.002*	0.003**	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Max 35°+	0.003	0.004**	0.004**	0.003**	0.005**	0.002	0.004**	0.004**	0.005**	0.004**
	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Constant	0.744***	0.826**	0.825**	0.775**	0.866**	0.848**	0.870**	0.845**	0.708**	0.837**
	(0.133)	(0.121)	(0.117)	(0.167)	(0.161)	(0.113)	(0.126)	(0.115)	(0.150)	(0.135)
Observations	21,643	21,294	21,321	21,394	21,308	21,291	21,222	21,172	21,215	21,944
R-sq	0.288	0.289	0.285	0.290	0.294	0.290	0.289	0.290	0.292	0.290

Table B6: Regression estimates on hospital admissions – leaving one year out

	NAPLAN average score Year 3							
Test year omitted:	2009	2010	2011	2012	2013	2014	2015	2016
1st trimester (reference category: 20-30 degrees)								
Max <20°	0.010***	0.003	0.005	0.004	0.005	0.003	0.003	0.005
	(0.003)	(0.003)	(0.004)	(0.002)	(0.004)	(0.004)	(0.003)	(0.005)
Max 30°-35°	-0.003***	-0.002***	-0.003***	-0.002***	-0.002***	-0.002***	-0.002***	-0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Max 35°+	-0.003**	-0.003**	-0.003**	-0.002	-0.003**	-0.003**	-0.003**	-0.003
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)
2nd trimester (reference category: 20-30 degrees)								
Max <20°	-0.002	-0.003	-0.002	-0.005	-0.004	-0.004	-0.006*	-0.000
	(0.004)	(0.003)	(0.003)	(0.005)	(0.003)	(0.004)	(0.003)	(0.002)
Max 30°-35°	-0.001	-0.002	-0.001	-0.002	-0.002	-0.001	-0.002	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Max 35°+	-0.002	-0.002	-0.002	-0.003	-0.003	-0.002	-0.004**	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
3rd trimester (reference category: 20-30 degrees)								
Max <20°	-0.001	0.001	0.001	-0.004	-0.003	0.002	-0.002	-0.001
	(0.004)	(0.004)	(0.002)	(0.004)	(0.005)	(0.004)	(0.006)	(0.003)
Max 30°-35°	-0.001**	-0.001**	-0.001*	-0.001**	-0.001	-0.001**	-0.002***	-0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
Max 35°+	-0.004***	-0.004***	-0.004***	-0.004***	-0.004***	-0.004***	-0.004***	-0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Constant	3.012***	3.046***	3.037***	3.118***	3.255***	3.120***	3.149***	3.131***
	(0.159)	(0.138)	(0.177)	(0.114)	(0.114)	(0.133)	(0.118)	(0.167)
Observations	10,784	10,100	10,108	10,174	10,194	10,170	10,108	10,129
R-sq	0.521	0.527	0.533	0.533	0.526	0.528	0.526	0.519

Table B7: Regression estimates on NAPLAN 3 – leaving one year out

		Number of hospital admissions				NAPLAN scores		
	Birth weight (grams)	Age 0 to 12	Ages 0 to 2	Ages 3 to 7	Ages 8 to 12	Year 3 (age 8)	Year 5 (age 10)	Year 7 (age 12)
Max <20	0.00	0.87	0.45	0.12	0.01	0.16	0.46	0.81
Max 30-35	0.22	0.39	0.17	0.02	0.21	0.18	0.05	0.98
Max 35+	0.07	0.14	0.08	0.04	0.01	0.48	0.40	0.11

Table B8: F-test of equivalence of trimester-specific effects in main regression results
Note: this table shows the p-values from a test of the equivalence of heat exposure across all trimesters (e.g. coefficient for under 20 in trimester 1 = under 20 in trimester 2 = under 20 in trimester 3)

Appendix C: Additional heat exposure metrics – hospital admissions

	Number of admissions				Days spent in hospital				Ever admitted to hospital			
	Ages 0-2	Ages 3-7	Ages 8-12	Ages 0-12	Ages 0-2	Ages 3-7	Ages 8-12	Ages 0-12	Ages 0-2	Ages 3-7	Ages 8-12	Ages 0-12
1st trimester (reference category: 25-30 degrees)												
Max <20°	-0.008	-0.000	-0.002	-0.027*	0.034	0.005	-0.002	0.014	0.002**	-0.001	-0.001	-0.004**
	(0.006)	(0.001)	(0.001)	(0.014)	(0.042)	(0.005)	(0.003)	(0.078)	(0.001)	(0.001)	(0.001)	(0.001)
Max 20°-25°	-0.002	0.002	0.002	-0.029*	-0.007	0.006	0.006	-0.050	0.001	0.001	0.000	-0.002*
	(0.004)	(0.002)	(0.002)	(0.016)	(0.035)	(0.005)	(0.004)	(0.087)	(0.001)	(0.001)	(0.001)	(0.001)
Max 30°-35°	0.001	-0.000	-0.000	0.000	-0.006	-0.002	-0.003***	-0.007	0.000	-0.000**	0.000	-0.000
	(0.001)	(0.000)	(0.000)	(0.002)	(0.008)	(0.003)	(0.001)	(0.017)	(0.000)	(0.000)	(0.000)	(0.000)
Max 35°-40°	0.002	-0.001*	-0.000	0.002	-0.009	-0.004	-0.002*	0.006	0.001	-0.001*	0.000	0.000
	(0.003)	(0.000)	(0.000)	(0.005)	(0.020)	(0.004)	(0.001)	(0.038)	(0.001)	(0.000)	(0.000)	(0.000)
Max 40°+	0.004	0.001	-0.001	-0.004	0.028	-0.009	-0.005**	0.022	0.000	-0.000	0.000	-0.000
	(0.005)	(0.001)	(0.001)	(0.010)	(0.031)	(0.011)	(0.002)	(0.051)	(0.001)	(0.001)	(0.001)	(0.001)
2nd trimester (reference category: 25-30 degrees)												
Max <20°	-0.017*	0.001	-0.001	0.019	-0.164**	0.003	-0.005	0.107	-0.001	0.002*	-0.000	0.003*
	(0.009)	(0.002)	(0.002)	(0.023)	(0.068)	(0.009)	(0.007)	(0.158)	(0.002)	(0.001)	(0.001)	(0.002)
Max 20°-25°	-0.005	-0.001	0.002*	0.001	-0.026	-0.001	0.009*	0.014	-0.003***	0.000	0.001	-0.001
	(0.006)	(0.001)	(0.001)	(0.019)	(0.038)	(0.010)	(0.005)	(0.115)	(0.001)	(0.001)	(0.001)	(0.002)
Max 30°-35°	0.000	0.000	-0.000	0.001	-0.005	0.003	-0.001**	-0.003	0.000	0.000*	-0.000	0.001**
	(0.001)	(0.000)	(0.000)	(0.003)	(0.008)	(0.002)	(0.001)	(0.016)	(0.000)	(0.000)	(0.000)	(0.000)
Max 35°-40°	0.001	0.001	0.000	0.008	0.008	0.005	0.001	-0.008	0.000	0.001***	-0.000	0.003***
	(0.002)	(0.001)	(0.001)	(0.006)	(0.014)	(0.004)	(0.002)	(0.031)	(0.001)	(0.000)	(0.001)	(0.001)
Max 40°+	-0.001	0.000	0.002*	-0.004	-0.015	0.006	0.003	-0.054	-0.000	0.001	0.001	0.001
	(0.004)	(0.001)	(0.001)	(0.008)	(0.027)	(0.007)	(0.002)	(0.044)	(0.001)	(0.001)	(0.000)	(0.001)
3rd trimester (reference category: 25-30 degrees)												
Max <20°	0.000	0.002	0.004**	-0.019	-0.100*	0.012*	0.005	-0.413***	(0.001)	0.001	0.002*	0.004**
	(0.009)	(0.002)	(0.002)	(0.020)	(0.057)	(0.007)	(0.004)	(0.135)	0.002	(0.002)	(0.001)	(0.002)
Max 20°-25°	-0.005	-0.001	0.002	-0.034	-0.067	0.018	0.010**	-0.136	(0.003)	-0.001	0.001	-0.001
	(0.005)	(0.001)	(0.002)	(0.027)	(0.043)	(0.018)	(0.005)	(0.192)	-0.000	(0.001)	(0.001)	(0.002)
Max 30°-35°	0.002*	-0.000	0.000	0.004*	0.026***	-0.001	0.002***	0.053***	(0.001)	-0.000	0.000	-0.001*
	(0.001)	(0.000)	(0.000)	(0.002)	(0.008)	(0.002)	(0.001)	(0.013)	0.000	(0.000)	(0.000)	(0.000)
Max 35°-40°	0.001	-0.001	0.001**	0.007	0.040**	-0.003	0.002	0.114***	(0.000)	-0.000	0.000*	-0.000
	(0.002)	(0.001)	(0.000)	(0.005)	(0.015)	(0.003)	(0.002)	(0.034)	0.000	(0.000)	(0.000)	(0.001)
Max 40°+	0.008**	0.000	-0.001	0.009	0.064*	0.004	-0.002	0.090*	(0.001)	-0.000	-0.000	-0.001
	(0.004)	(0.001)	(0.001)	(0.009)	(0.037)	(0.003)	(0.002)	(0.049)	0.001	(0.000)	(0.000)	(0.001)
1st trimester (reference category: 15-20 degrees)												
Min <5°	-0.001	0.004	0.003***	0.036*	-0.056	-0.001	0.007*	0.080	(0.001)	0.003*	0.002**	0.005**
	(0.006)	(0.002)	(0.001)	(0.018)	(0.046)	(0.006)	(0.003)	(0.177)	-0.004***	(0.001)	(0.001)	(0.002)
Min 5°-10°	-0.003	0.001	0.001	0.001	0.017	-0.005	-0.002	-0.028	(0.001)	-0.001	0.001	-0.000
	(0.005)	(0.001)	(0.001)	(0.013)	(0.042)	(0.006)	(0.003)	(0.081)	-0.003**	(0.001)	(0.001)	(0.002)

Min 10°-15°	-0.002	-0.001	0.001*	-0.002	-0.016	-0.003	0.005	-0.035	(0.001)	-0.000	0.001**	0.002
	(0.004)	(0.001)	(0.001)	(0.008)	(0.028)	(0.004)	(0.003)	(0.054)	-0.001	(0.001)	(0.000)	(0.001)
Min 20°-25°	0.004**	-0.000	0.002*	0.009*	0.029**	-0.000	0.004**	-0.014	(0.001)	0.000	0.001	0.001
	(0.002)	(0.000)	(0.001)	(0.005)	(0.014)	(0.003)	(0.002)	(0.028)	-0.000	(0.000)	(0.001)	(0.001)
Min 25°+	0.004*	-0.000	0.001	0.009	0.023	0.001	0.004	0.012	(0.001)	0.000	0.000	0.003**
	(0.002)	(0.001)	(0.001)	(0.007)	(0.016)	(0.005)	(0.003)	(0.036)	0.000	(0.000)	(0.001)	(0.001)
2nd trimester (reference category: 15-20 degrees)												
Min <5°	0.004	-0.002	0.002	-0.031	0.066	-0.001	0.000	-0.145	(0.001)	-0.002*	0.002	-0.003
	(0.009)	(0.002)	(0.002)	(0.023)	(0.057)	(0.008)	(0.006)	(0.143)	0.002	(0.001)	(0.001)	(0.003)
Min 5°-10°	-0.001	0.001	-0.002	0.004	0.034	0.007	-0.008*	-0.020	(0.002)	0.000	-0.001	0.001
	(0.006)	(0.001)	(0.002)	(0.023)	(0.046)	(0.008)	(0.005)	(0.109)	0.000	(0.001)	(0.001)	(0.002)
Min 10°-15°	-0.004	0.001	-0.000	0.001	-0.019	0.006*	-0.000	0.038	(0.001)	0.001**	0.001	0.002
	(0.004)	(0.001)	(0.002)	(0.010)	(0.029)	(0.004)	(0.003)	(0.060)	-0.000	(0.001)	(0.001)	(0.001)
Min 20°-25°	-0.004**	0.001*	0.001*	-0.001	-0.003	0.003	-0.001	0.009	(0.001)	0.000	0.001	-0.000
	(0.002)	(0.000)	(0.000)	(0.006)	(0.019)	(0.004)	(0.002)	(0.036)	-0.000	(0.000)	(0.000)	(0.001)
Min 25°+	-0.003*	0.001	0.000	0.001	0.002	0.003	0.000	0.055	(0.000)	0.000	0.000	-0.001
	(0.002)	(0.001)	(0.000)	(0.005)	(0.019)	(0.005)	(0.002)	(0.034)	-0.000	(0.000)	(0.000)	(0.001)
3rd trimester (reference category: 15-20 degrees)												
Min <5°	-0.004	-0.002	-0.005***	0.015	0.022	-0.017**	-0.016***	0.222	(0.001)	-0.002	-0.002**	-0.003*
	(0.009)	(0.002)	(0.001)	(0.020)	(0.046)	(0.008)	(0.005)	(0.147)	-0.001	(0.001)	(0.001)	(0.002)
Min 5°-10°	-0.001	0.002*	0.001	0.008	0.046	-0.007	-0.000	0.135	(0.002)	0.001**	0.000	-0.002
	(0.005)	(0.001)	(0.002)	(0.013)	(0.052)	(0.015)	(0.004)	(0.135)	-0.002**	(0.001)	(0.001)	(0.002)
Min 10°-15°	0.001	-0.000	0.000	-0.001	0.009	-0.002	-0.001	0.044	(0.001)	-0.000	-0.000	-0.001
	(0.004)	(0.001)	(0.001)	(0.008)	(0.031)	(0.006)	(0.003)	(0.056)	0.001	(0.001)	(0.001)	(0.001)
Min 20°-25°	0.001	0.000	-0.000	-0.002	0.022	-0.003	-0.002	-0.011	(0.001)	0.000	-0.000	-0.001**
	(0.002)	(0.000)	(0.000)	(0.005)	(0.018)	(0.002)	(0.002)	(0.040)	-0.000	(0.000)	(0.000)	(0.001)
Min 25°+	0.003	-0.000	0.000	0.003	0.025	-0.003	-0.001	0.011	(0.001)	0.000*	0.000	-0.000
	(0.002)	(0.000)	(0.000)	(0.004)	(0.015)	(0.003)	(0.001)	(0.028)	0.000	(0.000)	(0.000)	(0.000)
Constant	0.723***	0.091*	-0.129	0.647	-0.673	0.146	-0.012	-0.685	0.464***	0.033	-0.047	0.489***
	(0.217)	(0.051)	(0.111)	(0.965)	(1.250)	(0.378)	(0.344)	(5.867)	(0.049)	(0.037)	(0.081)	(0.099)
Observations	23,756	23,756	11,194	11,194	23,756	23,756	11,194	11,194	23,756	23,756	11,194	11,194
R-sq	0.283	0.065	0.126	0.371	0.257	0.104	0.154	0.360	0.208	0.069	0.120	0.236
Demographic controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Month-year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Month-community-sex FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Table C1: Regression estimates for admissions, with maximum and minimum temperature exposure

	Number of admissions				Days spent in hospital				Ever admitted to hospital			
	Ages 0-2	Ages 3-7	Ages 8-12	Ages 0-12	Ages 0-2	Ages 3-7	Ages 8-12	Ages 0-12	Ages 0-2	Ages 3-7	Ages 8-12	Ages 0-12
1st trimester (reference category: 25-30 degrees)												
Max <20°	-0.007*	0.000	-0.001*	-0.008	0.017	0.000	-0.002	0.044	-0.000	0.000	-0.001	-0.001
	(0.004)	(0.000)	(0.001)	(0.007)	(0.031)	(0.000)	(0.002)	(0.051)	(0.001)	(0.000)	(0.000)	(0.001)
Max 20°-25°	-0.004	-0.000*	0.002	-0.022**	-0.002	-0.001	0.003	-0.053	-0.001	-0.000*	0.001	-0.002*
	(0.003)	(0.000)	(0.001)	(0.010)	(0.027)	(0.001)	(0.003)	(0.056)	(0.001)	(0.000)	(0.001)	(0.001)
Max 30°-35°	0.001	-0.000*	-0.000	0.001	-0.004	-0.000*	-0.002***	-0.002	0.000*	-0.000	-0.000	0.000
	(0.001)	(0.000)	(0.000)	(0.002)	(0.008)	(0.000)	(0.001)	(0.015)	(0.000)	(0.000)	(0.000)	(0.000)
Max 35°-40°	0.001	-0.000*	-0.000*	0.001	-0.010	-0.000**	-0.002***	-0.010	0.000	-0.000**	0.000	-0.000
	(0.002)	(0.000)	(0.000)	(0.004)	(0.012)	(0.000)	(0.001)	(0.024)	(0.000)	(0.000)	(0.000)	(0.000)
Max 40°+	0.006	-0.000	-0.000	-0.001	0.031	-0.000	-0.003	0.022	0.001	-0.000	-0.000	0.001
	(0.004)	(0.000)	(0.001)	(0.008)	(0.027)	(0.000)	(0.002)	(0.040)	(0.001)	(0.000)	(0.000)	(0.001)
2nd trimester (reference category: 25-30 degrees)												
Max <20°	-0.012*	0.000	-0.001	-0.003	-0.073**	0.000	-0.006**	0.011	0.001	0.000	0.000	0.001
	(0.007)	(0.000)	(0.001)	(0.013)	(0.033)	(0.001)	(0.002)	(0.121)	(0.001)	(0.000)	(0.001)	(0.002)
Max 20°-25°	-0.005	0.000	0.001	-0.010	-0.006	0.000	0.004	-0.065	-0.002***	0.000	0.001	-0.002
	(0.004)	(0.000)	(0.001)	(0.012)	(0.029)	(0.000)	(0.003)	(0.077)	(0.001)	(0.000)	(0.001)	(0.002)
Max 30°-35°	-0.000	0.000	-0.000	0.001	-0.004	0.000	-0.001	-0.010	0.000	0.000	-0.000	0.001
	(0.001)	(0.000)	(0.000)	(0.002)	(0.007)	(0.000)	(0.001)	(0.013)	(0.000)	(0.000)	(0.000)	(0.000)
Max 35°-40°	-0.001	0.000***	-0.000	0.001	-0.007	0.000***	-0.000	-0.025	-0.000	0.000***	-0.000	0.001**
	(0.002)	(0.000)	(0.001)	(0.003)	(0.010)	(0.000)	(0.001)	(0.023)	(0.001)	(0.000)	(0.000)	(0.000)
Max 40°+	-0.004	0.000	0.002**	-0.006	-0.022	0.000	0.003	-0.047	-0.001	0.000	0.001	0.001
	(0.004)	(0.000)	(0.001)	(0.007)	(0.023)	(0.000)	(0.002)	(0.041)	(0.001)	(0.000)	(0.001)	(0.001)
3rd trimester (reference category: 25-30 degrees)												
Max <20°	-0.002	-0.000*	0.001*	0.002	-0.066**	-0.000	-0.002	-0.202**	(0.001)	-0.000*	0.001	0.004***
	(0.003)	(0.000)	(0.001)	(0.015)	(0.030)	(0.000)	(0.003)	(0.084)	-0.000	(0.000)	(0.001)	(0.001)
Max 20°-25°	-0.007	-0.000	0.001	-0.024	-0.049*	-0.000	0.007	0.002	(0.001)	-0.000	0.001	-0.003***
	(0.005)	(0.000)	(0.002)	(0.023)	(0.029)	(0.000)	(0.005)	(0.117)	-0.001	(0.000)	(0.001)	(0.001)
Max 30°-35°	0.002**	-0.000	0.000	0.003*	0.023***	-0.000	0.002**	0.039***	(0.001)	-0.000	0.000*	-0.000
	(0.001)	(0.000)	(0.000)	(0.002)	(0.008)	(0.000)	(0.001)	(0.012)	0.000	(0.000)	(0.000)	(0.000)
Max 35°-40°	0.002*	-0.000	0.000	0.005	0.034***	-0.000	0.002	0.077**	(0.000)	-0.000	0.000*	-0.000
	(0.001)	(0.000)	(0.000)	(0.004)	(0.012)	(0.000)	(0.002)	(0.032)	0.000	(0.000)	(0.000)	(0.000)
Max 40°+	0.010***	-0.000**	-0.001	0.012	0.075**	-0.000	-0.001	0.073	(0.000)	-0.000**	0.000	0.000
	(0.003)	(0.000)	(0.001)	(0.009)	(0.032)	(0.000)	(0.002)	(0.050)	0.001	(0.000)	(0.000)	(0.001)
Constant	0.838***	0.002	0.022	1.545***	2.736**	0.013**	0.082	3.985**	0.425***	0.003	0.037**	0.584***
	(0.118)	(0.004)	(0.029)	(0.311)	(1.083)	(0.006)	(0.063)	(1.650)	(0.023)	(0.004)	(0.015)	(0.019)
Observations	23,756	23,756	11,194	11,194	23,756	23,756	11,194	11,194	23,756	23,756	11,194	11,194
R-sq	0.283	0.042	0.124	0.369	0.257	0.052	0.152	0.359	0.208	0.046	0.119	0.234

Demographic controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Month-year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Month-community-sex FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Table C2: Regression estimates for admissions, with more granular bands of temperature exposure

Notes: This table shows regression estimates and cluster-robust standard errors in parentheses. Each column represents a different regression, with the outcome specified in the first row. Regressions are estimated with the model specified in equation (1). Admission counts and number of day are top-coded above the 97.5th percentile. As specified in equation (1), the regressions also include covariates (mother's age, Aboriginal status, whether mother's first pregnancy), month-year fixed effects and location-month-sex fixed effects (these are absorbed using the Stata 'areg', which affects the intercept but not the coefficients of interest).

Appendix D: Diagnoses

			ICD 10 chapters										
	Total admissions	Admission for infection	Infections	Respiratory	Digestive	Skin	Perinatal	Congenital	Misc	Accidents	Factors	Factors: Newborn admission no condition found	Num diagnoses per admission
1st trimester (reference category: 25-30 degrees)													
Max <20°	-0.007*	-0.004	-0.003	-0.004**	-0.000	-0.000	-0.001	-0.000	0.000	-0.001**	0.000	-0.000	-0.003
	(0.004)	(0.004)	(0.003)	(0.002)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.004)
Max 20°-25°	-0.004	-0.004	-0.001	-0.003*	-0.000	-0.001	0.002**	-0.001**	-0.000	-0.001	0.000	0.000	0.003
	(0.003)	(0.003)	(0.001)	(0.002)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.006)
Max 30°-35°	0.001	0.001	0.000	0.000	0.000	0.000	0.000**	0.000	0.001***	-0.000	-0.000	-0.000*	0.000
	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)
Max 35°-40°	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001**	0.000	0.000	-0.000	0.001
	(0.002)	(0.002)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)
Max 40°+	0.006	0.002	0.002	0.000	0.000	-0.000	0.001	0.000	0.001	0.000	0.000	0.000	0.001
	(0.004)	(0.003)	(0.001)	(0.002)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.005)
2nd trimester (reference category: 25-30 degrees)													
Max <20°	-0.012*	-0.009	-0.002	-0.009	0.001***	-0.001	-0.002	0.000	-0.000	0.002**	0.000	-0.000	0.006
	(0.007)	(0.007)	(0.002)	(0.006)	(0.000)	(0.001)	(0.002)	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)	(0.009)
Max 20°-25°	-0.005	-0.002	-0.001	-0.001	-0.000	-0.001	0.000	-0.001*	-0.000	0.000	-0.001**	-0.001**	-0.004
	(0.004)	(0.003)	(0.002)	(0.002)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.005)
Max 30°-35°	-0.000	-0.001	0.000	-0.000	-0.000	0.000	0.000	-0.000	0.000	0.000	0.000	-0.000	0.000
	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)
Max 35°-40°	-0.001	-0.001	-0.000	-0.000	-0.000	0.000	-0.000	-0.000	-0.000	0.000**	-0.000	-0.000	-0.001
	(0.002)	(0.001)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)
Max 40°+	-0.004	-0.002	0.000	-0.002	0.000	-0.001	-0.001*	-0.000	0.000	0.000	0.000	0.000	-0.003
	(0.004)	(0.003)	(0.002)	(0.002)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.004)
3rd trimester (reference category: 25-30 degrees)													
Max <20°	-0.002	-0.000	-0.000	-0.001	-0.001***	-0.000	-0.001	-0.000	-0.000	-0.000	0.001***	0.001***	-0.006
	(0.003)	(0.003)	(0.002)	(0.003)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.005)
Max 20°-25°	-0.007	-0.002	-0.001	-0.001	-0.000	-0.000	-0.003***	0.000	0.000	-0.000	-0.001***	-0.000	-0.009*
	(0.005)	(0.004)	(0.002)	(0.003)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)	(0.005)
Max 30°-35°	0.002**	0.001	0.000	0.001	-0.000	0.000	-0.000	-0.000	0.000*	-0.000	0.000**	0.000	0.003**
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Max 35°-40°	0.002*	0.001	-0.000	0.001	-0.000	-0.000	0.000	0.000	0.000	-0.000*	0.000	0.000**	0.004
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.003)

Max 40°+	0.010***	0.003	0.001	0.001	-0.000	0.000	0.001	0.000	-0.000	0.001***	0.001***	0.000	0.005
	(0.003)	(0.003)	(0.001)	(0.002)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.007)
Constant	0.838***	0.600***	0.194***	0.341***	0.028***	0.038**	0.047	0.029**	0.013	0.047***	0.012	0.006	0.988***
	(0.118)	(0.076)	(0.037)	(0.073)	(0.006)	(0.017)	(0.028)	(0.013)	(0.024)	(0.013)	(0.014)	(0.011)	(0.167)
Observations	23,756	23,756	23,756	23,756	23,756	23,756	23,756	23,756	23,756	23,756	23,756	23,756	23,756
R-sq	0.283	0.312	0.251	0.226	0.080	0.151	0.121	0.087	0.115	0.091	0.097	0.095	0.168
Demographic controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Month-year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Month-community-sex FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Table D1: Regression estimates on admissions and diagnoses at ages 0-2

Notes: Outcomes are the number of admissions for each diagnosis specific diagnoses. These are based on the primary diagnosis for the admission. We classify these diagnoses based on their ICD-10 'chapter', and we report estimates for all chapters for which at least 1 percent of children in the cohort has been admitted for that reason. We single out 'newborn admission no condition found' as this could conceivably be a precautionary admission e.g. in response to a heatwave around the time of birth. We separately construct a measure of whether an admission was for an infection of any type, based on infection-related ICD-10 codes (across multiple chapters) as set out by Miller et al. (2016): infections are the most common cause of hospital admission among children in the NT.

Appendix E: Additional heat exposure metrics – NAPLAN

	NAPLAN Year 7					
	Average	Reading	Writing	Grammar	Spelling	Numeracy
1st trimester (reference category: 25-30 degrees)						
Max <20°	-0.001	0.001	-0.001	-0.003	-0.001	-0.003
	(0.003)	(0.003)	(0.005)	(0.006)	(0.005)	(0.003)
Max 20°-25°	-0.012**	-0.010	-0.010	-0.010*	-0.019**	-0.010**
	(0.005)	(0.007)	(0.006)	(0.005)	(0.009)	(0.004)
Max 30°-35°	-0.001	-0.001	-0.001	-0.005**	-0.001	0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Max 35°-40°	-0.009***	-0.009***	-0.012***	-0.011***	-0.009***	-0.006***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Max 40°+	-0.005	-0.002	-0.006	-0.007	-0.005	-0.006
	(0.007)	(0.008)	(0.010)	(0.008)	(0.009)	(0.006)
2nd trimester (reference category: 25-30 degrees)						
Max <20°	0.004	0.017**	-0.006	0.006	-0.002	0.005
	(0.008)	(0.007)	(0.010)	(0.009)	(0.013)	(0.007)
Max 20°-25°	0.002	-0.003	-0.006	0.019**	0.005	-0.004
	(0.007)	(0.007)	(0.008)	(0.007)	(0.015)	(0.009)
Max 30°-35°	-0.000	-0.000	-0.001	-0.000	0.000	-0.001
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)
Max 35°-40°	0.000	-0.001	0.003	-0.001	-0.001	0.001
	(0.003)	(0.003)	(0.002)	(0.004)	(0.003)	(0.005)
Max 40°+	-0.006	-0.008*	0.004	-0.006	-0.014*	-0.005
	(0.004)	(0.005)	(0.005)	(0.004)	(0.007)	(0.004)
3rd trimester (reference category: 25-30 degrees)						
Max <20°	0.001	0.007	0.004	-0.005	0.001	-0.001
	(0.006)	(0.005)	(0.009)	(0.005)	(0.010)	(0.013)
Max 20°-25°	-0.010*	-0.007	-0.022**	-0.011	-0.006	-0.005
	(0.006)	(0.010)	(0.010)	(0.008)	(0.015)	(0.015)
Max 30°-35°	-0.002***	-0.001	-0.001*	-0.001*	-0.003**	-0.003**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Max 35°-40°	-0.002	0.001	-0.003	-0.002	-0.002	-0.004
	(0.002)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
Max 40°+	0.007	0.007	0.000	0.007	0.012*	0.007
	(0.005)	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)
Constant	5.851***	5.752***	5.558***	5.957***	6.046***	5.907***
	(0.140)	(0.151)	(0.138)	(0.164)	(0.227)	(0.200)
Observations	4,068	4,068	4,068	4,068	4,068	4,068
R-sq	0.553	0.498	0.468	0.472	0.455	0.492
Demographic controls	Y	Y	Y	Y	Y	Y
Interact 35+ with youngest in year	Y	Y	Y	Y	Y	Y
Month-year FE	Y	Y	Y	Y	Y	Y
Month-community-sex FE	Y	Y	Y	Y	Y	Y

Table E1: Regression estimates on NAPLAN 7 scores, with more granular bands of temperature exposure

	NAPLAN Year 7					
	Average	Reading	Writing	Grammar	Spelling	Numeracy
1st trimester (reference category: 25-30 degrees)						
Max <20°	-0.000	0.007	-0.002	0.002	-0.001	-0.008
	(0.004)	(0.007)	(0.005)	(0.007)	(0.009)	(0.008)
Max 20°-25°	-0.012***	-0.008	-0.009	-0.008	-0.018***	-0.015***
	(0.004)	(0.006)	(0.006)	(0.007)	(0.006)	(0.004)
Max 30°-35°	-0.002	-0.002	-0.002	-0.006*	-0.002	0.002
	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)
Max 35°-40°	-0.007***	-0.007***	-0.012***	-0.010***	-0.007**	-0.001
	(0.002)	(0.002)	(0.003)	(0.004)	(0.003)	(0.003)
Max 40°+	-0.007	-0.003	-0.009	-0.010	-0.008	-0.007
	(0.007)	(0.008)	(0.009)	(0.007)	(0.010)	(0.005)
2nd trimester (reference category: 25-30 degrees)						
Max <20°	0.013	0.028*	0.001	0.018	0.018	-0.002
	(0.012)	(0.015)	(0.011)	(0.013)	(0.020)	(0.009)
Max 20°-25°	0.015*	0.010	0.004	0.033***	0.024	0.002
	(0.008)	(0.009)	(0.010)	(0.012)	(0.015)	(0.009)
Max 30°-35°	0.001	0.001	0.001	0.001	0.002	0.000
	(0.003)	(0.003)	(0.002)	(0.004)	(0.004)	(0.003)
Max 35°-40°	0.004	-0.001	0.009***	0.002	0.004	0.006
	(0.005)	(0.005)	(0.003)	(0.005)	(0.007)	(0.007)
Max 40°+	-0.003	-0.006	0.008	-0.004	-0.010	-0.001
	(0.005)	(0.005)	(0.006)	(0.005)	(0.008)	(0.006)
3rd trimester (reference category: 25-30 degrees)						
Max <20°	-0.016	-0.015	-0.011	-0.020	-0.010	-0.023*
	(0.010)	(0.013)	(0.014)	(0.018)	(0.013)	(0.012)
Max 20°-25°	-0.033	-0.045***	-0.035	-0.038*	-0.014	-0.036
	(0.023)	(0.015)	(0.032)	(0.023)	(0.031)	(0.027)
Max 30°-35°	-0.002**	-0.001	-0.001	-0.002**	-0.004**	-0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Max 35°-40°	-0.002	0.001	-0.003	-0.002	-0.003	-0.004
	(0.003)	(0.003)	(0.004)	(0.004)	(0.005)	(0.005)
Max 40°+	0.006	0.008	0.000	0.007	0.008	0.005
	(0.006)	(0.005)	(0.007)	(0.005)	(0.010)	(0.007)
1st trimester (reference category: 15-20 degrees)						
Min <5°	0.001	-0.012	0.010	-0.009	-0.008	0.025
	(0.006)	(0.008)	(0.008)	(0.007)	(0.009)	(0.015)
Min 5°-10°	-0.000	-0.001	-0.004	-0.004	-0.005	0.014**
	(0.004)	(0.007)	(0.006)	(0.006)	(0.008)	(0.007)
Min 10°-15°	0.004	0.005	0.004	0.003	-0.004	0.012**
	(0.004)	(0.006)	(0.006)	(0.005)	(0.007)	(0.005)
Min 20°-25°	0.002	-0.001	0.005***	-0.001	-0.003	0.008**
	(0.002)	(0.003)	(0.002)	(0.002)	(0.004)	(0.003)
Min 25°+	0.007**	0.003	0.010***	0.005	0.005	0.012**
	(0.003)	(0.003)	(0.002)	(0.003)	(0.006)	(0.004)
2nd trimester (reference category: 15-20 degrees)						
Min <5°	-0.009	-0.013	-0.006	-0.023**	-0.012	0.008
	(0.011)	(0.015)	(0.011)	(0.011)	(0.024)	(0.008)
Min 5°-10°	-0.013	-0.019**	-0.005	-0.015	-0.021	-0.004
	(0.010)	(0.008)	(0.008)	(0.014)	(0.019)	(0.008)
Min 10°-15°	0.007	0.014*	0.004	0.006	0.014	-0.001
	(0.007)	(0.008)	(0.005)	(0.008)	(0.012)	(0.009)
Min 20°-25°	-0.001	-0.002	0.002	0.000	0.001	-0.005
	(0.003)	(0.004)	(0.002)	(0.004)	(0.005)	(0.003)
Min 25°+	-0.000	-0.002	0.003*	0.000	0.000	-0.002
	(0.003)	(0.005)	(0.002)	(0.004)	(0.004)	(0.002)
3rd trimester (reference category: 15-20 degrees)						
Min <5°	0.018	0.023	0.019	0.016	0.009	0.025*
	(0.016)	(0.020)	(0.018)	(0.022)	(0.019)	(0.015)
Min 5°-10°	0.019	0.033***	0.010	0.021	0.002	0.029**
	(0.017)	(0.012)	(0.025)	(0.020)	(0.025)	(0.011)

Min 10°-15°	-0.000	0.001	0.002	-0.001	-0.007	0.003
	(0.005)	(0.004)	(0.005)	(0.007)	(0.010)	(0.007)
Min 20°-25°	-0.001	0.001	-0.000	-0.003	-0.003	-0.002
	(0.002)	(0.003)	(0.002)	(0.002)	(0.004)	(0.003)
Min 25°+	0.001	0.002	-0.000	-0.001	0.000	0.003
	(0.003)	(0.004)	(0.002)	(0.003)	(0.003)	(0.004)
Constant	5.439***	5.654***	4.603***	5.913***	5.987***	5.225***
	(0.366)	(0.399)	(0.276)	(0.487)	(0.968)	(0.430)
Observations	4,068	4,068	4,068	4,068	4,068	4,068
R-sq	0.554	0.500	0.470	0.473	0.457	0.495
Demographic controls	Y	Y	Y	Y	Y	Y
Interact 35+ with youngest in year	Y	Y	Y	Y	Y	Y
Month-year FE	Y	Y	Y	Y	Y	Y
Month-community-sex FE	Y	Y	Y	Y	Y	Y

Table E2: Regression estimates for NAPLAN 7 scores, with maximum and minimum temperature exposure

Appendix F: School Test scores by domain

	Average	Reading	Writing	Grammar	Spelling	Numeracy
1st trimester (reference category: 20-30 degrees)						
Max <20°	0.004	0.007*	-0.001	0.006**	0.003	0.006
	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)	(0.004)
Max 30°-35°	-0.002***	-0.002***	-0.002***	-0.004***	-0.002**	-0.002**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Max 35+	-0.003**	-0.001	-0.002	-0.003**	-0.004***	-0.002
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
2nd trimester (reference category: 20-30 degrees)						
Max <20°	-0.003	-0.000	-0.018***	0.005	-0.006	0.004
	(0.003)	(0.005)	(0.005)	(0.004)	(0.005)	(0.006)
Max 30°-35°	-0.002	-0.003*	-0.000	-0.000	-0.002*	-0.002
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Max 35°+	-0.003	-0.005**	-0.002	-0.001	-0.003	-0.002
	(0.002)	(0.002)	(0.001)	(0.003)	(0.002)	(0.002)
3rd trimester (reference category: 20-30 degrees)						
Max <20°	-0.001	-0.004	-0.002	-0.003	0.000	0.004
	(0.004)	(0.003)	(0.005)	(0.004)	(0.005)	(0.007)
Max 30°-35°	-0.001**	-0.002***	-0.001**	-0.000	-0.001**	-0.001**
	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)
Max 35°+	-0.004***	-0.004***	-0.005***	-0.003***	-0.004***	-0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Constant	3.150***	3.200***	3.274***	3.113***	3.105***	2.974***
	(0.110)	(0.155)	(0.074)	(0.135)	(0.082)	(0.126)
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Month-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Month-community-sex FE	Yes	Yes	Yes	Yes	Yes	Yes
Interact 35+ with youngest in class	Yes	Yes	Yes	Yes	Yes	Yes
Observations	11,681	11,681	11,681	11,681	11,681	11,681
R-sq	0.522	0.432	0.512	0.429	0.412	0.429

Table F1: Regression estimates on school test scores at age 7 (Year 3), by domain

	Average	Reading	Writing	Grammar	Spelling	Numeracy
1st trimester (reference category: 20-30 degrees)						
Max <20°	0.002	0.007**	0.003	0.004	-0.005	0.004*
	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)	(0.002)
Max 30°-35°	-0.003***	-0.003***	-0.002	-0.005***	-0.002**	-0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Max 35+	-0.005***	-0.003*	-0.002*	-0.007***	-0.005***	-0.006***
	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)
2nd trimester (reference category: 20-30 degrees)						
Max <20°	0.002	0.005	0.004	0.001	-0.005	0.004
	(0.005)	(0.008)	(0.005)	(0.006)	(0.005)	(0.006)
Max 30°-35°	-0.001	-0.001	-0.001	-0.002*	-0.001	-0.001
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
Max 35°+	-0.003**	-0.002	-0.004**	-0.003**	-0.004***	-0.001
	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)
3rd trimester (reference category: 20-30 degrees)						
Max <20°	0.007**	0.002	0.013**	0.011***	0.002	0.008***
	(0.003)	(0.005)	(0.005)	(0.004)	(0.003)	(0.002)
Max 30°-35°	-0.002***	-0.001	-0.003***	-0.001**	-0.001	-0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Max 35°+	-0.003***	-0.002**	-0.005**	-0.001	-0.003**	-0.004***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
Constant	4.907***	4.916***	4.691***	5.079***	4.994***	4.881***
	(0.092)	(0.157)	(0.072)	(0.141)	(0.074)	(0.092)
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Month-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Month-community-sex FE	Yes	Yes	Yes	Yes	Yes	Yes
Interact 35+ with youngest in class	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8,013	8,013	8,013	8,013	8,013	8,013
R-sq	0.540	0.476	0.484	0.459	0.420	0.464

Table F2: Regression estimates on school test scores at age 10 (Year 5), by domain

	Average	Reading	Writing	Grammar	Spelling	Numeracy
1st trimester (reference category: 20-30 degrees)						
Max <20°	-0.002	0.002	0.000	-0.005	-0.001	-0.003
	(0.003)	(0.003)	(0.004)	(0.005)	(0.004)	(0.003)
Max 30°-35°	-0.001	-0.001	-0.002	-0.004*	-0.000	0.001
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Max 35+	-0.000	-0.001	-0.001	-0.002	-0.000	0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
2nd trimester (reference category: 20-30 degrees)						
Max <20°	0.002	0.013**	-0.011	0.011	-0.004	0.000
	(0.007)	(0.006)	(0.008)	(0.009)	(0.011)	(0.006)
Max 30°-35°	-0.001	-0.001	-0.002	-0.002	-0.001	-0.001
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)
Max 35°+	-0.004***	-0.003**	-0.004**	-0.005***	-0.005***	-0.004*
	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
3rd trimester (reference category: 20-30 degrees)						
Max <20°	-0.003	0.004	-0.007	-0.009*	-0.001	-0.003
	(0.006)	(0.004)	(0.006)	(0.005)	(0.008)	(0.007)
Max 30°-35°	-0.002***	-0.001	-0.001	-0.001*	-0.003**	-0.003**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Max 35°+	-0.005***	-0.004***	-0.003***	-0.005***	-0.007***	-0.004**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Constant	5.864***	5.732***	5.568***	6.020***	6.072***	5.936***
	(0.083)	(0.107)	(0.093)	(0.118)	(0.168)	(0.164)
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Month-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Month-community-sex FE	Yes	Yes	Yes	Yes	Yes	Yes
Interact 35+ with youngest in class	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,068	4,068	4,068	4,068	4,068	4,068
R-sq	0.552	0.498	0.468	0.471	0.454	0.491

Table F3: Regression estimates on school test scores at age 12 (Year 7), by domain

Appendix G: NAPLAN participation

	Present at NAPLAN test		
	Year 3 (age 8)	Year 5 (age 10)	Year 7 (age 12)
1st trimester (reference category: 20-30 degrees)			
Max <20°	0.001	0.001	-0.006**
	(0.002)	(0.001)	(0.002)
Max 30°-35°	0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.001)
Max 35°+	0.000	0.001**	-0.000
	(0.001)	(0.001)	(0.001)
2nd trimester (reference category: 20-30 degrees)			
Max <20°	0.001	0.000	-0.006
	(0.004)	(0.002)	(0.007)
Max 30°-35°	0.000	0.000	0.000
	(0.000)	(0.001)	(0.001)
Max 35°+	0.000	0.000	-0.000
	(0.000)	(0.001)	(0.001)
3rd trimester (reference category: 20-30 degrees)			
Max <20°	-0.002	-0.002	0.003
	(0.001)	(0.002)	(0.003)
Max 30°-35°	-0.000	-0.000	-0.001
	(0.000)	(0.000)	(0.001)
Max 35°+	-0.000	0.000	-0.001*
	(0.000)	(0.000)	(0.001)
Constant	0.895***	0.851***	0.876***
	(0.040)	(0.062)	(0.048)
Demographic controls	Yes	Yes	Yes
Month-year FE	Yes	Yes	Yes
Month-community-sex FE	Yes	Yes	Yes
Interact 35+ with youngest in class	Yes	Yes	Yes
Observations	14,203	9,919	5,450
R-sq	0.260	0.295	0.359

Table G1: Regression estimates on participation in NAPLAN test

Appendix H: Causal pathways

In the main text we summarise the first column of Table H1: analysis run on our full analysis cohort. However, if the congenital condition does not require admission to hospital, the child may be no more likely to be admitted to hospital than others, despite having their health impacted by heat exposure.

We cannot fully address this in our analysis, as health clinic data are not available for our cohort. However, we can limit our sample to children who have been admitted to hospital – or, even more specifically, to admissions for an infection. In doing so, we are better able to compare like-for-like, as selection into receiving a diagnosis is less likely to be determined by heat exposure itself (given that we found in our above analysis that infection-related admissions were not affected by heat exposure). Even within this subsample, we find no increase in likelihood of being diagnosed with a congenital condition as a result of prenatal heat exposure.

For hospital admissions, we find that when we run our analysis only on children with good health at birth (born at full term, weighing over 2500 grams and with an Apgar score of 8 or over), the estimated treatment effect of exposure to 35+ degrees in the third trimester halves and loses statistical significance, though we continue to see statistically significant estimates of exposure to 30-35 degrees (Table H4). This suggests that children with poor health at birth may be most affected by exposure to extreme heat, but impacts are not confined to this group. If instead of grouping children based on birth outcomes, we run our analysis on children with no labour or pregnancy complications, or for children who were not admitted to a special care nursery, the estimated impacts of heat exposure are no different from in our main analysis.

For school test scores, we find that the estimated effects of heat exposure are similar to our main estimates, or even marginally larger, for children with no health concerns at birth, and for children who were not admitted to a special care nursery (Table H5). However, the effects become smaller and lose

statistical significance among children with no obstetric or labour complications, and for children who were not admitted to hospital ages 0-2. Limiting analysis to these groups severely reduces the sample, and this may be the reason that treatment effects lose significance. The fact that we estimate smaller and less precise effects of heat exposure in these subsamples suggests that it is largely, but not exclusively, the children with perinatal health concerns and worse health in early childhood whose school test scores are affected.

	Ever diagnosed with a congenital condition, ages 0-7		
Sample:	Full cohort	Ever admitted to hospital	Ever admitted to hospital for an infection
1st trimester (reference category: 20-30 degrees)			
Max <20°	-0.001	-0.001	-0.002***
	(0.001)	(0.001)	(0.001)
Max 30°-35°	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Max 35°+	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
2nd trimester (reference category: 20-30 degrees)			
Max <20°	-0.000	-0.001	-0.002
	(0.001)	(0.001)	(0.001)
Max 30°-35°	0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)
Max 35°+	-0.000	-0.000	-0.001**
	(0.000)	(0.000)	(0.000)
3rd trimester (reference category: 20-30 degrees)			
Max <20°	-0.001	-0.001	-0.003
	(0.001)	(0.001)	(0.002)
Max 30°-35°	-0.000	-0.000	-0.001**
	(0.000)	(0.000)	(0.000)
Max 35°+	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
Constant	0.044***	0.092***	0.131***
	(0.012)	(0.019)	(0.034)
Observations	23,756	11,826	8,131
R-sq	0.097	0.128	0.170
Demographic controls	Y	Y	Y
Month-year conceived FE	Y	Y	Y
Month-community-sex FE	Y	Y	Y

Table H1: Regression estimates on probability of hospital diagnosis with congenital condition

This table shows estimation results from a version of equation (1), where instead of hospital admissions, or outcome is an indicator for whether the child has been ever diagnosed with a congenital condition during a hospital admission. This includes both primary and secondary diagnoses. Congenital conditions are defined as any ICD-10 code in chapter Q (Congenital malformations, deformations or chromosomal abnormalities). This differs from Appendix Table C1 because here we include secondary diagnoses, as well as primary diagnoses.

	Admissions ages 0 to 2					
	(1)	(2)	(3)	(4)	(5)	(6)
1st trimester (reference category: 20-30 degrees)						
Max <20°	-0.005	-0.007	-0.006	-0.007	-0.007*	-0.007
	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)
Max 30°-35°	0.001	0.000	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Max 35°+	0.001	0.001	0.002	0.001	0.001	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
2nd trimester (reference category: 20-30 degrees)						
Max <20°	-0.011*	-0.010	-0.011	-0.010	-0.011	-0.010
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Max 30°-35°	-0.001	-0.000	-0.001	-0.000	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Max 35°+	-0.002	-0.001	-0.002	-0.001	-0.002	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
3rd trimester (reference category: 20-30 degrees)						
Max <20°	-0.003	0.003	-0.004	0.002	-0.002	0.002
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Max 30°-35°	0.002**	0.002**	0.002**	0.002**	0.002**	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Max 35°+	0.004**	0.003**	0.004**	0.003**	0.003**	0.003**
	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
Gestational age		-0.955***		-0.951***		-0.882***
		(0.099)		(0.095)		(0.102)
Gestational age^2		0.012***		0.012***		0.011***
		(0.001)		(0.001)		(0.001)
Birth weight		-0.000***		-0.000***		-0.000***
		(0.000)		(0.000)		(0.000)
Apgar 5		-0.130***		-0.123***		-0.077***
		(0.020)		(0.020)		(0.017)
Labour complications (sum)			0.089***	0.084***		0.058***
			(0.016)	(0.015)		(0.016)
Obstetric complications (sum)			0.243***	0.162***		0.130***
			(0.040)	(0.036)		(0.036)

Admitted to special care nursery					0.760***	0.419***
					(0.129)	(0.078)
Constant	0.785***	22.200***	0.673***	21.805***	0.779***	19.475***
	(0.137)	(1.685)	(0.134)	(1.681)	(0.131)	(1.813)
Observations	23,756	23,756	23,756	23,756	23,756	23,756
R-sq	0.282	0.319	0.289	0.323	0.311	0.329
Birth outcomes (birth weight, gest age, Apgar5)		Y		Y		Y
Obstetric and labour complications			Y	Y		Y
Special care nursery admission					Y	Y

Table H2: Regression estimates controlling for health at birth covariates – admissions. This table shows estimation results from equation (1), but including additional covariates for indicators of health at birth, obstetric complications and postnatal care.

	NAPLAN average score - year 3						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1st trimester (reference category: 20-30 degrees)							
Max <20°	0.004	0.005	0.004	0.005	0.004	0.005	0.003
	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)
Max 30°-35°	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Max 35°+	-0.003**	-0.003**	-0.003**	-0.003***	-0.003**	-0.003**	-0.003**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
2nd trimester (reference category: 20-30 degrees)							
Max <20°	-0.003	-0.003	-0.003	-0.003	-0.004	-0.003	-0.003
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Max 30°-35°	-0.002	-0.001	-0.001	-0.001	-0.002	-0.001	-0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Max 35°+	-0.003	-0.002	-0.003	-0.002	-0.003	-0.002	-0.003
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
3rd trimester (reference category: 20-30 degrees)							
Max <20°	-0.001	-0.003	-0.000	-0.002	-0.002	-0.002	-0.000
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)
Max 30°-35°	-0.001**	-0.001	-0.001**	-0.001	-0.001*	-0.001	-0.001*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Max 35°+	-0.004***	-0.004***	-0.004***	-0.004***	-0.003***	-0.003***	-0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Gestational age		-0.041		-0.036		-0.063	
		(0.111)		(0.111)		(0.108)	
Gestational age^2		0.001		0.001		0.001	
		(0.001)		(0.001)		(0.001)	
Birthweight		0.000***		0.000***		0.000***	
		(0.000)		(0.000)		(0.000)	

Apgar 5		0.055***		0.053***		0.040***	
		(0.007)		(0.008)		(0.007)	
Labour complications (sum)			-0.020	-0.021*		-0.014	
			(0.012)	(0.012)		(0.013)	
Obstetric complications (sum)			-0.095***	-0.070***		-0.058***	
			(0.013)	(0.010)		(0.009)	
Admitted to special care nursery					-0.219***	-0.123***	
					(0.029)	(0.021)	
Admissions to age 7							-0.079***
							(0.017)
Constant	3.112***	2.662	3.168***	2.717	3.175***	3.482	3.208***
	(0.116)	(2.200)	(0.115)	(2.197)	(0.129)	(2.099)	(0.096)
Observations	11,681	11,681	11,681	11,681	11,681	11,681	11,681
R-sq	0.523	0.528	0.524	0.528	0.526	0.529	0.528
Controls:							
Birth outcomes (birth weight, gest. age, Apgar5)		Y		Y		Y	
Obstetric and labour complications			Y	Y		Y	
Special care nursery admission					Y	Y	
Admission in first 7 years							Y

Table H3: Regression estimates controlling for early life health covariates – school test scores. This table shows estimation results from equation (2), but including additional covariates for indicators of health at birth, obstetric complications and postnatal care.

	Benchmark	Drop if preterm, low birthweight or low Apgar 5	Drop if obstetric and labour complications	Drop if admitted to special care nursery
1st trimester (reference category: 20-30 degrees)				
Max <20°	-0.005	-0.005	0.002	-0.005
	(0.004)	(0.004)	(0.006)	(0.005)
Max 30°-35°	0.001	0.001	0.001	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Max 35°+	0.001	0.001	0.000	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)
2nd trimester (reference category: 20-30 degrees)				
Max <20°	-0.011*	-0.005	-0.016	-0.003
	(0.007)	(0.005)	(0.012)	(0.006)
Max 30°-35°	-0.001	-0.000	-0.002	-0.000
	(0.001)	(0.001)	(0.002)	(0.002)
Max 35°+	-0.002	-0.002	-0.002	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)
3rd trimester (reference category: 20-30 degrees)				
Max <20°	-0.003	0.002	0.005	-0.001
	(0.003)	(0.009)	(0.005)	(0.006)
Max 30°-35°	0.002**	0.002**	0.003**	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)
Max 35°+	0.004**	0.002	0.004*	0.003*
	(0.002)	(0.002)	(0.002)	(0.002)
Constant	0.785***	0.812***	0.734***	0.802***
	(0.137)	(0.104)	(0.082)	(0.094)
Observations	23,756	19,972	9,898	19,285
R-sq	0.282	0.295	0.369	0.293

Table H4: Subsample regression estimates– admissions ages 0 to 2. This table shows estimation results from equation (1), run on the specific subpopulations indicated in the column heading.

	Benchmark	Drop if preterm, low birthweight or low Apgar 5	Drop if obstetric and labour complications	Drop if admitted to special care nursery	Drop if admitted to hospital ages 0-2
1st trimester (reference category: 20-30 degrees)					
Max <20°	0.004	0.003	0.002	0.005	0.003
	(0.003)	(0.004)	(0.003)	(0.004)	(0.002)
Max 30°-35°	-0.002***	-0.002***	-0.000	-0.003***	-0.002
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
Max 35°+	-0.003**	-0.003***	-0.002	-0.003***	-0.000
	(0.001)	(0.001)	(0.003)	(0.001)	(0.002)
2nd trimester (reference category: 20-30 degrees)					
Max <20°	-0.003	-0.002	-0.003	-0.002	-0.008
	(0.003)	(0.004)	(0.006)	(0.004)	(0.005)
Max 30°-35°	-0.002	-0.000	-0.001	-0.000	-0.003
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Max 35°+	-0.003	-0.002	-0.003	-0.002	-0.005*
	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)
3rd trimester (reference category: 20-30 degrees)					
Max <20°	-0.001	-0.007	0.007**	-0.004	0.000
	(0.004)	(0.006)	(0.003)	(0.005)	(0.003)
Max 30°-35°	-0.001**	-0.001**	-0.000	-0.001***	-0.001
	(0.000)	(0.001)	(0.002)	(0.001)	(0.001)
Max 35°+	-0.004***	-0.005***	-0.003	-0.005***	-0.002
	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)
Constant	3.112***	3.152***	3.005***	3.119***	3.221***
	(0.116)	(0.122)	(0.217)	(0.101)	(0.061)
Observations	11,681	9,901	5,051	9,517	6,178
R-sq	0.523	0.521	0.513	0.523	0.455

Table H5: Subsample regression estimates – average school test scores at age 8. This table shows estimation results from equation (2), run on the specific subpopulations indicated in the column heading.

Appendix I: Effect sizes in context

Study	Exposure	Outcome	Estimate in SDs	Context
This study	Additional day of max 30-35C in third trimester	School test scores age 8	-0.0007	NT, Australia
		School test scores age 10	-0.0015	
		School test scores age 12	-0.0016	
	Additional day of max 35C+ in third trimester	School test scores age 8	-0.0028	
		School test scores age 10	-0.0023	
		School test scores age 12	-0.0040	
Cho 2017	Additional day of max 34C+ in summer (May-Sept) before test (Nov)	College entrance - English	0.0064	South Korea
		College entrance – Maths	0.0042	
Park et al 2021	Additional school days above max 26.7C in year before test	School test scores grades 3-8	0.0007	USA
Park et al 2020	Additional day of max 37C+ in prior year	Grades 10-11 (PSAT)	-0.00098	USA
	Additional day of max 32-37 in prior year		-0.00064	USA
Park et al 2022	Temperatures on exam day of 32C or above	High school exit test	0.13	USA

Chapter 5: Income and immunity: The consequences of a pre- and neo-natal income shock on childhood infection risk

Abstract

We analyse the impact of a temporary reduction in income, occurring in utero and early infancy, on health in early childhood. We identify the income reduction through the gradual rollout of the so-called 'income management' policy in Aboriginal communities in Australia's Northern Territory in 2007. This policy changed the delivery method of welfare payments but not their value – however, implementation challenges meant that many families did not receive their payments on time. Using linked administrative data, we find that children who were exposed to the policy rollout in utero or in their first three months of life (the 'fourth trimester') were at higher risk of severe infection requiring hospitalisation. They spent, on average, 4.7 more days in hospital between birth and their 8th birthday. Most of this impact is concentrated in hospitalisations for infection, which increased by 23 percent. These admissions are driven by a range of infection types: bacterial, viral and respiratory. We link our findings to the 'immune programming hypothesis', i.e. maternal stress and poor nutrition during key stages in immune system development can permanently weaken the child's immune system. Our findings highlight the importance of attention to key phases in child development when designing policies that affect households' financial resources, even temporarily.

5.1 Introduction

The conditions that children are exposed to in utero and early infancy can have long-lasting impacts on their health and development. Research by economists and other social scientists tells us that availability of financial resources in utero and early childhood are key determinants of a child's environment, and can therefore have long-term impacts on measures of adult human capital, including health, education and earnings potential (see Almond, Currie and Duque (2018) for a review).

Most research on this topic to date has studied the introduction of major elements of modern social security systems which lead to large and sustained changes in family income. The evidence shows that social security policies are a powerful tool to invest in human capital. But what is more relevant to policymakers in high income countries today is not the possible impact of major overhauls to entitlements, which happen rarely, but the impacts of changes in conditionality and administration of social security.

The US National Scientific Council on the Developing Child (2023) argues that policymakers should embrace an 'early childhood lens' in a broad range of policy domains, beyond those traditionally considered within the realms of childhood policy. In this chapter, we analyse one such area that policymakers should view through an early childhood lens: administration and implementation of social security. Because these types of incremental policy changes are often accompanied by implementation challenges, they may have unanticipated impacts on child development.

Implementation challenges are often glazed over as temporary inconveniences. However, for children affected by these challenges, the impacts are not necessarily temporary. Medical and neurological research shows that during key developmental stages in pregnancy and early childhood, children's brains and bodies are highly plastic. Shocks to family income, which can impact children's nutrition and parents' stress levels, even if temporary, can impact children's developmental trajectories. This can lead to long-term

effects on children's health and cognitive function (Adkins, Leclerc, and Marshall-Clarke 2004; Melville and Moss 2013; Anand and Scalzo 2000; Marques et al. 2013).

In this chapter, we analyse a change to the administration of means-tested benefits, and the impact of this change on health throughout early childhood. Australia's income management policy was introduced in 2007, and targeted at Aboriginal communities in the Northern Territory – a population with very low average incomes, who are four times more likely to face poverty than the general population (Altman 2017). The policy did not change households' benefit eligibility or the value of their payments, but imposed conditions on what recipients could spend their money on. To implement these conditions, the policy restricted where recipients could shop, which caused stress and reduced their purchasing power. More acutely, challenges in implementing these restrictions meant that some households experienced a temporary reduction in income for weeks or months. We study the impact of exposure to the policy change in utero or as a newborn, on children's health to age 8.

This setting provides a useful test case to understand the general impacts of changes to social security administration – and accompanying implementation challenges – on children in very low-income households. Aboriginal communities in Australia's Northern Territory are unique in many ways (discussed below in Section 2.1) and this episode is particularly instructive to Australian policymakers. However, recipients' experiences of adapting to changes in administration of social security, learning how to navigate new government systems, and interacting with poorly implemented policies – is by no means unique to Aboriginal communities. We know, for example, that when the US SNAP program moved from physical food stamps to its current electronic benefits card, recipients' food insecurity increased temporarily (Lovett and Xue 2017). Similarly, in the UK's transition from paper vouchers to a debit card for its Healthy Start food voucher scheme, many recipients were temporarily unable to access their funds (Defeyter et al. 2022).

We therefore already know that these types of policy changes can be disruptive due to implementation challenges. Our analysis of the introduction of income management allows us to further understand the impacts of such disruption not just on consumption, as detailed in studies from the US and the UK, but on child development. We can do this because income management was rolled out gradually, across 73 Aboriginal communities over a 13-month period, creating a natural experiment that makes it possible to estimate the causal effect of the policy change. In earlier research, Doyle, Schurer and Silburn (2022) and Cobb-Clark et al. (2021) use the same administrative datasets and identification strategy to analyse the short-term effects of the policy introduction: they find reductions in birthweight and in school attendance, respectively. In this chapter, we analyse the policy's longer-term impacts on children's health trajectories. We isolate the effect of this policy change on children's health to age 8, if the change occurred when they were in utero or as a newborn – key stages in early development – relative to exposure later in childhood.

We find that children who were exposed to the introduction of income management in utero or in their first three months of life appear to be less healthy in early childhood. They spend on average 4.7 more days in hospital from birth to their 8th birthday (a 40 percent increase from baseline). Most of this impact is concentrated in hospital admissions for infection, which increased by 23 percent. These admissions are driven by multiple types of infections: bacterial, viral and respiratory.

These adverse effects persist but become smaller in magnitude and lose statistical significance for communities receiving income management later in the rollout period. This informs our interpretation that, while the income management policy itself did not improve outcomes, it was the temporary income reduction resulting from initial implementation challenges that drives the large effect we find in the rollout period overall.

Hospital admissions are not necessarily an indication of worse health but, in our case, they appear to be. It is possible that admissions could increase if hospitals were used as a substitute for primary care, with no change in underlying infection rates. This is not the

case; analysing clinic data for a subsample of children for whom primary care data are available, we find an increase in clinic visits for infections. This implies that it is increased illness, and not a substitution away from primary care, driving our results for hospital admissions in the full sample.

In addition to these headline effects, our analysis provides two insights into how these effects develop through infancy and early childhood, which make important contributions to the literature on the early origins of lifetime human capital.

First, the pattern of effects is consistent with the 'immune programming hypothesis' (Marques et al. 2013). This hypothesis suggests that poor nutrition and maternal stress in utero and infancy can lead to a permanent weakening of the immune system, resulting in higher susceptibility to infectious disease later in life (Bermick and Schaller 2022; Marques et al. 2013; Nielsen et al. 2011). We find impacts of the policy change on viral, bacterial and upper respiratory infections through to age 8. At earlier ages, we find effects on soft tissue and lower respiratory infections. We also find an increase in asthma diagnoses – another condition resulting from a weaker immune system. Because the impacts are not concentrated in a single type of infection, our findings do not appear to reflect recurring infections or continued complications from a specific health issue. Instead, these patterns, of an increase in multiple types of infections, suggest that this cohort of children are more susceptible to infections in general, which fits the immune programming hypothesis.

Second, we interrogate the relationship between birth outcomes and health in childhood. Doyle, Schurer and Silburn (2022) found that income management reduced birthweight for babies in utero during the policy rollout. In this chapter, we analyse the longer-term health outcomes for this cohort of children as they grow up; we find that a prediction of child health outcomes based on the policy's initial impact on birthweight would severely under-estimate the true effect. However, when we combine several measures of newborn health, our models can explain around a third of the treatment effect on hospital admissions. Therefore, while birthweight alone may be of limited use as an early indicator

of how adverse in utero experiences affect childhood health, a combination of measures perform better as predictors. This finding underscores the shortcomings of birthweight as a primary measure of health at birth (Conti et al. 2020). It also suggests that further research to identify the right combination of measures – looking beyond the more common measures of birthweight and preterm birth – may make it possible to predict which children are at greatest risk of adverse effects from pre- and peri-natal shocks, and hence to target early support towards those children.

Our estimated effect sizes are large in magnitude. This may be because, as Torche (2019) describes, children already exposed to disadvantages experience more severe effects from early life insults. This policy change affected Aboriginal people living in remote communities in Australia's Northern Territory, a highly disadvantaged population. There is good reason to suspect that the effects of an income reduction may be accentuated in remote Aboriginal communities relative to the rest of the country, because of low baseline incomes, distance to comprehensive healthcare, limited access to fresh, affordable and nutritious food, and exposure to disease and crowded housing.

The historical and geographical context of Aboriginal communities in the Northern Territory is unique, but many of the issues Aboriginal communities face are not. Most advanced economies have pockets of disadvantage, with communities who also experience crowded housing, and limited access to fresh food and healthcare. While the effects of such a policy change in a higher-resource setting may be smaller on average, the impacts experienced by the most disadvantaged children may be similar to those we find. Our findings show that even within a high-income country, children in under-resourced communities are at high risk of worse health when they are exposed to perinatal income reductions.

While our findings relate specifically to income management policy, the general finding is that implementation challenges matter. Even a temporary change in households' financial resources can have long-term effects if it occurs during key stages in childhood development. These effects are often unseen because they do not occur immediately, but

they have tangible costs. An additional 4.7 days in hospital per child in this cohort would have costed an estimated AU$4.5 million to the NT public hospital system over the eight years we analyse. Policymakers, of course, do not plan on such implementation challenges – but they can take steps to avoid these unintended effects. The cost-savings from doing so are potentially large.

5.2 Policy and background

5.2.1 Aboriginal communities in the Northern Territory of Australia and child health

The policy we analyse in this chapter was targeted specifically at Aboriginal people living in Aboriginal communities in the Northern Territory (NT) of Australia (hereafter referred to as NT communities). Around half of the Aboriginal people in the Northern Territory live in very remote communities (Australian Bureau of Statistics 2023), many located long distances from towns and cities. The characteristics of NT communities vary greatly, with populations ranging from 20 to 2,000 people. The policy did not apply to non-Aboriginal people,[46] or to Aboriginal people living in towns and cities. However, it did apply to Aboriginal people living in 'town camps' – that is, Aboriginal communities located on the outskirts of towns or cities.

Because of their remoteness, NT communities have limited access to healthcare services and infrastructure, and poor quality housing (Foster and Hall 2021). Residents also face high prices for food and other essentials. All of these factors contribute to poor health outcomes.

[46] Very few non-Aboriginal people live in NT communities, and those who do could apply for an exemption from the policy.

Children living in NT communities face particularly poor health outcomes. Despite the long distances that residents of NT communities must travel to reach hospitals, children born in NT communities have, on average, just under 3 hospital admissions between ages 0-7, compared with 2 admissions for Aboriginal children not born in NT communities, and 0.7 admissions for non-Aboriginal children (Figure 1). The most common reason for admission is infection. In our analysis, we estimate a change from this already high base.

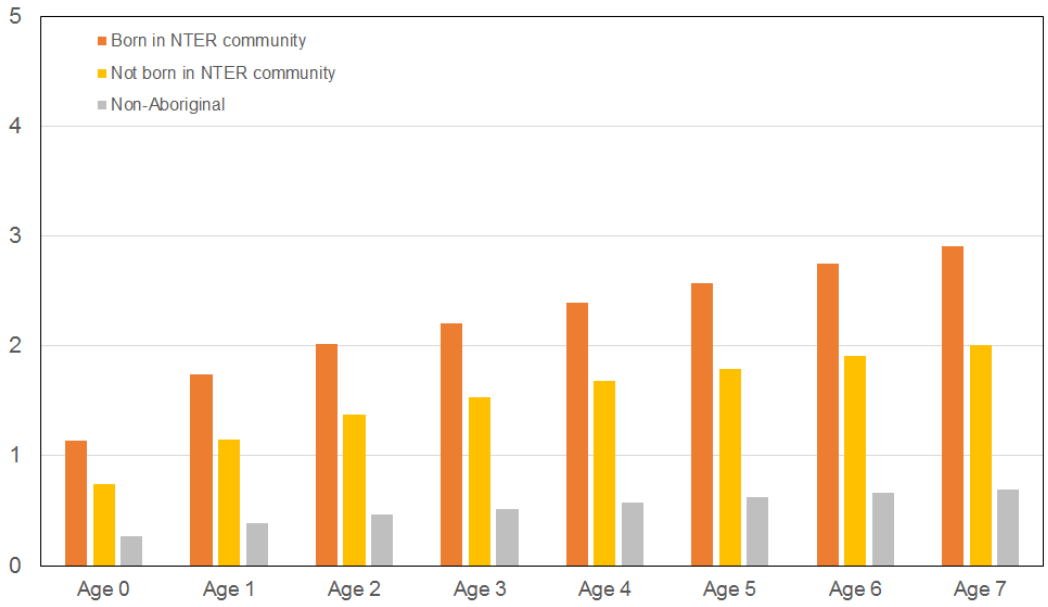


Figure 1: Cumulative number of hospital admissions by age, for children born in the NT from 2000-2007.
Note: The bars show averages for Aboriginal children in 'NTER' communities (those we study in this chapter, which were targeted by the income management policy), Aboriginal children born in towns and cities (not affected by the income management policy), and non-Aboriginal children. Authors' calculations based on administrative data.

By focusing on hospital admissions, we isolate only the most severe cases of illness that require treatment in hospital. However, illness and infection not requiring hospital admission are also common in NT communities: Bar-Zeev et al. (2012) find that infants in remote communities visit their clinic on average once every two weeks, most commonly presenting with respiratory, skin and gastrointestinal conditions.

5.2.2 The income management policy

In Australia, the federal government pays social security benefits (often referred to as 'welfare'), including for unemployment, parenting payments, disability support and the aged pension. There is no time limit on claims, meaning that households may receive these payments indefinitely, provided they remain eligible.

Income management represented a change in the way these payments were administered. Most people who receive government transfers within Australia have the funds deposited into their bank account, and can spend the money however they choose. But this changed under income management, which imposes conditions on how funds should be spent. Under income management in the NT, 50% of each regular payment was deposited into the recipients' bank accounts, as usual. But the remaining 50% of regular payments, and 100% of lump-sum payments, was 'quarantined'. The quarantined money was held in an account with Centrelink, the government agency that administers the payments. Those quarantined funds were earmarked to be spent on essentials, such as food, utility bills and housing.[47] They could not be withdrawn as cash, and could not be spent on alcohol, tobacco, gambling services, or pornography. To enforce these spending restrictions, payments were made directly from the income managed account to suppliers and stores on behalf of the recipient.

Importantly, the intention of income management was to change only the way that payments were delivered, and not the value, frequency, or eligibility rules. The policy was compulsory for benefit recipients living in NT communities, but did not apply to people outside of those communities. Statistics reported by the Australian Institute of Health and Welfare (2010) suggest that around three-quarters of adults in NT communities received benefits during the income management rollout period. We do not

[47] 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.*

have data on the average value of payments within NT communities, but in general for low-income households receiving benefits in Australia, these make up around 80 percent of total household income, and at the time of the introduction of income management, were worth around A$210 per week.[48] In addition, at the time of the policy change, parents were eligible for a universal A$4,000 'Baby Bonus' lump-sum payment after the birth of a child, 100% of which was quarantined.

Income management was introduced as part of what was called the Northern Territory Emergency Response. This was a package of policies introduced from mid-2007, in response to a report documenting child abuse and neglect in Aboriginal communities (Wild and Anderson 2007). The policies were wide-ranging, though income management was a central part of the package, and was rolled out on a different timeline to other policies (see Doyle, Schurer and Silburn (2022) for details).

The intention of income management was to ensure that a higher share of payments is spent on essentials that benefit children. However, the weight of evidence suggests it is unlikely that this was achieved. While comprehensive spending data are unavailable, Brimblecombe et al. (2010) analyse grocery spending patterns across 10 NT communities and find no change following the introduction of income management. In the immediate aftermath of the policy introduction, Doyle, Schurer and Silburn (2022) find a decrease in birthweight, and Cobb-Clark et al. (2021) find a reduction in school attendance: neither of which are consistent with an increased provision of essentials to children.

The reason that the policy did not have the intended effect may be that it does not impose a binding constraint to most households. While we do not have spending data from NT communities, we know that most low-income households in the NT already spent more than 50% of their income on essentials (Doyle, Schurer, and Silburn 2022).

[48] This is based on Australian Bureau of Statistics estimates of Household Income and Income Distribution for 2005-06, Table 6523.0.55.001. These data relate to the national average – it is likely that payments are somewhat higher and other income lower in our context, as there are fewer employment opportunities in NT communities.

Implementation challenges

The policy rollout began in September 2007, three months after it was announced in July 2007. The complexity of the policy and the short timeframe meant that there was little time to anticipate and overcome implementation problems. These implementation problems led to a sharp, but temporary, reduction in some recipients' income – and a more modest reduction in purchasing power for other recipients.

The reduction in income for many households was due to administrative requirements. Recipients were required to meet with a caseworker to jointly decide how the quarantined portion of their funds will be allocated. Before this meeting, quarantined funds would accrue in recipients' income management account, and they could not be accessed until after meeting with a caseworker. We do not have data on when individual recipients met with their caseworker, so we cannot say exactly how many recipients this rule affected, and for how long. But aggregate data suggest that in the early stages of the rollout (November 2007), around one-fifth of recipients were temporarily unable to access their quarantined funds for this reason. In the first few months of the rollout, unallocated funds accounted for 30 to 50 percent of all quarantined funds (Figure 2). This situation improved throughout the rollout but was not entirely resolved. By the end of the policy rollout in October 2008, around 1 percent of recipients had unallocated funds (Australian Institute of Health and Welfare 2010, 24). While this is a small share of recipients, it still represents a large amount of money withdrawn from circulation in affected communities: around A$4 million by March 2009 (Australian Institute of Health and Welfare 2010, 30).

For recipients who did not meet with a caseworker within 13 weeks of the policy's introduction in their community, the income shock was more severe: they had their full payment (not only the quarantined portion) suspended. This affected around 8½ percent of all income managed individuals during the rollout period (Australian Institute of Health and Welfare 2010).

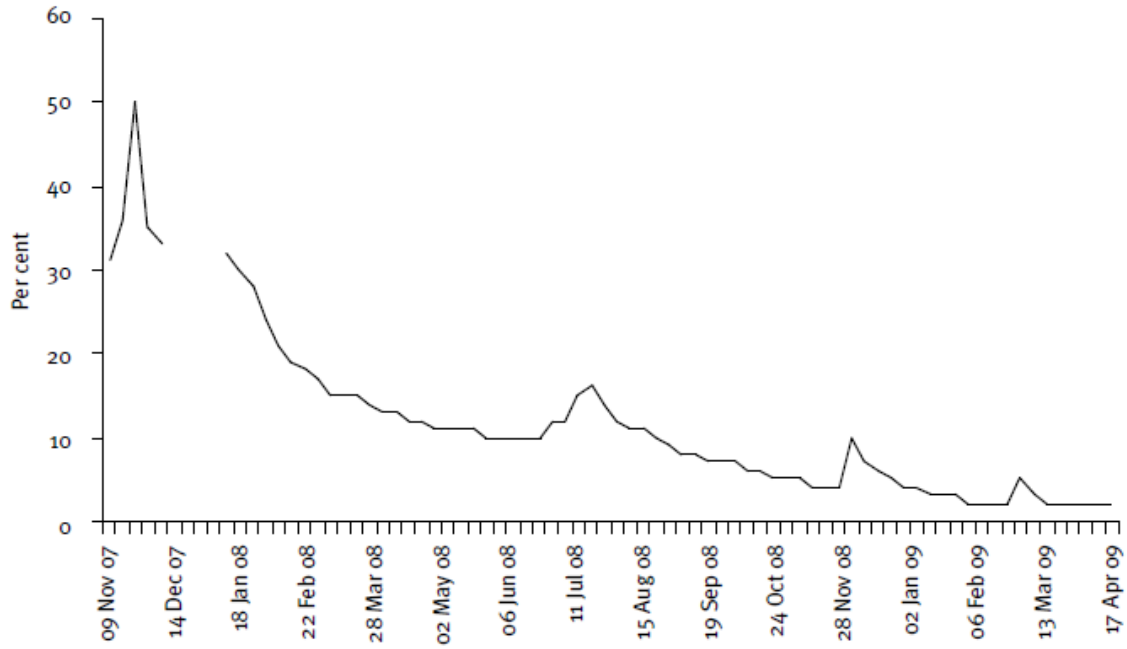


Figure 2: Unallocated funds as a proportion of all quarantined funds to 27 March 2009.

Note: Data unavailable from 14 December 2007 to 10 January 2008. Source: Figure 6 in Australian Institute of Health and Welfare 2010. Data

Recipients who did meet with a caseworker did not experience the same disruptions in their payments. But they experienced less acute reductions in their purchasing power. This is because during the rollout period, any changes to allocations of quarantined funds had to be made manually by Centrelink. For example, if recipients wished to travel into town, they would need to organise this with a caseworker, and arrange for their funds to be sent as credit to the stores they planned to visit. This introduced frictions, and meant that recipients could not 'shop around' for the lowest prices.[49]

[49] These arrangements applied for the first year that the policy operated. In late 2008, Centrelink introduced a payment card (the 'BasicsCard') which could be loaded with quarantined funds, and spent at any participating retailer. This did not eliminate frictions for recipients, but alleviated them substantially. To accept payments from the BasicsCard, retailers must receive approval from the government agency that provides income support. Retailers that have not received approval are blocked from accepting the BasicsCard at their card terminal – any attempt to the card at an unapproved card terminal would be

Heterogeneous policy effects

Among households receiving regular benefits, there are two different types of 'treatment'. First, as described above, around one in five households had half or more of their transfer income withheld as 'quarantined' funds they could not access, for between a few weeks and a few months. Second, the remaining households were affected less acutely, but experienced inconvenience and lower purchasing power. In addition to budgeting difficulties, both groups likely experienced substantial stress during the rollout. As Yu, Duncan and Gray explain (2008, 20), 'people were required to master new, complex and often changing procedures with a minimum of information or explanation'.

Unfortunately, data on social security receipt are not available. The information we have on unallocated funds comes from aggregated reporting (Australian Institute of Health and Welfare 2010) and therefore we do not know which of the children in our population were most and least acutely affected. However, we use estimation methods which allow for heterogeneous treatment effects.

5.3 Conceptual framework

We know that the treatment we analyse reduced birthweight (Doyle, Schurer, and Silburn 2022), but its impact on health in early childhood is an open question. It is possible that children with lower birthweight were able to catch up to their peers, with no longer-term impacts. But equally there may be children within this cohort who were affected by the treatment even if it did not affect their birthweight. We know other in utero shocks can affect health and development despite no impact on standard measures of health at birth (e.g., see Schwandt (2019)).

declined. Most major retailers (supermarkets, department stores, clothing store chains, petrol stations) were approved to accept it. If a retailer sells both restricted and unrestricted items and has been approved to accept the BasicsCard, then the retail staff must enforce the restrictions. That is, if a customer tries to purchase alcohol using their BasicsCard, the retail staff must refuse to process the transaction.

Our analysis draws on the fetal origins hypothesis, which states that the prenatal environment can have long-term impacts on children's health and development, and on their later-life human capital. This hypothesis is supported by substantial empirical evidence (Bharadwaj, Eberhard, and Neilson 2018; Currie and Rossin-Slater 2015; Risnes et al. 2011).

Research on the fetal origins hypothesis originally focused on the epigenetic mechanism of 'fetal programming' of chronic disease, which is the idea that experiences during pregnancy can switch 'on' the genes that increase risk of metabolic disease later in life (Almond and Currie 2011). This is the mechanism that Barker (1998) proposed in explaining the link between low birthweight (implicitly, poor maternal nutrition during pregnancy) and the child's latent risk of chronic diseases in middle age.

However, since this literature first emerged, our understanding of the link between early life conditions and adult human capital has expanded in two ways.

First, either implicitly or explicitly, in the mechanisms linking early experiences to adult human capital. The fetal origins literature finds impacts of in utero exposures on a wide range of outcomes, from chronic disease in middle age, to employment outcomes, to behavioural patterns, to school test scores (Chorniy, Currie, and Sonchak 2020; Almond, Currie, and Duque 2018). To produce this wide range of outcomes, it seems likely that a diverse set of mechanisms, beyond 'fetal programming' of metabolic syndromes, are at work. We discuss this further in Section 7.

Second, while this body of literature initially focused on the prenatal environment, it now extends to cover periods after birth. As Torche (2019) describes, research in neurobiology and developmental psychology has posited the more general idea of 'sensitive' or 'critical' periods during childhood, in which the brain and nervous system develop, and insults experienced during these periods can permanently impair development. Eshaghina et al. (2024) support this empirically, finding evidence of different sensitive periods for different developmental outcomes for children in Denmark. The idea that there are

multiple 'sensitive' periods beyond pregnancy helps to unify empirical evidence in the social sciences, which finds long-term impacts on human capital from both the in utero environment and the childhood environment.

In this chapter, we consider exposure to the policy change in utero or as a newborn. Of course, we know that there were likely impacts of the introduction of income management on older children as well. A number of studies find that changes to family income throughout childhood have lasting impacts on outcomes in adulthood (Braga, Blavin, and Gangopadhyaya 2020; Bailey et al. 2020). However, because the in utero and newborn periods are particularly 'sensitive periods' (Adkins, Leclerc, and Marshall-Clarke 2004), we expect that effect may be larger or longer-lasting for children exposed during these periods. For example, Hoynes, Schanzenbach and Almond (2016) find that the impact of an effective increase in income (US food stamps) on adult health was largest if introduced before birth, but that there were still effects if it was introduced in childhood. We may expect a similar pattern for an income reduction. Our detailed data, which contains exact date of birth and date of the policy change in each community, allows us to also isolate exposure in the newborn period (the so-called 'fourth trimester'), as well as during pregnancy.

5.4 Data and sample

We use linked administrative data on children born in the NT, retaining only the subpopulation of all children whose mother lived in an NT community or town camp at the time of birth. Almost all of these children are Aboriginal (for summary statistics of main variables used in analysis, see Appendix Table A1).

5.4.1 Sample selection

We limit our sample to children born during the income management rollout period, and in the preceding year. We define our sample based on date of conception, given the

potential for sample definition dates based on date of birth to bias analysis (Strand, Barnett, and Tong 2011a). We therefore include all children conceived between July 2005 and April 2008, giving us a sample size of 2,279 children.

We focus on hospital inpatient admissions. A patient may be admitted to hospital directly if they are referred for hospital care, or may be admitted after visiting the Emergency Department. We consider all inpatient admissions from ten days after birth through to age 7.[50] Ending our analysis at age 7 is due to data availability: we have hospital admission data through to mid 2017, by which point the latest-born children in our cohort would have just had their 8th birthday.[51]

We exclude cases where the child was admitted for a non-medical reason (e.g. accompanying family members). We also exclude duplicated 'statistical admissions' and transfers, where there are multiple admission records for a single episode of care. In these cases, we count only one admission, but retain all diagnoses.

Public hospital care in Australia is provided at no cost to the patient, and hence there is no financial barrier to receiving care. While Australia also has a parallel private healthcare system, there is just one private hospital in the NT (in Darwin), which is co-located with a public hospital, and receives very few paediatric admissions.

We also use data from primary care clinics, although these data do not cover the universe of primary care encounters. They are available for two-thirds of NT communities, though some are smaller communities, and hence only around half of the children in our

[50] We exclude the birth admission, and admissions within 10 days of when the child was discharged from their birth admission. We do this because we do not want our results to reflect the initial impact of the policy on admissions to special care nurseries or neonatal intensive care units. We know from Doyle, Schurer and Silburn (2022) that the policy reduced birthweight, and that children with low birthweight are automatically admitted to a special care nursery. We do not include admissions within 10 days of birth because, in the NT, women who live remotely travel long distances to hospital, and stay close to the hospital for some time after birth for monitoring. This means the mother and baby are more likely to be readmitted to hospital for minor concerns. However, we run a version of our model including admissions in the first 10 days: estimates are slightly larger in magnitude but otherwise little changed (see Appendix Table E6).

[51] While we are able to analyse admissions through to age 7, most admissions occur in a child's first two years (see Figure 1), and hence the exact end point makes little difference to our conclusions.

population were born in these communities.[52] These data are available from 2010 onwards, therefore we are unable to observe clinic visits for our cohort from birth. Instead, we observe admission for children at ages 3 to 7. For this analysis, we limit our sample to children who were born in the communities for which clinic data are available, giving a sample size of 1,191 children.

5.4.2 Outcome measure

Our main outcome measures are inpatient admissions for any reason (ever, number of admissions and number of days spent at hospital), and admissions for infection (as a primary diagnosis).[53]

To identify diagnoses, in particular infections, we use codes from the International Statistical Classification of Disease and Related Health Problems (ICD-10), encompassing diagnoses from multiple ICD chapters. We follow Miller et al. (2016), who define six sub-categories of infection: bacterial, viral, upper respiratory, lower respiratory, soft tissue and gastrointestinal.[54] We amend the definitions slightly, to include additional infection-related diagnoses which were present in our data, but not in Miller et al.'s list (see Appendix G for full list of codes).

Primary care codes differ from the ICD-10 codes used in the hospital admission data, which means we cannot construct the same breakdowns of infection by type across the hospital and primary care datasets. Instead, we identify infections as all diagnoses in the primary care data classified as an 'infection' in the summary published by the WONCA

[52] There are an additional 4 community clinics which do not contribute data to the NT Government's clinical records system. The remaining communities do not have their own primary care clinic.

[53] In this definition, we also include admissions where a primary diagnosis is deemed to be caused by an infection. For example, one of the most common primary diagnoses is cutaneous abscess (ICD-10 code L02) and, in the majority of cases, the secondary diagnosis is 'streptococcus and staphylococcus as the cause of diseases classified to other chapters' (ICD-10 code B95).

[54] Otitis media and other ear infections are included in upper respiratory infections, and skin infections are included in soft tissue infections. Miller et al's definition also includes genitourinary infections, however we exclude these from our analysis as the number of admissions for genitourinary infection among children aged 0-7 in our population is negligible.

International Classifications Committee (2018), and we identify two common subcategories: respiratory and soft tissue (eye or skin) infections.

5.5 Empirical strategy

5.5.1 Validity of the policy experiment

Income management was introduced gradually between September 2007 and October 2008, with 32 distinct rollout dates. The rollout was done at the community level.

Our causal identification is based on this staggered rollout. Our analysis includes only children who were born in NT communities, where income management was eventually introduced, meaning all children were exposed to income management at some point. Provided that the rollout timing is exogenous, and that there were parallel trends in early- and late-adopting communities before the policy implementation, we can assign a causal interpretation to our estimates.

Before the policy was introduced, the federal government set out a timeline for when each community would receive income management. The reason for the gradual rollout was that Centrelink staff needed to visit communities before income management began: they did so in two teams, with one team sequentially visiting all communities in northern NT and another team visiting all communities in southern NT. While visiting, staff would meet with residents to give them an opportunity to allocate their income managed funds, and would issue the local community store with a license to accept income managed funds (Australian Institute of Health and Welfare 2010).[55]

[55] Stores needed to meet conditions of selling sufficient fresh food and essentials. However, when stores did not meet these conditions, most were still allocated 'provisional' licenses, meaning this did not unduly delay the rollout.

We do not have information on the logic behind the rollout timeline. Doyle, Schurer and Silburn (2022) suggest the rollout may be, at least partly, scheduled to avoid visiting communities prone to flooding during the wet season. However, they show that there were no *a priori* differences in newborn health, or community demographics between communities that received income management earlier or later. They also show that there was no clear geographic pattern in the rollout, in terms of whether larger communities, more remote communities, or town camps systematically received treatment earlier or later.

In Appendix Table B1, we repeat Doyle, Schurer and Silburn's (2022) analysis on differences between earlier and later-adopting communities, based on our main sample and outcome measures. We find no systematic differences in health or community characteristics for earlier- or later-adopting communities, which suggests the rollout was indeed exogenous to any pre-existing differences across communities in child health and hospitalisations.

To formally test for parallel trends, we test the joint statistical significance of the pre-trends in the lead up to the policy introduction, in the way recommended by Borusyak, Jaravel and Spiess (2024). We find that the parallel trends assumption holds (see Appendix Table B2, and Figure 3 below).

However, over several years leading up to the policy introduction, we observe a trend decrease in the number of hospital admissions per child (see Appendix Figure A1). This trend may bias our estimates downwards; if the trend continued into our analysis period, the comparison group (more likely to be born earlier in the sample period) would have more admissions than those born later. We revisit this point in our robustness tests, in which we test the sensitivity of our findings to the length of the pre-implementation period included in our sample.

5.5.2 Definition of treatment status

In our analysis, we define a child as 'treated' if income management was introduced in the mother's community of residence during pregnancy, or on their first 90 days of life. This treatment is defined based on the child's exact date of birth and the exact date the policy was introduced in their community.

The comparison group consists of children living in NT communities, but who were born more than 90 days, and up to 30 months,[56] before income management was introduced in their community. To be clear, all children in our sample were affected by the policy change at some time before they were 30 months (two and a half years) old. the treatment group were affected during pregnancy or as a newborn, and the comparison group were affected from when they were 3 months old or later. Appendix Table B3 presents the number of children in our analysis sample in the treatment and comparison groups, by quarter of birth.

We define the 'cutoff' point for treatment as 90 days old. This covers the so-called 'fourth trimester' of pregnancy (the newborn period). While much research confirms that the in utero period is a key stage in development (Almond, Currie, and Duque 2018), research has also demonstrated that the newborn stage is a particularly sensitive period, in which nutritional interventions can have large impacts. For instance, Gunnsteinsson et al. (2022) find that providing a vitamin supplement soon after birth improves growth and reduces fevers in infancy. This fits with the broader scientific literature which states that, alongside the in utero period, the newborn period is the most active stage of epigenetic imprinting (Simeoni et al. 2014), in which nutrition (Wiedmeier et al. 2011) and maternal stress (Bermick and Schaller 2022, 303) can have long-term impacts on developmental trajectories. As described by Georgieff et al. (2015), while there are many potential

[56] Appendix Table E5 shows the sensitivity of our estimates to changes in the length of the pre-period we include (and hence in the age of children in the comparison group at the time of the policy introduction). We find that including no pre-period – meaning the treatment group are instead compared with children born 3 to 18 months at the time of the policy change – leads to somewhat larger and strongly statistically significant estimates.

'sensitive periods' in early childhood development, the fetal and neonatal period are times of particularly high nutrient requirements, given rapid metabolism at these stages. Therefore, poor nutrition during these periods creates particularly high risks to children's health and development.[57]

5.5.3 Estimation model

We use the following generalised difference-in-difference specification:

$$y_{itj} = \beta_0 + \beta_1 treatment_{itj} + \beta_2 X_{itj} + \eta_j + \theta_t + \epsilon_{itc}$$

where y_{itj} represents the outcome (probability of admission, number of admissions, or days spent in hospital) for child *i,* born at time *t* in community *j*. η_j denotes fixed effects for the mother's community of residence when the child was born, and θ_t represents fixed effects for the year-quarter that the baby was born. The vector of covariates X_{itj} includes sex of the baby, the mother's age at birth (in bands), an indicator for whether it was the mother's first pregnancy, the season that the baby was conceived in, and rainfall in the mother's place of residence in the 90 days to birth.[58] We also estimate an event study based on this model,[59] which allows us to separately estimate the impact of exposure depending on when it occurred relative to birth (i.e. in each trimester of pregnancy, or after birth).

β_1 represents our treatment coefficient. As described in Section 2.4, the policy change represented a severe income reduction to some households, and a milder reduction and

[57] Breastfeeding rates are high in NT communities (Longmore et al. 2020), and therefore we expect maternal nutrition and stress to be the predominant mechanism of any effect, for exposure both during pregnancy and as a newborn. However, the income reduction may also affect parents' ability to purchase infant formula if they are not breastfeeding.

[58] We control for season and rainfall because in the NT, babies born in the summer/wet season are generally less healthy, with potential long-term impacts on health (M.-A. Doyle 2023).

[59] The event study specification is given by: $y_{itj} = 0 + \sum_{n=1}^{4} 1(treatment_{itj}=n) + 2X_{itj} + j + t + itc$. In this specification, we have four categories of treatment: rollout occurring in the first trimester, second trimester, or third trimester or pregnancy, or in the baby's first 3 months of life (the so-called 'fourth trimester').

source of stress to others. Therefore, we expect there will be heterogeneous treatment effects.

Recent econometric literature has highlighted the issues with using standard two-way fixed effects models in cases with heterogeneous treatment effects across time and groups. There are now a range of estimators available that address these issues (Callaway and Sant'Anna 2020; Imai and Kim 2020; Borusyak, Jaravel, and Spiess 2024; Gardner 2021; Wooldridge 2021), each using different assumptions. Our data structure is slightly unusual for a generalised difference-in-difference application. This is because, while we have balanced longitudinal data at the individual level (which allows us to track outcomes through to age 8), for the purposes of this estimation, our data are made up of repeated cross-sections of children born in each community in each time period. Each time period-location cell is made up of a varying number of observations depending on the number of births in that period, and some small communities have no births in some quarters. Furthermore, each child has individual-level characteristics (our X_{itj} variables) which we include as covariates. This data structure discounts the possibility of using Callaway and Sant'Anna's (2020) estimator, which is one of the more commonly used estimators. With that estimator it is possible to include covariates, but not time-varying covariates.[60]

We instead use Borusyak, Jaravel and Spiess' (2024) imputation estimator, which is more flexible and hence accommodates our data structure, allowing for individual-level covariates and varying numbers of observations (some zero) in each time period-community cell. Furthermore, as both Borusyak Jaravel and Spiess (2024) and de Chaisemartin and D'Haultfoeuille show (2023), this estimator is more efficient than alternatives. The increased efficiency and ability to control for individual-level covariates reduces our standard errors: this is desirable given our small population with highly

[60] To be clear, the x_{itj} covariates we use are constant within the individual over time, but the dimension across which time variation is relevant for choice of estimator is time variation within a given treatment unit (i.e. community) and as the individuals born within each period are different, the values of our x_{itj} variables will change from quarter to quarter within a given community.

variable outcomes. We provide more detail on this in Appendix C, and present robustness tests using alternative estimators in Appendix Table E7.

5.6 Results

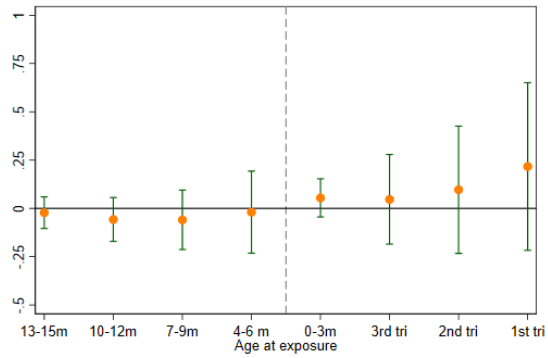
We present our primary estimates in two ways.

We first present event study estimates for our high-level outcomes of hospital admissions for any reason, number of days spent in hospital, and hospital admissions for infection (Figure 3). These figures break down the impact of the policy change 'event' based on the child's age at the time of the policy introduction. This allows us to do two things. First, to confirm the validity of the natural experiment. These figures show no statistically significant effects or trends in our comparison group. Second, with the event studies, we can see whether the effects of the policy introduction differ by timing of exposure. We can see, for instance, whether the magnitude of the policy's effect differed for children for whom the policy was introduced in their community in their first, second, or third trimester in utero, or as a newborn.

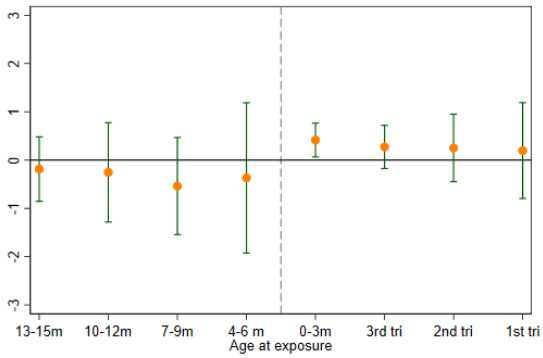
We then present a table with overall difference-in-difference estimate, which shows the average treatment effects across all treated children, regardless of the timing of the policy introduction. This allows us to summarise our findings as point estimates, and to present a larger number of outcomes: in our case, diagnoses for specific types of infections.

5.6.1 Event study estimates

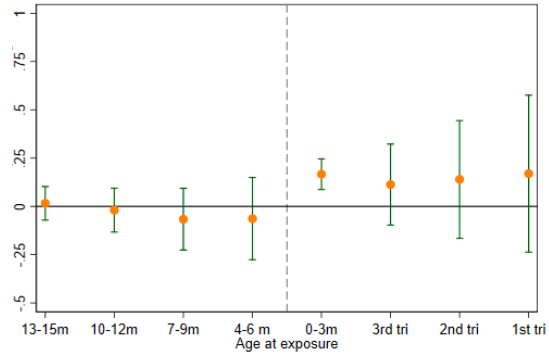
Figure 3 presents the impact of exposure to the introduction of income management on hospital admissions and days spent in hospital, cumulatively for ages 0–7 years. All estimates are relative to exposure at 16–21 months, following recommendations by Borusyak, Jaravel and Spiess (2024). that the reference period should not be the one immediately preceding the treatment.



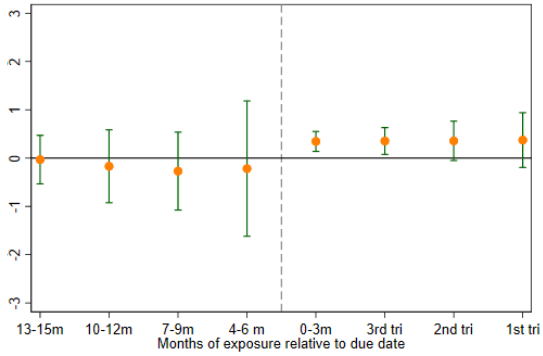
A. Probability ever admitted to hospital for any reason



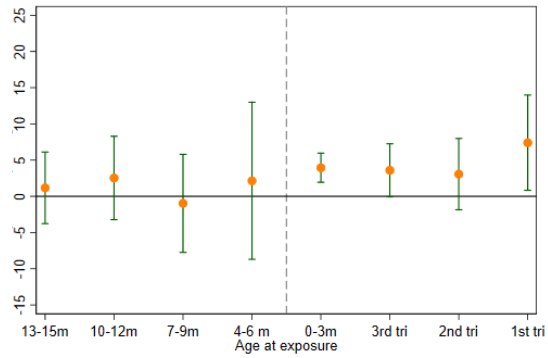
B. Number of times admitted to hospital for any reason



C. Probability ever admitted to hospital for infection



D. Number of times admitted to hospital for infection



E. Number of days in hospital

Figure 3: Event study for impact of introduction of income management on hospital admissions at ages 0 to 5 years.

Note: These figures show the estimated effect using the Borusyak, Jaravel and Spiess (2024) imputation estimator. The reference period is children aged 16-21 months old at the time of the policy.

Figure 3(a) shows the impact of the policy change on the probability of ever being admitted to hospital for any reason. With 95% confidence intervals spanning zero for all periods of exposure, there is no statistically significant difference in hospital admission for children exposed to income management during pregnancy or as a newborn, relative to exposure at 16-21 months.

Figures 3(d) and 3(e) show the impact of the policy change on the number of times a child is admitted to hospital for an infection, and days spent in hospital, respectively. If the policy change happened when the child was 0–3 months old, or in the third trimester of pregnancy, there is a statistically significant increase in admissions for infection, equivalent to around 0.4 additional admissions per child. There is a similarly sized estimate for children exposed in the first or second trimesters of pregnancy, though the 95% confidence intervals are wider, spanning zero. We also find an increase in days spent in hospital, of around 4 or 5 days per child, for those exposed to the policy as a newborn, in their third trimester, or their first trimester. While the magnitude of the estimate on exposure in the second trimester is similar, the 95% confidence intervals span zero. We do not find impacts on any outcome for children exposed when they were 4 months and over, relative to exposure at 16-21 months.

In sum, the plots in Figure 3 suggest that the introduction of income management increased the number of infection-related hospital admissions, and the number of days spent in hospital. Across the whole treatment group, we find no statistically significant impact on the probability of children ever being admitted to hospital for any reason or ever admitted for an infection – although we do find small effects on total number of admissions for children treated in the first three months of life (Figure 3(b)). As the share of children ever admitted was already high before the policy change, it makes sense that most of the effect of the income reduction would come through the intensive margin (i.e. number of admissions, days spent in hospital) and not the extensive margin (i.e. whether ever admitted). Across almost all outcomes, the point estimates remain more or less constant for different periods of exposure, however, the 95% confidence intervals around

our point estimates become wider for children exposed to the policy earlier, reflecting greater uncertainty around these estimates.[61]

5.6.2 Difference in difference estimates

Table 1 presents point estimates for a treatment effect aggregated over all periods (i.e. exposure during pregnancy or as a newborn). Panel A presents the same outcomes that are shown in Figure 3. The estimates suggest that the introduction of income management increased the number of days spent in hospital by 4.7, from a base of 11.5 days, and number of admissions for infection by 0.37 admissions per child, from a base of 1.6 admissions. Both effects are statistically significant at the 5% level.

We do not find a significant effect on overall admissions, but this does not mean that other admissions reduced at the same time that admissions for infections increased. Instead, the point estimate for the change in total admissions for any reason (0.36 admissions) is very close to that for admissions for an infection (0.37 admissions), but not statistically significant. This likely reflects greater baseline variation in total admissions, and hence a larger standard error.

Panel B of Table 1 breaks down the treatment effect on admissions by type of infection. We find a statistically significant increase in the probability of ever being admitted to hospital for three types of infection: viral (8 percentage points, significant at the 1% level), bacterial (13 percentage points, significant at the 1% level) and upper respiratory (7 percentage points, significant at the 5% level). Income management doubles the share of children ever admitted to hospital for these three reasons, with very similar point estimates on the number of admissions for these reasons.

[61] According to de Chaisemartin and D'Haultfoeuille (2023), the estimator we use is known to produce somewhat less precise estimates following the first treatment period. The confidence bands around the pre-treatment periods are also wider, reflecting the fact that the imputation method uses a different estimator for pre- and post-treatment estimates (see Borusyak, Jaravel and Spiess (2024) for more details on the reasons for this).

	Whether ever admitted ages 0-7 years (probability)			Number of admissions ages 0-7 years (cumulative)		
	Untreated mean	Treatment effect	95% confidence interval	Untreated mean	Treatment effect	95% confidence interval
Panel A: main outcomes						
Admission for any reason	0.75	0.09	-0.15 - 0.33	2.52	0.36	-0.16 - 0.83
Admission for infection	0.61	0.15	-0.07 - 0.37	1.62	0.37**	0.08 - 0.66
Days in hospital				11.48	4.68**	0.97 - 8.40
Panel B: Type of infection						
Viral	0.05	0.08***	0.03 - 0.14	0.06	0.08***	0.02 - 0.14
Bacterial	0.12	0.13***	0.03 - 0.23	0.15	0.17**	0.04 - 0.30
Upper respiratory	0.11	0.07**	0.01 - 0.13	0.15	0.09**	0.02 - 0.16
Lower respiratory	0.31	0.09	-0.14 - 0.31	0.65	0.12	-0.06 - 0.29
Soft tissue	0.27	0.01	-0.19 - 0.22	0.39	-0.08	-0.39 - 0.23
Gastrointestinal	0.23	-0.07	-0.25 - 0.12	0.34	-0.01	-0.15 - 0.12
Sample size		2,279			2,279	

Table 1: Impact of income management introduction on hospital admissions

Note: Treatment is defined as exposure to income management within the mother's community before the child is 90 days old. Excludes admissions within 10 days of birth admission. Covariates: child's sex, whether they were the mother's first pregnancy, rainfall in the 90 days to birth, whether conceived in the wet season. Fixed effects for community at birth and quarter-year of birth. Standard errors clustered by community. Treatment effects and standard errors estimated using Boruysak, Jaravel and Spiess' (2024) estimator (Stata command: did_imputation (Borusyak 2023)).
*** 1%; ** 5%, * 10%

5.6.2 Sensitivity of our estimates

Our estimates presented here are robust to a host of alternative specifications and estimators, presented and described further in Appendix E. However, Table 2 presents the two estimation decisions that have most impact on our estimates and hence affect how we interpret our findings.

First, as discussed in Section 2 above, while we are unable to link data on actual receipt (or delays in receipt) of social security benefits with our outcomes, we may assume that implementation challenges were most acute at the beginning of the rollout period. We can see from Figure 2 that the share of income managed funds that were unallocated was highest in the final months of 2007, and dramatically lower in 2008. If the effects we find were driven by these implementation challenges, then we may expect the impact to be much smaller for communities that received income management later in the rollout period. To test this, we rerun our analysis excluding the communities receiving income management in 2007. Our estimate on days spent in hospital declines from 4.7 days to 1.6 days and loses statistical significance. The estimated effect on number of admissions remains little changed but also loses statistical significance.

Second, our benchmark specification is based on time and community fixed effects. We could instead use sibling fixed effects. Here we control for all unobserved family-level characteristics that might influence a child's health outcomes, such as genetic background, parental education, and family environment. When we do this, our estimates reduce in magnitude. This approach suggests income management led to an additional 1.6 days in hospital and 0.31 additional admissions for infection. The estimates also become much more precise, and statistically significant at the 1% level. However, there are two reasons we prefer our benchmark model. The first is sample size: only around 15 percent of children born in the treatment period have a sibling born within the sample period, meaning use of sibling fixed effects discards information from a large number of children without siblings. Second, some families move between communities, meaning siblings may be born and spend their early years in different communities. This is the case for around 10% treated children with a sibling in our cohort. Hence, while this approach gives us more precise estimates, it does not strictly fit within our identification strategy, which relies on holding community (i.e. unit of treatment) characteristics constant.

Put together, these sensitivity tests imply two things: first, given that our estimates lose statistical significance when we exclude the earliest-adopting communities, this supports

our interpretation that much of the total effect we find is due to initial implementation challenges. However, the fact that the point estimates remain positive among later-adopting communities suggests there may be some more modest impact from ongoing implementation challenges and reduced purchasing power as a result of the income management policy. Second, the range in estimates we find between our benchmark model and the sibling fixed effects model suggests it is likely that the impact of income management was somewhat less than that implied by our benchmark model.

	N	Pr. Ever admitted	N admissions	Days in hospital	Pr ever admitted for infection	N adm. for infection
Benchmark specification	2,279	0.090	0.335	4.683**	0.147	0.370**
		(0.123)	(0.253)	(1.895)	(0.113)	(0.150)
Exclude first-adopting communities	1,856	0.095	0.348	1.587	0.162	0.399*
		(0.137)	(0.126)	(1.106)	(0.284)	(0.213)
Sibling fixed effects	2,465	-0.023***	0.245***	1.579***	0.0511***	0.305***
		(0.007)	(0.061)	(0.349)	(0.006)	(0.038)

Table 2: Impact of income management introduction on hospital admissions – selected robustness tests

Note: The benchmark includes fixed effects for community at birth and quarter-year of birth, the 'siblings fixed effects' specification does not include any additional control for location. The 'exclude first-adopting communities' specification excludes all communities receiving income management in 2007. Standard errors clustered by community. Treatment effects and standard errors estimated using Boruysak, Jaravel and Spiess' (2024) estimator (Stata command: did_imputation (Borusyak 2023)).
*** 1%; ** 5%, * 10%

5.6.3 Does the increase in admissions reflect worse health?

It is possible that the increase in admissions reflects an increase in a family's likelihood of taking their child to hospital for a given level of illness, instead of worse underlying health. Care-seeking behaviour may well have changed in response to the income management rollout. It is plausible that income management, and other policy changes occurring around that time, may have led families to withdraw from accessing routine care due to mistrust of public institutions.

For such behaviour to drive our estimates, it would require that care-seeking responses to income management differed based on the age the child was at the time of the policy introduction. That is, for instance, that a family whose child was 1 at the time of the policy change would be less willing to access healthcare years later, compared with a family whose child was a newborn at the time.

If this were the case, there could be two possible explanations for the increase in admissions we find. First, the treatment group may go to hospital as a substitute for primary care (that is, for less severe infections), which would mean an increase in hospital admissions, a reduction in primary care use, but no change in underlying health. Second the treatment group may avoid seeking routine healthcare. This would mean a reduction in primary care, but could lead to more infections going untreated and becoming severe enough to require hospitalisation.

We can test these possibilities using primary care clinic data. If either of these scenarios were true, we would expect a reduction in primary care visits. If instead the increase in hospital admissions truly reflects worse health, we may expect to see an increase in primary care consultations alongside the increase in hospital admissions. Table 3 shows that this was indeed the case. Focusing on infections, we find that the treatment group had on average one more primary care consultation for an infection than the comparison group between ages 3-7; this increase is concentrated in respiratory infections, for which the treatment group had 1.3 more consultations than the comparison group, statistically significant at the 5% level.

We interpret this as evidence that the increase in hospital admissions represents a worsening of children's health, instead of a change in health care seeking behaviour.

	Any consultation for infection		Number of consultations for infection	
	Untreated mean	Treatment effect	Untreated mean	Treatment effect
Any infection	0.86	0.0101	8.14	1.079
		(0.0393)		(0.686)
Eye or skin infection	0.76	0.0119	3.62	-0.0833
		(0.0197)		(0.148)
Respiratory infection	0.79	0.233***	3.87	1.303**
		(0.0437)		(0.633)
Sample size		1,191		1,191

Table 3: Consultations for infection at primary care clinics, ages 3-7 years
Note: clinic data available for 2/3rds of communities, and from 2010 onwards only. Analysis sample is limited to children born in communities that had a clinic contributing data for the full period 2010-2015. Data are to-coded at 2.5% at the upper end to eliminate extreme outliers.

5.6.4 Could the increase in admissions have been predicted earlier?

Our longitudinal data present an opportunity to unpack the relationship between the policy change, birth outcomes, and early childhood health. We ask the question: could the increase in hospital admissions have been predicted by the policy's shorter-term effect on birthweight and other indicators of newborn health?

To answer this question, we re-estimate our models, sequentially adding controls for birthweight, as well as a set of other indicators of perinatal health: maternal anaemia during pregnancy, fetal distress during delivery and admission to a special care nursery after birth. If the estimated treatment effect declines with these additions, this is an indication that the treatment effect could have been, at least partly, predicted using these variables observed much earlier in life.

We find that controlling for birthweight in our model leads to only a small reduction in the estimated treatment effect (Table 4, columns 1 to 2) – by around 7 percent.

We then consider our additional measures of perinatal health. All measures reduce the treatment effect somewhat. The largest reduction comes from including an indicator for whether the baby was admitted to a special care nursery, reducing the treatment effect by

21 percent (columns 1 and 5). Including all variables in the model (column 6), the treatment effect on number of admissions for infection is reduced by around one-third.

Overall, this analysis suggests that the income reduction impacted in utero and newborn health in ways that are not fully reflected in birthweight. Crucially, this means that a prediction of the policy's impact on children's health based on the impact on birthweight alone would capture only about 7% of the effect we find, severely under-estimating the true effect.

However, based on just a handful of additional variables measured at or before birth, we can predict around one-third of the policy's impact on later health. This is promising because it means that it may be possible to identify and target support at the subset of newborns – out of a larger population affected by a particular shock – who are most at risk of poor health in childhood. Further research, perhaps using machine learning methods in a larger population, could identify more systematically which indicators of health are the strongest and most robust predictors of childhood infection.

	(1)	(2)	(3)	(4)	(5)	(6)
Treatment effect	0.37**	0.34*	0.36**	0.33**	0.29*	0.26
	(0.15)	(0.19)	(0.16)	(0.15)	(0.15)	(0.18)
Controls						
Birthweight		✓				✓
Anaemia during pregnancy			✓			✓
Fetal distress during delivery				✓		✓
Admitted to special care nursery					✓	✓

Table 4: Treatment effect on number of admissions for infection, controlling for measures of perinatal health

Note: The treatment effects reported here relate to the benchmark model shown in Table 1, on the number of admissions for infection. We include controls in the model for various measures of perinatal health. We present treatment effects from 0 months to the 8th birthday, as in Table 1, when using controls relating to the pregnancy and birth.

5.7 Discussion: immune programming

We have found a broad-based increase in infections as a result of an income reduction in utero or as a newborn. Viral, bacterial and upper respiratory infections all increased through to age 7. These effects are concentrated at early ages (up to age 3; Appendix Figure F1). At these younger ages, we see additional effects on soft tissue (primarily skin) infections, and in lower respiratory infections.

What mechanisms could explain the link between the treatment (a decrease in income and increase in maternal stress in the perinatal period) and hospitalisation for infection years later? It is unlikely that the treated group have higher exposure to pathogens causing infection – they live in the same communities as children in the comparison group. Instead, it must be that the treated group are either more susceptible to contracting infections, or are at higher risk of severe symptoms requiring medical treatment if they do contract an infection. Either case would suggest that the treatment group have a weaker immune response relative to the comparison group.

Our findings fit with the 'immune programming hypothesis' described by Marques et al. (2013). This is, that both maternal stress and poor nutrition in utero and during breastfeeding can program a permanent change in the way a child's immune system develops. The precise physiological mechanisms are still not well understood. But evidence from human and animal studies find that perinatal nutrition and stress are associated with long-term functioning of both innate (Spencer 2013) and adaptive immunity (Marques et al. 2013). Adkins, Leclerc and Marshall-Clarke (2004) describe the perinatal period as a 'unique developmental window' for immunity, given the plasticity of the immune system at this time.

Much of the evidence on immune programming relates to nutrition, both in utero (Wellinghausen 2001) and as a newborn (Spencer 2013). These are the periods when

epigenetic imprinting is most active (Simeoni et al. 2014), and therefore when the child's environment can affect not just their health contemporaneously, but can lead to long-term changes in their developmental trajectories. But in addition to nutrition, there is substantial evidence that maternal stress during pregnancy can have epigenetic effects (Nielsen et al. 2011), and as Marques et al. (2013) describe, this, too, affects the child's developing immune system.

Turning specifically to nutrition, it is well documented empirically that malnourished children, and even children with micronutrient deficiency, are more susceptible to infections contemporaneously (Walson and Berkley 2018; Govers et al. 2022). But there is some empirical evidence of permanent effects on immunity from exposures during the prenatal and neonatal period. In a range of human and animal studies, maternal micronutrient intake during pregnancy and breastfeeding (e.g. folate, vitamin D and omega-3) has been linked to immune function in childhood (Bermick and Schaller 2022). For instance, using data from The Gambia and Bangladesh, Moore et al. (2006) find that poor nutrition in pregnancy (caused by being born in the 'hungry season') is associated with smaller thymus size one year after birth, even after controlling for nutritional status at the time of measurement.[62] They suggest this may be an early indicator of functional deficits of the immune system, potentially lasting into adulthood. As Bourke et al. (2016) note, this body of research suggests there may therefore be a double burden of poor prenatal nutrition, both increasing the risk of metabolic disease in adulthood (Barker's (1998) 'thrifty phenotype' hypothesis), and permanently weakening the immune system.

While not specifically testing the immune programming hypothesis, population-level empirical research has found associations between perinatal health and childhood infection rates (Miller et al. 2016; Nielsen et al. 2011). This research generally uses birth outcomes – birthweight and gestational age – as indicators of perinatal health. However, these estimates provide a noisy signal on the impact of immune programming, because maternal stress and nutrition need not affect birth outcomes to impact children's

[62] T-cells, which fight infections, develop in the thymus.

development: this is clearest for nutrition, as we know that even when mothers consume sufficient calories and macronutrients, maternal micronutrient deficiency is associated with higher childhood risk of infection, asthma and allergic reactions (Marques et al. 2013, 6; Wellinghausen 2001). Therefore, it is not necessarily only preterm and low birthweight children who are at risk of 'immune programming'; babies born full term and within the healthy birthweight range can be affected. This reasoning could explain why much of the effect we find cannot be predicted by birthweight.

We have identified effects of the policy change on health in early childhood. In principle, the effects of pre- and neo-natal exposures to poor nutrition and stress could go on to affect adult human capital in two ways. First, as described above, there may be permanent impacts on how effectively the immune system functions, directly impacting health in adulthood. Perinatal immune programming is linked not just to weaker immunity to infections as measured in our analysis, but can also lead to asthma, allergies, and autoimmune disorders – there may also be a link with neurodevelopmental disorders such as autism spectrum disorder and schizophrenia, with an effect mediated through the immune system (Marques et al. 2013). Second, even in the absence of permanent effects on the immune system, there may be a feedback loop between health and cognitive outcomes in childhood, whereby increased frequency and severity of infections may reduce children's ability to develop their cognitive skills (MAL-ED Network Investigators et al. 2018; Alderman et al. 2001), a process that Cunha and Heckman (2007) refer to as 'self-productivity'.

To explore possible impacts on other elements of immune system functioning further, we test for treatment effects on two additional diagnoses in our data: asthma and allergies. We use the primary care clinic data, as these conditions would not necessarily require hospitalisation. We find that the treated group were almost twice as likely to be diagnosed with asthma, with a 3.3 percentage point increase in the share of children diagnosed, off a baseline of 4% of children between ages 3 to 7 (Appendix Table F1). We do not find

any increase in primary care visits for allergies. We take the increase in asthma diagnoses as supporting evidence of the immune programming hypothesis.

Immune system development has not been discussed as part of the 'early origins' literature in the social sciences. However, piecing together evidence from the medical literature and from animal studies (Marques et al. 2013; McDade and Kuzawa 2004), epidemiological evidence on the associations between the perinatal environment and infection risk (Miller et al. 2016; S. E. Moore et al. 2006; Nielsen et al. 2011), and our findings, we believe this is a likely mechanism. We also note that our findings on infection and asthma risk are consistent with a range of other findings in the field of economics. For instance, De Gendre et al. (2021) find that an unconditional cash transfer at birth, the opposite of the treatment considered in our study, significantly reduces the risk of bronchiolitis, which is the most common type of infection in infancy, and a precursor to asthma. Gunnsteinsson et al. (2022) find that a vitamin A supplement at birth, which promotes normal immune system development, supports growth and reduces fevers (associated with infection) in infancy.

We therefore propose that immune programming explains our findings and may be the mechanism behind many other findings within the early origins literature.

5.8 Conclusions

In this chapter, we estimate the impact on health in early childhood of a temporary reduction in social security payments for low-income households, occurring in utero or as a newborn. We do this by analysing the introduction of 'income management', a change in the way that government transfers were paid to residents in remote Aboriginal communities in the Northern Territory of Australia. Due to implementation challenges, this policy change meant that many households received only half of their full payment for up to 13 weeks, and a small number had their payments temporarily suspended after

13 weeks. For all households, even if they did not experience a suspension of their payments, the policy introduction led to stress and reduced purchasing power.

We estimate the impact of exposure to the policy change during pregnancy or in the newborn period (0-3 months), compared with exposure after 3 months of age. We find a 4½ day increase in the average number of days spent in hospital from ages 0 to 7, a 40% increase on baseline. This was driven by an increase in hospital visits for infection. We also find an increase in primary care visits for respiratory infections, which strengthens our interpretation that the increase in hospital visits is the result of worse health, and not the result of a change in care seeking behaviour.

Our findings are consistent with the literature on the early origins of human capital, which shows that shocks to the perinatal environment can have long-lasting effects. In this case, we hypothesise that these effects are due to poorer nutrition and maternal stress, as a result of the reduction in income during a critical window in fetal and neonatal development. While household spending and consumption data are not available for us to empirically test this channel, it is supported by aggregate evidence (see Section 2.3). This temporary reduction in income appears to have impacted children's immune system development, and hence increased their risk of severe infection requiring hospitalisation. Putting this together with existing research on immunity (S. E. Moore et al. 2006; McDade et al. 2001), the immune programming hypothesis (Marques et al. 2013), and our additional finding of increases in asthma diagnoses, it appears likely that these impacts on immune functioning will be permanent for this cohort of children.

Our estimates are large, and contrast with studies such as Borra et al. (2024) and Hawkins et al. (2023) for Spain and California, respectively. Both find no impact on childhood health and healthcare use from the introduction or removal of generous cash transfers at or around birth. In contrast, we find an increase of 4½ additional days spent in hospital – though some sensitivity tests suggest the magnitude could be substantially lower, closer to 1.6 days. The fact that we find effects where other studied did not is likely due to differences in the populations studied. Residents in Aboriginal communities in the NT

face high poverty rates. An income reduction in this context is likely to affect maternal nutrition and stress more acutely than a shock of a similar magnitude in other high-income countries where cash transfers have been studied. Furthermore, the impact of weaker immunity on developing severe infection symptoms is likely lower for children who do not face a particularly high risk of infection. Children in Aboriginal communities were already at high risk of infection (Carville et al. 2007). Elevated exposure to pathogens through crowded, poor quality housing and infrastructure (Foster and Hall 2021), and limited availability of nutritious food within remote communities (Leonard et al. 2017; Tonkin et al. 2020) likely impacts childhood immunity contemporaneously as well. While certain shocks during sensitive periods may affect children throughout the income distribution, the large estimates that we find underscore the importance of paying particular attention to impacts among disadvantaged groups.

Aboriginal communities in Australia are unique, but there are unfortunately many families living in deep poverty in advanced economies that offer means-tested benefits. Children in these families also face poor nutrition, elevated risk of infection and imperfect access to healthcare (US Census Bureau 2024; Torjesen 2023). The effect sizes we find are most likely to generalise to these very low-income groups, and not to a general population (as studied by Borra et al.) or even to all households eligible for means-tested benefits (as studied by Hawkins et al.).

Our findings make three key contributions to the evidence base on the early origins of human capital.

First, we highlight the importance of policy implementation and administration on health in early childhood. We find that even a temporary reduction in social security payments can have sustained impacts on children, if experienced in utero or as a newborn. Our findings lend support to the call for an ‘early childhood lens’ to policymaking across a broad range of policy domains (National Scientific Council on the Developing Child 2023).

Second, we interpret our findings in light of the 'immune programming hypothesis'. We are the first to estimate the causal effect of a perinatal shock on immune functioning in childhood. We propose the immune system as a mechanism through which shocks or investments in utero and very early life can have long-term effects.

Third, we unpack the link between indicators of perinatal health and childhood health. The effect we find on child health is orders of magnitude larger than what could have been predicted given the initial impact on birthweight. However, we find that use of a handful of additional measures of maternal and perinatal health help to predict one-third of the treatment effect. This suggests that use of a broader set of measures of perinatal health in research could greatly improve our ability to identify children early who are most at risk of adverse outcomes, and hence who may benefit most from targeted support. This element of our research was exploratory: further work could select and evaluate specific measures.

As well as broadening our understanding of the nature of early life shocks in general, our research documents the specific costs of this episode in Australian policy. In effect our estimates reflect the combined impacts of the policy implementation, and the policy itself. We cannot definitively disentangle the two. However, the fact that our estimates reduce in magnitude and lose statistical significance for communities which received income management later in the rollout period suggests the large impacts we estimate are driven by implementation challenges. The income management policy itself did not improve child health, but did not cause the large adverse effects we estimate. We therefore conclude that the poor implementation of income management had substantial costs for the children affected, their families, and the healthcare system. An additional 4½ days in hospital per child in this cohort would have costed an estimated $4.5m to the NT public hospital system over eight years.[63] Longer-term effects may be larger, with the

[63] The Australian Independent Hospital Pricing Authority (2014) estimates that the average cost of a child spending one night in hospital in remote NT was AU$1,973 in 2013/14. Given our estimate of 4.6 additional days in hospital, this implies an additional AU$9,076 in hospital costs per child. With just under 500

international literature suggesting there may be intergenerational effects of perinatal shocks (Veenendaal et al. 2013; Costa 2023).

children in the treatment group, this would suggest that the policy introduction cost the public hospital system around AU$4.5 million in the six years following its rollout.

Appendix A: Descriptive statistics

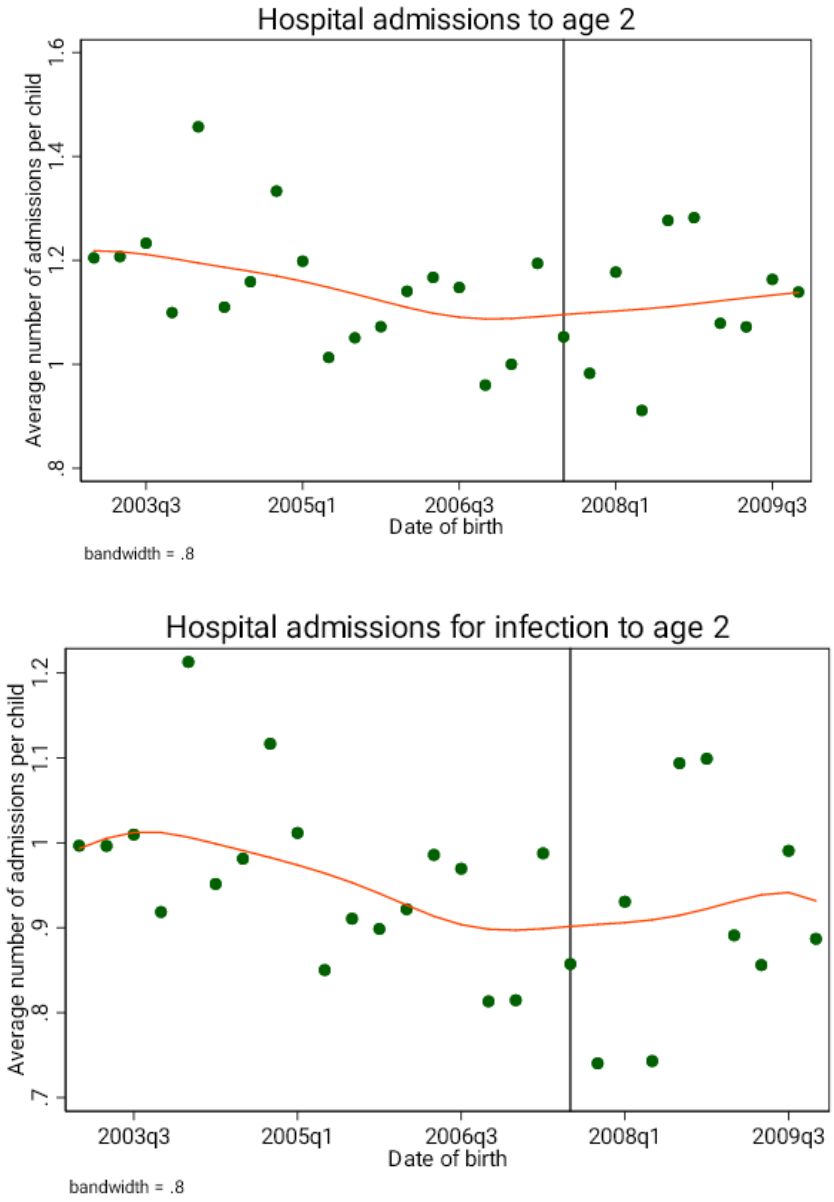


Figure A1: Trend decrease in number of hospital admissions per child before the policy change

Note: These figures show trends in the number of hospital admissions per child over time, generated based on locally weighted regressions using the 'Lowess' command in Stata. The horizontal axis represents the child's date of birth. The vertical line shows the date that income management was introduced in the first community., We show outcomes through to children's 2nd birthdays instead of through to their 6th birthdays (as in our main analysis) to better facilitate comparisons before and after the policy change. Children whose outcomes were not at all affected by income management are those born two years before the introduction (i.e. before Quarter 3, 2006).

	Mean	SD	Min	Max
Health at birth				
Average birthweight	3079.52	641.86	300.00	5600.00
Low birthweight (pr)	0.14	0.35	0.00	1.00
Preterm (pr)	0.15	0.36	0.00	1.00
Small for gestational age (pr)	0.20	0.40	0.00	1.00
Admitted to special care nursery (pr)	0.25	0.43	0.00	1.00
Anaemia during pregnancy (pr)	0.06	0.24	0.00	1.00
Fetal distress during delivery (pr)	0.12	0.33	0.00	1.00
Mother, baby and birth characteristics				
Mother's age	23.48	6.04	13.00	45.00
Mother's first pregnancy (pr)	0.25	0.43	0.00	1.00
Boy (pr)	0.52	0.50	0.00	1.00
Rainfall in 90 days to birth (ml)	3.36	3.78	0.00	17.72
Conceived in wet season/summer	0.52	0.50	0.00	1.00
Hospital inpatient admissions				
Ever admitted ages 0-7 (pr)	0.78	0.41	0.00	1.00
Number admissions	2.55	3.18	0.00	93.00
Total days in hospital ages 0-7	13.15	25.27	0.00	508.00
Number of diagnoses per admission	3.21	2.30	1.00	33.00
Admissions by age				
Avg admissions age 0	0.82	1.27	0.00	16.00
Avg admissions age 1	0.60	1.08	0.00	17.00
Avg admissions age 2	0.27	0.65	0.00	12.00
Avg admissions age 3	0.18	0.53	0.00	12.00
Avg admissions age 4	0.18	0.54	0.00	13.00
Avg admissions age 5	0.18	0.51	0.00	11.00
Avg admissions age 6	0.17	0.49	0.00	13.00
Avg admissions age 7	0.15	0.48	0.00	13.00
Ever admitted age 0-7 for:				
Any infection (pr)	0.65	0.48	0.00	1.00
Viral infection (pr)	0.04	0.20	0.00	1.00
Bacterial infection (pr)	0.10	0.30	0.00	1.00
Upper respiratory infection (pr)	0.11	0.31	0.00	1.00
Lower respiratory infection (pr)	0.34	0.48	0.00	1.00
Soft tissue infection (pr)	0.24	0.43	0.00	1.00
Gastrointestinal infection (pr)	0.31	0.46	0.00	1.00
Genito-urinary infection (pr)	0.03	0.17	0.00	1.00
Number admissions for infection age 0-7	1.70	2.32	0.00	30.00
Population - children in NT communities				
Share of NT population aged 0-7	26			
Share of NT hospital admissions aged 0-7	50			

Table A1: Summary statistics for children born in NT communities in the years before income management rollout (born Jul 2006-Jun 2007)

Appendix B: Causal identification and parallel trends

	Outcome: Days from first day of rollout to day community received income management				Outcome: Order in which community received income management (i.e. first, second, etc)			
Admissions ages 0-5 per child (avg. of children born 2004-2006)	-2.946	-3.726	5.702	-1.798	-0.197	-0.322	-0.156	-0.189
	(14.57)	(15.35)	(37.40)	(15.25)	(1.320)	(1.413)	(3.400)	(1.407)
Days in hospital ages 0-5 per child (avg of children born 2004-2006)	-0.713	-0.722	-5.960	-0.787	-0.0578	-0.0593	-0.439	-0.0638
	(2.923)	(2.963)	(5.629)	(2.861)	(0.241)	(0.246)	(0.504)	(0.239)
Rainfall		-1.917	-0.0312	-0.856		-0.307	-0.351	-0.234
		(5.864)	(12.99)	(6.047)		(0.523)	(1.126)	(0.540)
Total population			-0.0362				-0.00401	
			(0.0403)				(0.00388)	
Median age			5.837				0.668	
			(6.810)				(0.611)	
Share of population aged 65+			-4.147				-0.490	
			(10.48)				(0.937)	
Avg people per household			32.48***				3.449***	
			(13.34)				(1.301)	
Median personal income			0.516				0.0581*	
			(0.380)				(0.0337)	
Labour force to population ratio			-0.502				-0.0390	
			(1.016)				(0.0975)	
Missing Census data				26.67				1.839
				(26.49)				(2.393)
Constant	225.9***	233.5***	-112.6	216.5***	17.09***	18.30***	-19.66	17.13***
	(31.83)	(43.28)	(255.3)	(46.19)	(2.862)	(3.916)	(24.33)	(4.215)
Observations	85	85	54	85	85	85	54	85
R-squared	0.008	0.009	0.188	0.023	0.006	0.009	0.207	0.017

Table B1: community characteristics and treatment timing

Test of joint significance of pre-trend coefficients

	Whether ever admitted ages 0-5		**Number of admissions ages 0-5**	
	F-stat	P-value	F-stat	P-value
Any admission	0.49	0.74	0.47	0.76
Days in hospital			1.16	0.33
Admissions for infection				
Any infection	0.37	0.83	0.20	0.94
Viral	0.07	0.99	0.11	0.98
Bacterial	2.55	0.05	2.94	0.03
Upper respiratory	0.32	0.86	0.60	0.67
Lower respiratory	0.11	0.98	0.18	0.95
Soft tissue	1.24	0.30	0.79	0.53
Gastrointestinal	0.27	0.90	0.41	0.80

F-test for joint significance of coefficients on income management introduction 3-6 months, 6-9 months, 9-12 months, 12-15 months, 15-18 months after birth. Regression on pre-treatment sample only. Excludes admissions within 10 days of birth admission. Covariates: child's sex, whether they were the mother's first pregnancy, rainfall in the 90 days to birth, whether conceived in the wet season. Fixed effects for community at birth and quarter-year of birth.

Table B2: Parallel trends

	Treated	Untreated
Q1 2006	0	13
Q2 2006	0	233
Q3 2006	0	223
Q4 2006	0	221
Q1 2007	0	276
Q2 2007	0	240
Q3 2007	14	173
Q4 2007	66	160
Q1 2008	137	90
Q2 2008	196	16
Q3 2008	219	2
Total	**632**	**1647**

Table B3: Number of children in treatment and comparison groups, by quarter of analysis period

Appendix C: Estimation methods

In this chapter, we estimate the impact of the introduction of income management using a generalised difference-in-difference framework, based on the staggered rollout of the policy across NT communities. We expect the effects of the treatment to be different for different groups, because the policy implementation issues led to a sharp drop in income for around one-fifth of households, and for others it led to a reduction in purchasing power and increase in stress – but no change in income. See discussion in Section 2.4. Furthermore, the effects may change over time – or in our case, may vary depending on the timing of exposure to the policy change (i.e. whether it occurred at the beginning or the end of pregnancy, or as a newborn).

This policy was rolled out gradually across communities. Until recently, most researchers would have used a linear model with fixed effects for time and location ('two-way fixed effects') to estimate the effect of policy change such as this. However, recent econometrics research has highlighted the issues with using standard two-way fixed effects models in cases with heterogeneous treatment effects, as the way that such estimators aggregate across time and groups can lead to unexpected weighting of certain time-group cells (de Chaisemartin and D'Haultfœuille 2023). The results can be misleading because of negative weights that the estimator places on some individual treatment effects, leading to bias or potentially to estimated effects in the opposite direction of the true effect. There are now a range of estimators available, with different approaches and assumptions to correct these issues (Callaway and Sant'Anna 2020; Imai and Kim 2020; Borusyak, Jaravel, and Spiess 2024; Gardner 2021; Wooldridge 2021).

Our data structure is slightly unusual for a generalised difference-in-difference application, because our data are made up of repeated cross-sections – while we have longitudinal data for each individual, they represent repeated cross sections for each unit of treatment (i.e. community). Each time period-location cell is made up of a varying number of observations (and some are empty), depending on how many children were born in a particular community in that quarter. And each child has their own

characteristics for which we want to control with covariates – including sex, maternal age, and weather conditions during pregnancy. We are also able to observe treatment timing precisely, given that we have exact date of birth, and exact date of the policy introduction in each community. One option would be to collapse the data to a community-level time series of average health outcomes for children born in each quarter. This would convert the data structure to a longitudinal (by community) format more amenable to all available estimators. However, we have a small population with high variability in outcomes. We want to maximise our ability to draw inferences from the data. We gain more statistical power by controlling for individual-level covariates, and by using the information we have on exact timing of treatment. Therefore, we keep our data at the individual level. This data structure discounts the possibility of using Callaway and Sant'Anna's (2020) estimator, which is one of the more commonly-used estimators. With this estimator it is possible to include covariates, but not time-varying covariates.

Instead, we use Bousyak, Jaravel and Spiess' (2024) imputation estimator. This estimator is suitable for our data structure, as it allows us to estimate a model with individual-level observations, identifying treatment based on exact date of birth, but with fixed effects at a more aggregated (community and quarter-year) level, and it allows time-varying covariates. It does this by estimating a model using observations from the comparison group, and using this model to predict outcomes for the treatment group. The treatment effect is the sum of the differences between the model's predictions, and actual outcomes for the treated observations. This estimator is more efficient than alternatives (Borusyak, Jaravel, and Spiess 2024; de Chaisemartin and D'Haultfœuille 2023), which reduces our standard errors: again, this is desirable given our small population with highly variable outcomes. In our robustness tests we run our model with two other estimators (see Table D7), Gardner's (2021) two stage difference in differences, and de Chaisemartin and & D'Haultfoeuille's (2020) estimator. Both allow time-varying covariates and allow us to retain individual-level data, within repeated cross-sections of treatment units. Across all alternatives, our estimates are similar in magnitude, with wide confidence intervals. As noted above, they are more precisely estimated using the imputation estimator.

Appendix D: Alternative sample period

	Whether ever admitted ages 0-7			Number of admissions ages 0-7		
	Untreated mean	Treatment effect	95% confidence interval	Untreated mean	Treatment effect	95% confidence interval
Hospital inpatient admissions	0.74	0.093	-0.173 - 0.364	2.39	0.348	-0.208 - 0.904
Hospital admissions for infection	0.59	0.162	-0.0858 - 0.409	1.33	0.399*	-0.0176 - 0.815
Days in hospital				8.91	1.587	-0.581 - 3.756
Sample size		1,856			1,856	

Table D1: Restricting sample to communities receiving IM from Jan 2008 onwards (dropping those with presumably greatest implementation issues)

Treatment is defined as exposure to income management within the mother's community before the child is 3 months old. Excludes admissions within 10 days of birth admission. Covariates: child's sex, whether they were the mother's first pregnancy, rainfall in the 90 days to birth, whether conceived in the wet season. Fixed effects for community at birth and quarter-year of birth. Standard errors clustered by community. Treatment effects and standard errors estimated using Boruysak, Jaravel and Spiess' (2024) estimator (Stata command: did_imputation).

Appendix E: Robustness tests

We conduct a series of robustness tests. First, we test robustness to model specification. We run our baseline model, including alternative variables to control for maternal and child characteristics, for seasonal variation, and alternative ways of specifying time and location fixed effects (e.g. quarter-year FE vs month-year FE). Our estimates are robust to all of these alternative specifications – they remain of similar magnitude, and statistically significant at the 5% level or higher (Tables E1, E2, and E3).

Second, given the large number of outcomes we analyse, we adjust for multiple hypothesis testing (Table E4). We adjust for the false discovery rate, using Benjamini, Krieger, and Yekutieli's (2006) sharpened Q-values, with Stata code provided by Andersen (2008). In doing so, we test two distinct sets of hypotheses: first, whether the treatment affected health in general: for this test, we include the five high-level outcomes shown in Panel A of Table 1. Second, whether the treatment affected the probability of ever being admitted to hospital for any specific type of infection, or the number of admissions for any specific type of infection (outcomes in Panel B of Table 1). We run adjustments separately for each of these hypotheses. The resulting sharpened-Q values (which can be interpreted as adjusted p-values) are presented in Appendix Table E4, alongside the p-values from the unadjusted analysis. Our main estimates on number of admissions for infection and days spent in hospital, and number of admissions for viral, bacterial and upper respiratory infections discussed above remain statistically significant at the 5% level. However, our estimates on whether ever admitted for a specific type of infection are no longer statistically significant at the 5% level.

Third, we change the length of the pre-implementation period included in our model and the time horizon over which we measure hospital admissions (Appendix Tables E5 and E6). Our main model has a pre-period of one year. That is, the 'comparison' group includes all children born in NT communities in the year before the policy rollout began, as well as children born during the rollout period, but at least three months before the policy change happened in their community. Our inclusion of the one-year pre-period

provides a larger comparison group, allowing for more precise estimates, and crucially, allows us to estimate the event study model. However, there was a trend decline in hospital admissions in the years preceding the policy change (see Appendix Figure A1). If this trend continued, including a longer pre-period would bias our estimates downwards, because the treatment group (who are, on average, born later than the comparison group) would have fewer admissions due to this trend. Roth et al. (2023) suggest that in a case with linear trends, it is best to limit the length of the pre-period included in the model. Indeed, if we include no pre-period, our estimates are larger and all are statistically significant. If instead we include a longer pre-period, this leads to little change in our point estimates, but they lose statistical significance. Because we have good reason to expect that the downward trend in hospital admissions will bias our estimates down with a longer pre-period, we conclude that our findings are robust to varying definitions of the comparison group.

Separately in terms of time horizon of our outcome measure, in Table E6, we redefine hospital admissions within scope of analysis to include those within 10 days of the baby's initial discharge from hospital for their birth admission. Doing so increases the estimate on number of days spent in hospital, but with no change in other estimates.

Finally, we run our analysis using alternative estimators, as described in Appendix C. Appendix Table E7 presents the estimates from our main model (using Borusyak, Jaravel and Spiess' (2024) imputation estimator), alongside standard two ways fixed effects linear model estimates, estimates from de Chaisemartin and D'Haultfoeuille's (2020) estimator, and Gardner's (2021) estimator. Across all alternatives, our estimates are similar in magnitude, though they are not as precisely estimated; this is consistent with Borusyak, Jaravel and Spiess' (2024) finding that their estimator provides substantial efficiency gains over alternatives.

	N	Pr. Ever admitted	N admissions	Days in hospital	Pr ever admitted - infection	N admissions - infection
Benchmark: sex of baby, whether first pregnancy, mother's age	2,319	0.110	0.241	4.602**	0.162	0.350**
		(0.112)	(0.247)	(1.883)	(0.104)	(0.147)
Sex of baby	2,319	0.105	0.251	5.084***	0.152	0.338**
		(0.119)	(0.241)	(1.956)	(0.111)	(0.165)
Sex of baby, whether first pregnancy	2,319	0.105	0.243	4.972**	0.153	0.334**
		(0.119)	(0.252)	(2.062)	(0.111)	(0.170)
Sex of baby, mother's age	2,319	0.112	0.259	4.714***	0.162	0.358**
		(0.111)	(0.227)	(1.785)	(0.103)	(0.139)
Benchmark + plural birth	2,319	0.111	0.233	4.581**	0.163	0.343**
		(0.112)	(0.247)	(1.883)	(0.104)	(0.148)
Sex of baby, mother's age, number of pregnancies carried to >20 weeks	2,319	0.1000	0.224	4.232***	0.151	0.333***
		(0.101)	(0.206)	(1.508)	(0.0947)	(0.121)
No controls	2,319	0.0999	0.238	5.041**	0.148	0.332**
		(0.120)	(0.244)	(1.977)	(0.113)	(0.165)

Table E1: Choice of maternal and child control variables

	N	Pr. Ever admitted	N admissions	Days in hospital	Pr ever admitted - infection	N admissions - infection
Benchmark: rainfall (ml) and conceived in wet season	2,279	0.0902	0.335	4.683**	0.147	0.370**
		(0.123)	(0.253)	(1.895)	(0.113)	(0.150)
Benchmark + preterm	2,279	0.0882	0.297	4.307**	0.144	0.331*
		(0.125)	(0.274)	(2.075)	(0.113)	(0.172)
Flu prevalence	2,280	0.0724	0.472	4.782**	0.126	0.426***
		(0.121)	(0.326)	(1.940)	(0.102)	(0.155)
Benchmark + flu prevalence	2,278	0.0737	0.442	4.704**	0.123	0.405**
		(0.121)	(0.334)	(1.894)	(0.105)	(0.162)
Days over 35 degrees	2,279	0.0889	0.362	4.772**	0.151	0.397***
		(0.124)	(0.238)	(1.913)	(0.110)	(0.138)
Days under 0 degrees	2,279	0.0889	0.362	4.730**	0.150	0.394***
		(0.124)	(0.242)	(2.031)	(0.109)	(0.134)
Days rainfall >50ml	2,279	0.0926	0.338	4.817**	0.145	0.363**
		(0.122)	(0.250)	(1.940)	(0.112)	(0.150)
Days humidity > 85%	2,279	0.0885	0.340	4.651**	0.148	0.377**
		(0.124)	(0.256)	(2.028)	(0.112)	(0.150)
All weather variables	2,279	0.0933	0.336	4.814**	0.146	0.369**
		(0.121)	(0.261)	(2.119)	(0.112)	(0.151)
No season controls	2,281	0.0890	0.360	4.747**	0.149	0.387***
		(0.124)	(0.240)	(1.934)	(0.110)	(0.139)

Note: all weather variables measured based on 90 days to birth; flu prevalence is average throughout pregnancy

Table E2: Choice of season controls

Time FE	Location FE	N	Pr. Ever admitted	N admissions	Days in hospital	Pr ever admitted - infection	N admissions - infection
Benchmark: quarter-year of birth	Community	2,279	0.0902	0.335	4.683**	0.147	0.370**
			(0.123)	(0.253)	(1.895)	(0.113)	(0.15)
Quarter-year of conception	Community	2,279	0.0848	0.251	5.271***	0.137	0.362**
			(0.129)	(0.252)	(2.03)	(0.118)	(0.167)
Month-year of birth	Community	2,279	0.105	0.243	5.155**	0.136	0.354**
			(0.239)	(0.264)	(2.146)	(0.12)	(0.176)
Month-year of conception	Community	2,279	0.112	0.259	4.809***	0.149	0.383***
			(0.122)	(0.232)	(1.78)	(0.11)	(0.14)
Quarter-year of birth	Region (interacted)	2,279	0.111	0.233	4.621**	0.147	0.358**
			(0.123)	(0.253)	(1.896)	(0.113)	(0.151)
Quarter-year of conception	Region (interacted)	2,279	0.1000	0.224	4.260***	0.138	0.356***
			(0.113)	(0.207)	(1.451)	(0.101)	(0.119)
Month-year of birth	Region (interacted)	2,279	0.0999	0.238	5.244**	0.132	0.358**
			(0.13)	(0.253)	(2.046)	(0.12)	(0.166)
Month-year of conception	Region (interacted)	2,279	0.0902	0.335	4.683**	0.147	0.370**
			(0.123)	(0.253)	(1.895)	(0.113)	(0.15)

Table E3: Choice fixed effects

	Whether ever admitted ages 0-7			Number of admissions ages 0-7		
	Untreated mean	TE	95% CI	Untreated mean	TE	95% CI
Any admission	0.750	0.0902	-0.151 - 0.332	2.520	0.335	-0.160 - 0.830
P-value		*0.465*			*0.184*	
Sharpened-Q value		*0.245*			*0.245*	
Any infection	0.610	0.147	-0.0740 - 0.367	1.620	0.370**	0.0765 - 0.664
P-value		*0.193*			*0.0135*	
Sharpened-Q value		*0.134*			*0.039*	
Days in hospital				11.480	4.683**	0.969 - 8.396
P-value					*0.0134*	
Sharpened-Q value					*0.039*	
<u>Admissions for infection</u>						
Viral	0.050	0.0833***	0.0256 - 0.141	0.060	0.0786***	0.0202 - 0.137
P-value		*0.00469*			*0.00839*	
Sharpened-Q value		*0.057*			*0.044*	
Bacterial	0.120	0.131***	0.0347 - 0.228	0.150	0.170**	0.0363 - 0.304
P-value		*0.00772*			*0.0127*	
Sharpened-Q value		*0.057*			*0.044*	
Upper respiratory	0.110	0.0692**	0.00880 - 0.130	0.150	0.0931**	0.0219 - 0.164
P-value		*0.0247*			*0.0104*	
Sharpened-Q value		*0.057*			*0.044*	
Lower respiratory	0.310	0.0855	-0.141 - 0.312	0.650	0.115	-0.0623 - 0.292
P-value		*0.460*			*0.203*	
Sharpened-Q value		*0.614*			*0.192*	
Soft tissue	0.270	0.0114	-0.192 - 0.215	0.390	-0.0823	-0.393 - 0.229
P-value		*0.912*			*0.604*	
Sharpened-Q value		*0.938*			*0.473*	
Gastrointestinal	0.230	-0.0654	-0.252 - 0.121	0.340	-0.0128	-0.145 - 0.120
P-value		*0.492*			*0.850*	
Sharpened-Q value		*0.538*			*0.473*	
Sample size		2,279			2,279	

Table E4: Adjustment for multiple hypothesis testing

Treatment is defined as exposure to income management within the mother's community before the child is 3 months old. Excludes admissions within 10 days of birth admission. Covariates: child's sex, whether they were the mother's first pregnancy, rainfall in the 90 days to birth, whether conceived in the wet season. Fixed effects for community at birth and quarter-year of birth. Standard errors clustered by community. Treatment effects and standard errors estimated using Boruysak, Jaravel and Spiess' (2021) estimator (Stata command: did_imputation). Q-values use Anderson's (2007) code, which is based on Benjamini Krieger Yekutieli (2006).

	N	Pr. Ever admitted	N admissions	Days in hospital	Pr ever admitted - infection	N admissions - infection
Benchmark: incl 1 year before rollout period	2,279	0.0902	0.335	4.683**	0.147	0.370**
		(0.123)	(0.253)	(1.895)	(0.113)	(0.150)
Include longer pre-period (conceived 2005 onwards)	2,743	0.103	0.314	4.026	0.174	0.395**
		(0.118)	(0.286)	(2.465)	(0.107)	(0.198)
No pre-period	1,258	0.135	1.002**	4.659***	0.247***	0.832***
		(0.105)	(0.453)	(1.063)	(0.0662)	(0.247)

Table E5: Length of pre-period

	Whether ever admitted ages 0-7			Number of admissions ages 0-7		
	Untreated mean	Treatment effect	95% confidence interval	Untreated mean	Treatment effect	95% confidence interval
Hospital inpatient admissions	0.78	0.09	-0.170 - 0.340	2.77	0.36	-0.143 - 0.872
Days in hospital				14.41	6.89***	3.746 - 10.03
Admissions for infection						
Any infection	0.61	0.13	-0.0990 - 0.354	1.64	0.39***	0.119 - 0.655
Viral	0.06	0.06***	0.0276 - 0.144	0.06	0.08***	0.0223 - 0.140
Bacterial	0.12	0.13***	0.0377 - 0.230	0.15	0.18***	0.0424 - 0.308
Upper respiratory	0.11	0.07**	0.0102 - 0.131	0.15	0.10***	0.0249 - 0.168
Lower respiratory	0.31	0.0905	-0.138 - 0.319	0.65	0.120	-0.0599 - 0.300
Soft tissue	0.27	-0.00944	-0.223 - 0.204	0.4	-0.0772	-0.367 - 0.212
Gastrointestinal	0.23	-0.0650	-0.252 - 0.122	0.34	-0.0116	-0.145 - 0.122
Genitourinary	0.04	0.0246	-0.0139 - 0.0631	0.04	0.0679	-0.0296 - 0.166
Sample size		2,279			2,279	

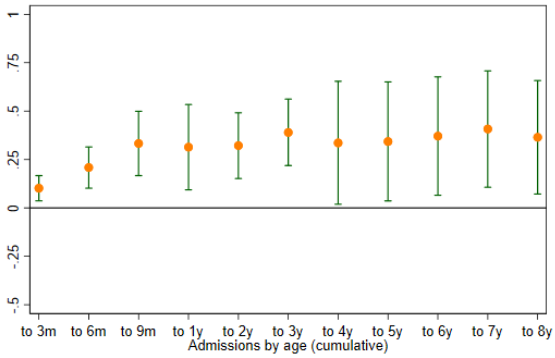
Table E6 Main results including admissions in first 10 days after birth discharge

Treatment is defined as exposure to income management within the mother's community before the child is 3 months old. Excludes birth admission. Covariates: child's sex, whether they were the mother's first pregnancy, rainfall in the 90 days to birth, whether conceived in the wet season. Fixed effects for community at birth and quarter-year of birth. Standard errors clustered by community. Treatment effects and standard errors estimated using Boruysak, Jaravel and Spiess' (2021) estimator (Stata command: did_imputation).

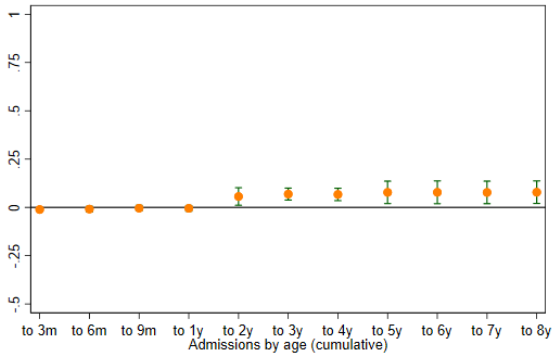
	Borusyak, Jaravel and Spiess imputation estimator		OLS fixed effects		Gardner two-stage DID		de Chaisemartin & D'Haultfoeuille	
	Treatment effect	95% confidence interval	Treatment effect	95% confidence interval	Treatment effect	95% confidence interval	Treatment effect	95% confidence interval
Ever admitted	0.0902	-0.151 - 0.332	0.0344	-0.04 to 0.11	0.0643	-0.11 to 0.24	-0.02	-0.21 to 0.18
Number of admissions	0.335	-0.160 - 0.830	0.353	-0.36 to 1.07	0.0912	-0.4 to 0.58	1.49	-0.84 to 3.83
Days in hospital	4.683**	0.969 - 8.396	4.450	-1.3 to 10.2	2.694	-0.72 to 6.11	5.46	-6.75 to 17.66
Ever admitted for infection	0.147	-0.0740 - 0.367	0.111**	0.01 to 0.21	0.0978	-0.07 to 0.26	0.11	-0.09 to 0.3
Number of admissions for infection	0.370**	0.0765 - 0.664	0.244	-0.37 to 0.86	0.176	-0.21 to 0.56	0.61	-0.72 to 1.94
Sample size	2,279		2,539		2,539		1,008	

Table E7 Alternative estimators

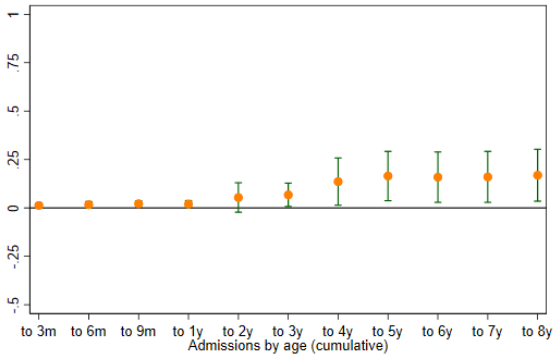
Appendix F: Additional estimates



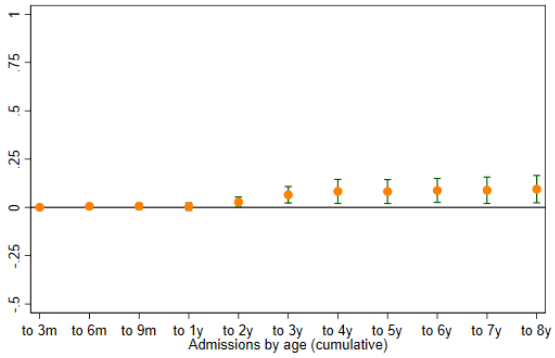
A. Any infection



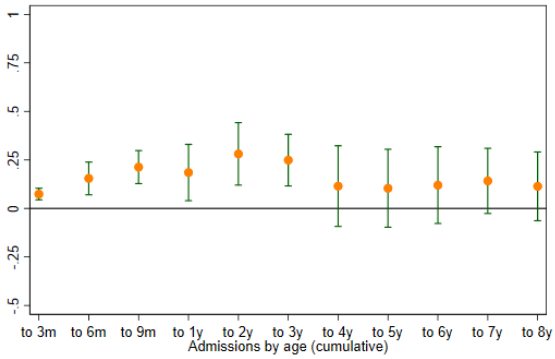
B. Viral infection



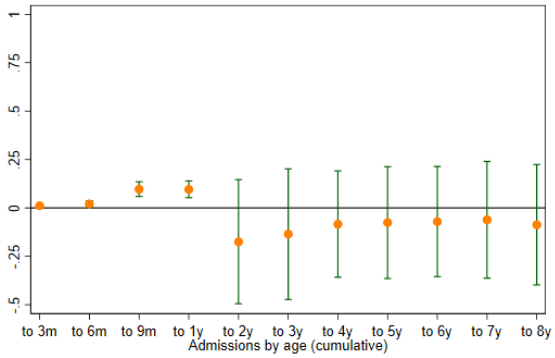
C. Bacterial infection



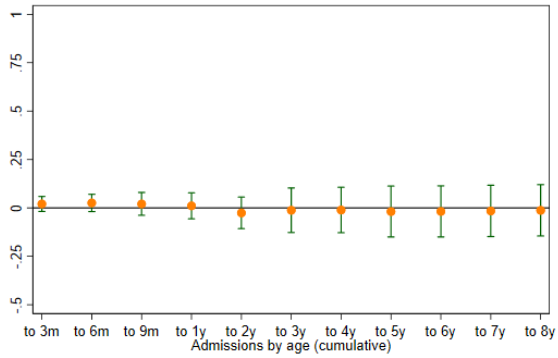
D. Upper respiratory



E. Lower respiratory



F. Soft tissue



G. Gastrointestinal

Figure F1: Treatment effects on number of admissions, by time horizon of outcome measure

	Untreated mean	Treatment effect
Any asthma diagnosis	0.04	0.033***
		(0.008)
Any allergy diagnosis	0.06	-0.006
		(0.040)
Sample size		1,191

Table F1: consultations for immune system-related diagnoses at community clinics, age 3-7

Note: clinic data available for 2/3rds of communities, and from 2010 onwards only. Analysis sample is limited to children born in communities that had a clinic contributing data for the full period 2010-2015. Allergies include allergic rhinitis, allergic contact dermatitis, allergy/allergic reaction not otherwise specified, and allergic conjunctivitis.

Appendix G: ICD 10 codes for infection

Source: Miller et al. 2016, with our additions in italics.

ICD-10 code	description	Included if secondary + diagnosis	Primary diagnosis may be accompanied by additional diagnoses
Gastrointestinal			
A00.0	Cholera due to Vibrio cholerae 01, biovar cholerae (ICD10)	0	0
A00.1	Cholera due to Vibrio cholerae 01, biovar eltor (ICD10)	0	0
A00.9	Cholera, unspecified (ICD10)	0	0
A01.0	Typhoid fever (ICD10)	0	0
A01.1	Paratyphoid fever A (ICD10)	0	0
A01.2	Paratyphoid fever B (ICD10)	0	0
A01.3	Paratyphoid fever C (ICD10)	0	0
A01.4	Paratyphoid fever, unspecified (ICD10)	0	0
A02.0	Salmonella enteritis (ICD10)	0	0
A02.2	Localised salmonella infections (ICD10)	0	1
A02.8	Other specified salmonella infections (ICD10)	0	0
A02.9	Salmonella infection, unspecified (ICD10)	0	0
A03.0	Shigellosis due to Shigella dysenteriae (ICD10)	0	0
A03.1	Shigellosis due to Shigella flexneri (ICD10)	0	0
A03.2	Shigellosis due to Shigella boydii (ICD10)	0	0
A03.3	Shigellosis due to Shigella sonnei (ICD10)	0	0
A03.8	Other shigellosis (ICD10)	0	0
A03.9	Shigellosis, unspecified (ICD10)	0	0
A04.0	Enteropathogenic Escherichia coli infection (ICD10)	0	0
A04.1	Enterotoxigenic Escherichia coli infection (ICD10)	0	0
A04.2	Enteroinvasive Escherichia coli infection (ICD10)	0	0
A04.3	Enterohemorrhagic Escherichia coli infection (ICD10)	0	0
A04.4	Other intestinal Escherichia coli infections (ICD10)	0	0
A04.5	Campylobacter enteritis (ICD10)	0	0
A04.6	Enteritis due to Yersinia enterocolitica (ICD10)	0	0
A04.7	Enterocolitis due to Clostridium difficile (ICD10)	0	0
A04.8	Other specified bacterial intestinal infections (ICD10)	0	0
A04.9	Bacterial intestinal infection, unspecified (ICD10)	0	0
A05.0	Foodborne staphylococcal intoxication (ICD10)	0	0
A05.1	Botulism food poisoning (ICD10)	0	0
A05.2	Foodborne Clostridium perfringens [Clostridium welchii] intoxication (ICD10)	0	0
A05.3	Foodborne Vibrio parahaemolyticus intoxication (ICD10)	0	0
A05.4	Foodborne Bacillus cereus intoxication (ICD10)	0	0
A05.8	Other specified bacterial foodborne intoxications (ICD10)	0	0
A05.9	Bacterial foodborne intoxication, unspecified (ICD10)	0	0
A06.0	Acute amebic dysentery (ICD10)	0	0
A06.1	Chronic intestinal amebiasis (ICD10)	0	0
A06.2	Amebic nondysenteric colitis (ICD10)	0	0
A06.3	Amoeboma of intestine (ICD10)	0	0
A06.4	Amebic liver abscess (ICD10)	0	0
A06.9	Amebiasis, unspecified (ICD10)	0	0
A07.0	Balantidiasis (ICD10)	0	0
A07.1	Giardiasis [lambliasis] (ICD10)	0	0
A07.2	Cryptosporidiosis (ICD10)	0	0
A07.3	Isosporiasis (ICD10)	0	0
A07.8	Other specified protozoal intestinal diseases (ICD10)	0	0
A07.9	Protozoal intestinal disease, unspecified (ICD10)	0	0
A08.0	Rotaviral enteritis (ICD10)	0	0
A08.1	Acute gastroenteropathy due to Norwalk agent (ICD10)	0	0
A08.2	Adenoviral enteritis (ICD10)	0	0
A08.3	Other viral enteritis (ICD10)	0	0
A08.4	Viral intestinal infection, unspecified	0	0

A08.5	Other specified intestinal infections (ICD10)	0	0
A09	Infectious gastroenteritis and colitis, unspecified (ICD10)	0	0
A09.0	Other and unspecified gastroenteritis and colitis of infectious origin	0	0
A09.9	Gastroenteritis and colitis of unspecified origin	0	0
A21.3	Gastrointestinal tularemia (ICD10)	0	0
A22.2	Gastrointestinal anthrax (ICD10)	0	0
B37.88	Candidiasis of other sites (ICD10)	0	0
B71.0	Hymenolepiasis (ICD10)	0	0
B71.9	Cestode infection, unspecified (ICD10)	0	0
B76.0	Ancylostomiasis (ICD10)	0	0
B78.9	Strongyloidiasis, unspecified (ICD10)	0	0
B81.0	Anisakiasis (ICD10)	0	0
B96.81	Helicobacter pylori [H. pylori] as the cause of diseases classified to other chapters (ICD10)	0	0
K52.8	Other specified noninfective gastroenteritis and colitis (ICD10)	0	0
B78.0	Intestinal strongyloidiasis, trichuriasis, enterobiasis or other intestinal parasites	0	0
B79		0	0
B80		0	0
B81.8		0	0
B82.0		0	0
B83.9		0	0
Bacterial		0	0
A02.1	Salmonella sepsis (ICD10)	0	0
A17.0	Tuberculous meningitis (ICD10)	0	1
A17.1	Meningeal tuberculoma (ICD10)	0	1
A17.8	Other tuberculosis of nervous system (ICD10)	0	1
A17.9	Tuberculosis of nervous system, unspecified (ICD10)	0	1
A18.0	Tuberculosis of bones and joints (ICD10)	0	1
A19.0	Acute miliary tuberculosis of a single specified site	0	0
A19.1	Acute miliary tuberculosis of multiple sites	0	0
A20.3	Plague meningitis	0	0
A20.7	Septicemic plague (ICD10)	0	0
A21.7	Generalized tularaemia	0	0
A22.7	Anthrax sepsis (ICD10)	0	0
A23.0	Brucellosis due to Brucella melitensis (ICD10)	0	0
A23.1	Brucellosis due to Brucella abortus (ICD10)	0	0
A23.2	Brucellosis due to Brucella suis (ICD10)	0	0
A23.3	Brucellosis due to Brucella canis (ICD10)	0	0
A23.8	Other brucellosis (ICD10)	0	0
A23.9	Brucellosis, unspecified (ICD10)	0	0
A24.1	Acute and fulminating melioidosis	0	0
A25.0	Spirillosis (ICD10)	0	0
A25.1	Streptobacillosis (ICD10)	0	0
A25.9	Rat-bite fever, unspecified (ICD10)	0	0
A32.1	Listerial meningitis and meningoencephalitis	0	0
A32.7	Listerial sepsis	0	0
A39.0	Meningococcal meningitis (ICD10)	0	1
A39.1	Waterhouse-Friderichsen syndrome (ICD10)	0	1
A39.2	Acute meningococcaemia	0	0
A39.3	Chronic meningococcaemia	0	0
A39.4	Meningococcemia, unspecified (ICD10)	0	0
A39.5	Meningococcal heart disease	0	0
A39.8	Other meningococcal infections (ICD10)	0	0
A39.8	Other meningococcal infections	0	1
A39.9	Meningococcal infection, unspecified (ICD10)	0	0
A40.0	Sepsis due to streptococcus, group A	0	0
A40.1	Sepsis due to streptococcus, group B	0	0
A40.2	Sepsis due to streptococcus, group D	0	0
A40.3	Sepsis due to Streptococcus pneumoniae (ICD10)	0	0
A40.8	Other streptococcal sepsis	0	0
A40.9	Streptococcal sepsis, unspecified (ICD10)	0	0
A41.0	Sepsis due to Staphylococcus aureus	0	0

A41.1	Sepsis due to other specified staphylococcus	0	0
A41.2	Sepsis due to unspecified staphylococcus (ICD10)	0	0
A41.3	Sepsis due to Hemophilus influenzae (ICD10)	0	0
A41.4	Sepsis due to anaerobes (ICD10)	0	0
A41.51	Sepsis due to Escherichia coli [E. Coli] (ICD10)	0	0
A41.52	Sepsis due to Pseudomonas (ICD10)	0	0
A41.58	Sepsis due to other Gram-negative organisms (ICD10)	0	0
A41.8	Other specified septicaemia (ICD10)	0	0
A41.9	Sepsis, unspecified (ICD10)	0	0
A44.0	Systemic bartonellosis	0	0
A48.3	Toxic shock syndrome	0	0
A52.0	Cardiovascular syphilis (I98.0*) (ICD10)	0	1
A52.1	Symptomatic neurosyphilis (ICD10)	0	0
A52.1	(no description found)	0	1
A52.2	Asymptomatic neurosyphilis (ICD10)	0	0
A52.3	Neurosyphilis, unspecified	0	0
A52.7	Other symptomatic late syphilis (ICD10)	0	0
A52.7	(no description found)	0	1
A52.8	Late syphilis, latent (ICD10)	0	0
A65	Nonvenereal syphilis (ICD10)	0	0
A74.0	Chlamydial conjunctivitis (ICD10)	0	0
A74.8	Other chlamydial diseases (ICD10)	0	0
A78	Q fever (ICD10)	0	0
A79.0	Trench fever (ICD10)	0	0
A79.1	Rickettsialpox due to Rickettsia akari (ICD10)	0	0
A79.8	Other specified rickettsioses (ICD10)	0	0
A79.9	Rickettsiosis, unspecified (ICD10)	0	0
B95.1	Streptococcus, group B, as the cause of diseases classified elsewhere (ICD10)	0	0
G00.0	Haemophilus meningitis (ICD10)	0	0
G00.1	Pneumococcal meningitis (ICD10)	0	0
G00.1	(no description found)	1	0
G00.2	Streptococcal meningitis (ICD10)	0	0
G00.3	Staphylococcal meningitis (ICD10)	0	0
G00.8	Other bacterial meningitis (ICD10)	0	0
G00.9	Bacterial meningitis, unspecified (ICD10)	0	0
G01	Meningitis in bacterial diseases classified elsewhere	0	0
G03.1	Chronic meningitis (ICD10)	0	0
G05.0	Encephalitis, myelitis and encephalomyelitis in bacterial diseases classified elsewhere	0	0
G06.0	Intracranial abscess and granuloma (ICD10)	0	0
G06.1	Intraspinal abscess and granuloma (ICD10)	0	0
G06.2	Extradural and subdural abscess, unspecified (ICD10)	0	0
G07	Intracranial and intraspinal abscess and granuloma in disease classified elsewhere (ICD10)	0	0
I30.1	Infective pericarditis	0	0
I32.0	Pericarditis in bacterial diseases classified elsewhere	0	0
I33.0	Acute and subacute infective endocarditis	0	0
I41.0	Myocarditis in bacterial diseases classified elsewhere	0	0
J39.0	Retropharyngeal and parapharyngeal abscess (ICD10)	0	0
J39.1	Other abscess of pharynx (ICD10)	0	0
J85.3	Abscess of mediastinum	0	0
M00.90	Pyogenic arthritis, unspecified, multiple sites (ICD10)	0	0
M00.91	Pyogenic arthritis, unspecified, shoulder region (ICD10)	0	0
M00.92	Pyogenic arthritis, unspecified, upper arm (ICD10)	0	0
M00.93	Pyogenic arthritis, unspecified, forearm (ICD10)	0	0
M00.94	Pyogenic arthritis, unspecified, hand (ICD10)	0	0
M00.95	Pyogenic arthritis, unspecified, pelvic region and thigh	0	0
M00.96	Pyogenic arthritis, unspecified, lower leg (ICD10)	0	0
M00.97	Pyogenic arthritis, unspecified, ankle and foot (ICD10)	0	0
M00.98	Pyogenic arthritis, unspecified, other site (ICD10)	0	0
M00.99	Pyogenic arthritis, unspecified, site unspecified (ICD10)	0	0
M01.0	Meningococcal arthritis	0	0
M01.1	Tuberculous arthritis	0	0

M46.2	Osteomyelitis of vertebra	0	0
M46.3	Infection of intervertebral disc (pyogenic)	0	0
M46.4	Discitis, unspecified	0	0
M46.5	Other infective spondylopathies	0	0
M49.0	Tuberculosis of spine	0	0
M49.1	Brucella spondylitis	0	0
M49.2	Enterobacterial spondylitis	0	0
M49.3	Spondylopathy in other infectious and parasitic diseases classified elsewhere	0	0
M72.6	Necrotizing fasciitis	0	0
M73.0	Gonococcal bursitis	0	0
M73.1	Syphilitic bursitis	0	0
M86.0	Acute haematogenous osteomyelitis	0	0
M86.10	Other acute osteomyelitis, unspecified site (ICD10)	0	0
M86.11	Other acute osteomyelitis, shoulder region (ICD10)	0	0
M86.12	Other acute osteomyelitis, upper arm (ICD10)	0	0
M86.13	Other acute osteomyelitis, forearm (ICD10)	0	0
M86.14	Other acute osteomyelitis, hand (ICD10)	0	0
M86.15	Other acute osteomyelitis, pelvic region and thigh (ICD10)	0	0
M86.16	Other acute osteomyelitis, lower leg (ICD10)	0	0
M86.17	Other acute osteomyelitis, ankle and foot (ICD10)	0	0
M86.18	Other acute osteomyelitis, other site (ICD10)	0	0
M86.19	Other acute osteomyelitis, multiple sites (ICD10)	0	0
M86.2	Subacute osteomyelitis	0	0
M86.4	Chronic osteomyelitis with draining sinus	0	0
M86.5	Other chronic osteomyelitis	0	0
M86.60	Other chronic osteomyelitis, unspecified site (ICD10)	0	0
M86.67	Other chronic osteomyelitis, ankle and foot (ICD10)	0	0
M86.68	Other chronic osteomyelitis, other site (ICD10)	0	0
M86.69	Other chronic osteomyelitis, multiple sites (ICD10)	0	0
M86.8	Other osteomyelitis	0	0
M86.90	Unspecified osteomyelitis, multiple sites (ICD10)	0	0
M86.91	Unspecified osteomyelitis, shoulder region (ICD10)	0	0
M86.92	Unspecified osteomyelitis, upper arm (ICD10)	0	0
M86.93	Unspecified osteomyelitis, forearm (ICD10)	0	0
M86.94	Unspecified osteomyelitis, hand (ICD10)	0	0
M86.95	Unspecified osteomyelitis, pelvic region and thigh (ICD10)	0	0
M86.96	Unspecified osteomyelitis, lower leg (ICD10)	0	0
M86.97	Unspecified osteomyelitis, ankle and foot (ICD10)	0	0
M86.98	Unspecified osteomyelitis, other site (ICD10)	0	0
M86.99	Unspecified osteomyelitis, site unspecified (ICD10)	0	0
M90.0	Tuberculosis of bone	0	0
P36.0	Sepsis of newborn due to streptococcus, group B	0	0
P36.1	Sepsis of newborn due to other and unspecified streptococci	0	0
P36.2	Sepsis of newborn due to Staphylococcus aureus	0	0
P36.3	Sepsis of newborn due to other and unspecified staphylococci	0	0
P36.4	Sepsis of newborn due to Escherichia coli	0	0
P36.5	Sepsis of newborn due to anaerobes	0	0
P36.8	Other bacterial sepsis of newborn	0	0
P36.9	Bacterial sepsis of newborn, unspecified	0	0
P37.0	Congenital tuberculosis	0	0
P37.2	Neonatal (disseminated) listeriosis	0	0
A38	Scarlet fever	0	0
A28.1	Cat-scratch disease	0	0
A18	Tuberculosis of other organs	0	0
A24.2	Melioidosis	0	0
A24.3		0	0
A24.4		0	0
A42.8	Other forms of actinomycosis	0	0
A49	Bacterial infection of unspecified site	0	0
B95.0	Stepcoccus or staphylococcus ac cause of disease classified elsewhere	1	0
B95.1		1	0
B95.2		1	0
B95.4		1	0

B95.5		1	0
B95.7		1	0
B95.8		1	0
Lower respiratory			
A06.5	Amebic lung abscess (ICD10)	0	1
A15.0	Tuberculosis of lung	0	0
A15.1	Tuberculosis of lung, confirmed by culture only	0	0
A15.2	Tuberculosis of lung, confirmed histologically	0	0
A15.3	Tuberculosis of lung, confirmed by unspecified means	0	0
A15.4	Tuberculosis of intrathoracic lymph nodes	0	0
A15.5	Tuberculosis of larynx, trachea and bronchus	0	0
A15.6	Tuberculosis pleurisy	0	0
A15.7	Primary respiratory tuberculosis	0	0
A15.8	Other respiratory tuberculosis	0	0
A15.9	Respiratory tuberculosis unspecified, confirmed bacteriologically and histologically	0	0
A16.0	Tuberculosis of lung, bacteriologically and histologically negative	0	0
A16.1	Tuberculosis of lung, bacteriological and histological examination not done	0	0
A16.2	Tuberculosis of lung, without mention of bacteriological or histological confirmation	0	0
A16.3	Tuberculosis of intrathoracic lymph nodes, without mention of bacteriological or histological confirmation	0	0
A16.4	Tuberculosis of larynx, trachea and bronchus, without mention of bacteriological or histological confirmation	0	0
A16.5	Tuberculous pleurisy, without mention of bacteriological or histological confirmation	0	0
A16.7	Primary respiratory tuberculosis, without mention of bacteriological or histological confirmation	0	0
A16.8	Other respiratory tuberculosis, without mention of bacteriological or histological confirmation	0	0
A16.9	Respiratory tuberculosis unspecified, without mention of bacteriological or histological confirmation	0	0
A19.2	Acute miliary tuberculosis, unspecified (ICD10)	0	0
A19.8	Other miliary tuberculosis (ICD10)	0	0
A19.9	Miliary tuberculosis, unspecified (ICD10)	0	0
A20.2	Pneumonic plague (ICD10)	0	0
A21.2	Pulmonary tularemia (ICD10)	0	0
A22.1	Pulmonary anthrax (ICD10)	0	0
A24.0	Glanders	0	0
A31.0	Pulmonary mycobacterial infection (ICD10)	0	0
A37.0	Whooping cough due to Bordetella pertussis (ICD10)	0	0
A37.1	Whooping cough due to Bordetella parapertussis (ICD10)	0	0
A37.8	Whooping cough due to other Bordetella species (ICD10)	0	0
A37.9	Whooping cough, unspecified (ICD10)	0	0
A42.0	Pulmonary actinomycosis (ICD10)	0	0
A48.1	Legionnaires disease	0	0
A70	Chlamydia psittaci infections (ICD10)	0	0
A70	(no description found)	0	1
B01.2	Varicella pneumonia (ICD10)	0	1
B05.2	Measles complicated by pneumonia (ICD10)	0	1
B34.0	Adenovirus infection, unspecified (ICD10)	0	0
B37.1	Pulmonary candidiasis (ICD10)	0	0
B38.0	Acute pulmonary coccidioidomycosis (ICD10)	0	0
B38.1	Chronic pulmonary coccidioidomycosis (ICD10)	0	0
B38.2	Pulmonary coccidioidomycosis, unspecified (ICD10)	0	0
B39.0	Acute pulmonary histoplasmosis capsulati	0	0
B39.1	Chronic pulmonary histoplasmosis capsulati	0	0
B39.2	Pulmonary histoplasmosis capsulati, unspecified (ICD10)	0	0
B39.3	Disseminated histoplasmosis capsulati	0	0
B39.4	Histoplasmosis capsulati, unspecified	0	0

B39.5	Histoplasmosis duboisii (ICD10)	0	0
B39.9	Histoplasmosis, unspecified (ICD10)	0	0
B40.0	Acute pulmonary blastomycosis	0	0
B40.1	Chronic pulmonary blastomycosis	0	0
B40.2	Pulmonary blastomycosis, unspecified	0	0
B40.3	Cutaneous blastomycosis	0	0
B40.7	Disseminated blastomycosis	0	0
B40.8	Other forms of blastomycosis	0	0
B41.0	Pulmonary paracoccidioidomycosis	0	0
B42.0	Pulmonary sporotrichosis	0	1
B44.0	Invasive pulmonary aspergillosis	0	0
B44.1	Other pulmonary aspergillosis	0	0
B45.0	Pulmonary cryptococcosis	0	0
B58.3	Pulmonary toxoplasmosis (ICD10)	0	1
B59	Pneumocystosis (ICD10)	0	0
J05.0	Acute obstructive laryngitis [croup] (ICD10)	0	0
J09	Influenza due to certain identified influenza virus	0	0
J10.0	Influenza with pneumonia, influenza virus identified (ICD10)	0	0
J10.1	Influenza due to other influenza virus with respiratory manifestations (ICD10)	0	0
J11.0	Influenza with pneumonia, virus not identified (ICD10)	0	0
J11.1	Influenza with other respiratory manifestations, virus not identified (ICD10)	0	0
J12.0	Adenoviral pneumonia (ICD10)	0	0
J12.1	Respiratory syncytial virus pneumonia (ICD10)	0	0
J12.2	Parainfluenza virus pneumonia (ICD10)	0	0
J12.3	Human metapneumovirus pneumonia	0	0
J12.8	Other viral pneumonia (ICD10)	0	0
J12.9	Viral pneumonia, unspecified (ICD10)	0	0
J13	Pneumonia due to Streptococcus pneumoniae (ICD10)	0	0
J14	Pneumonia due to Hemophilus influenzae (ICD10)	0	0
J15.0	Pneumonia due to Klebsiella pneumoniae (ICD10)	0	0
J15.1	Pneumonia due to Pseudomonas (ICD10)	0	0
J15.2	Pneumonia due to staphylococcus (ICD10)	0	0
J15.3	Pneumonia due to streptococcus, group B (ICD10)	0	0
J15.4	Pneumonia due to other streptococci (ICD10)	0	0
J15.5	Pneumonia due to Escherichia coli (ICD10)	0	0
J15.6	Pneumonia due to other aerobic Gram-negative bacteria (ICD10)	0	0
J15.7	Pneumonia due to Mycoplasma pneumoniae (ICD10)	0	0
J15.8	Pneumonia due to other specified bacteria (ICD10)	0	0
J15.9	Unspecified bacterial pneumonia (ICD10)	0	0
J16.0	Chlamydial pneumonia	0	0
J16.8	Pneumonia due to other specified infectious organisms (ICD10)	0	0
J17.0	Pneumonia in bacterial diseases classified elsewhere (ICD10)	1	0
J17.1	Pneumonia in viral diseases classified elsewhere (ICD10)	1	0
J17.2	Pneumonia in mycoses (ICD10)	1	0
J17.3	Pneumonia in parasitic diseases	0	0
J17.8	Pneumonia in other diseases classified elsewhere (ICD10)	1	0
J18.0	Bronchopneumonia, unspecified organism (ICD10)	0	0
J18.1	Lobar pneumonia, unspecified	0	0
J18.8	Other pneumonia, unspecified organism (ICD10)	0	0
J18.9	Pneumonia, unspecified	0	0
J20.0	Acute bronchitis due to Mycoplasma pneumoniae (ICD10)	0	0
J20.1	Acute bronchitis due to Hemophilus influenzae (ICD10)	0	0
J20.2	Acute bronchitis due to streptococcus (ICD10)	0	0
J20.3	Acute bronchitis due to coxsackievirus (ICD10)	0	0
J20.4	Acute bronchitis due to parainfluenza virus (ICD10)	0	0
J20.5	Acute bronchitis due to respiratory syncytial virus (ICD10)	0	0
J20.6	Acute bronchitis due to rhinovirus (ICD10)	0	0
J20.7	Acute bronchitis due to echovirus (ICD10)	0	0
J20.8	Acute bronchitis due to other specified organisms (ICD10)	0	0
J20.9	Acute bronchitis, unspecified (ICD10)	0	0
J21.0	Acute bronchiolitis due to respiratory syncytial virus	0	0
J21.1	Acute bronchiolitis due to human megapneumovirus	0	0
J21.8	Acute bronchiolitis due to other specified organisms	0	0

J21.9	Acute bronchiolitis, unspecified (ICD10)	0	0
J22	Unspecified acute lower respiratory infection (ICD10)	0	0
J22.0	(no description found)	0	0
J40	Bronchitis, not specified as acute or chronic	0	0
J41.0	Simple chronic bronchitis (ICD10)	0	0
J41.1	Mucopurulent chronic bronchitis (ICD10)	0	0
J41.8	Mixed simple and mucopurulent chronic bronchitis	0	0
J42	Unspecified chronic bronchitis (ICD10)	0	0
J44.0	Chronic obstructive pulmonary disease with acute lower respiratory infection	0	0
J47	Bronchiectasis (ICD10)	0	0
J65	Pneumoconiosis associated with tuberculosis	0	0
J85.0	Gangrene and necrosis of lung	0	0
J85.1	Abscess of lung with pneumonia	0	0
J85.2	Abscess of lung without pneumonia (ICD10)	0	0
J86.0	Pyothorax with fistula	0	0
J86.9	Pyothorax without fistula	0	0
P23.0	Congenital pneumonia due to viral agent	0	0
P23.2	Congenital pneumonia due to staphylococcus	0	0
P23.3	Congenital pneumonia due to staphylococcus, group B	0	0
P23.4	Congenital pneumonia due to Escherichia coli	0	0
P23.5	Congenital pneumonia due to Pseudomonas	0	0
P23.6	Congenital pneumonia due to other bacterial agents	0	0
P23.8	Congenital pneumonia due to other organisms	0	0
P23.9	Congenital pneumonia, unspecified	0	0
B95.6	Pneumoniae streptococcus	0	0
Upper respiratory			
A36.0	Pharyngeal diphtheria (ICD10)	0	0
A36.1	Nasopharyngeal diphtheria (ICD10)	0	0
A36.2	Laryngeal diphtheria (ICD10)	0	0
B05.3	Measles complicated by otitis media (ICD10)	0	1
H65.0	Acute serous otitis media (ICD10)	0	0
H65.1	Other acute nonsuppurative otitis media (ICD10)	0	0
H65.2	Chronic serous otitis media (ICD10)	0	0
H65.3	Chronic mucoid otitis media (ICD10)	0	0
H65.4	Other chronic nonsuppurative otitis media (ICD10)	0	0
H65.9	Nonsuppurative otitis media, unspecified (ICD10)	0	0
H66.0	Acute suppurative otitis media (ICD10)	0	0
H66.1	Chronic tubotympanic suppurative otitis media (ICD10)	0	0
H66.2	Chronic atticoantral suppurative otitis media (ICD10)	0	0
H66.3	Other chronic suppurative otitis media (ICD10)	0	0
H66.4	Suppurative otitis media, unspecified (ICD10)	0	0
H66.9	Otitis media, unspecified (ICD10)	0	0
H67.0	Otitis media in bacterial diseases classified elsewhere	0	0
H67.8	Otitis media in other diseases classified elsewhere (ICD10)	1	0
H68.0	Eustachian salpingitis	0	0
H70.0	Acute mastoiditis (ICD10)	0	0
H70.1	Chronic mastoiditis (ICD10)	0	0
H70.2	Petrositis (ICD10)	0	0
H70.8	Other mastoiditis and related conditions (ICD10)	0	0
H70.9	Mastoiditis, unspecified (ICD10)	0	0
H72.0	Central perforation of tympanic membrane (ICD10)	0	0
H72.1	Attic perforation of tympanic membrane (ICD10)	0	0
H72.2	Other marginal perforations of tympanic membrane (ICD10)	0	0
H72.8	Other perforations of tympanic membrane (ICD10)	0	0
H72.9	Perforation of tympanic membrane, unspecified (ICD10)	0	0
H73.0	Acute myringitis	0	0
H75.0	Mastoiditis in infectious and parasitic diseases classified elsewhere	0	0
H83.0	Labyrinthitis	0	0
H92.1	Otorrhoea	0	0
J00	Acute nasopharyngitis [common cold] (ICD10)	0	0
J01.0	Acute maxillary sinusitis (ICD10)	0	0

J01.1	Acute frontal sinusitis (ICD10)	0	0
J01.2	Acute ethmoidal sinusitis (ICD10)	0	0
J01.3	Acute sphenoidal sinusitis (ICD10)	0	0
J01.4	Acute pansinusitis	0	0
J01.8	Other acute sinusitis (ICD10)	0	0
J01.9	Acute sinusitis, unspecified (ICD10)	0	0
J02.0	Streptococcal pharyngitis (ICD10)	0	0
J02.8	Acute pharyngitis due to other specified organisms	0	0
J02.9	Acute pharyngitis, unspecified (ICD10)	0	0
J03.0	Streptococcal tonsillitis	0	0
J03.8	Acute tonsillitis due to other specified organisms	0	0
J03.9	Acute tonsillitis, unspecified (ICD10)	0	0
J04.0	Acute laryngitis (ICD10)	0	0
J04.1	Acute tracheitis (ICD10)	0	0
J04.2	Acute laryngotracheitis (ICD10)	0	0
J05.1	Acute epiglottitis (ICD10)	0	0
J06.0	Acute laryngopharyngitis (ICD10)	0	0
J06.8	Other acute upper respiratory infections of multiple sites (ICD10)	0	0
J06.9	Acute upper respiratory infection, unspecified (ICD10)	0	0
J32.0	Chronic maxillary sinusitis (ICD10)	0	0
J32.1	Chronic frontal sinusitis (ICD10)	0	0
J32.2	Chronic ethmoidal sinusitis (ICD10)	0	0
J32.3	Chronic sphenoidal sinusitis (ICD10)	0	0
J32.4	Chronic pansinusitis	0	0
J32.8	Other chronic sinusitis (ICD10)	0	0
J32.9	Chronic sinusitis, unspecified (ICD10)	0	0
J34.0	Abscess, furuncle and carbuncle of nose	0	0
J35.0	Chronic tonsillitis (ICD10)	0	0
J36	Peritonsillar abscess (ICD10)	0	0
Soft tissue			
A31.1	Cutaneous mycobacterial infection (ICD10)	0	0
A46	Erysipelas (ICD10)	0	0
B00.0	Eczema herpeticum (ICD10)	0	0
B35.0	Tinea barbae and tinea capitis	0	0
B35.1	Tinea unguium	0	0
B35.2	Tinea manuum	0	0
B35.3	Tinea pedis	0	0
B35.4	Tinea corporis	0	0
B35.5	Tinea imbricate	0	0
B35.6	Tinea cruris	0	0
B35.8	Other dermatophytoses	0	0
B35.9	Dermatophytosis, unspecified (ICD10)	0	0
B36.0	Pityriasis versicolour	0	0
B36.1	Tinea nigra	0	0
B36.2	White piedra	0	0
B36.3	Black piedra	0	0
B36.8	Other specified superficial mycoses	0	0
B37.2	Candidiasis of skin and nail (ICD10)	0	0
B85.0	Pediculosis due to Pediculus humanus capitis (ICD10)	0	0
B85.1	Pediculosis due to Pediculus humanus corporis (ICD10)	0	0
B85.2	Pediculosis, unspecified (ICD10)	0	0
B85.3	Phthiriasis (ICD10)	0	0
B85.4	Mixed pediculosis and phthiriasis (ICD10)	0	0
B86	Scabies (ICD10)	0	0
B87.9	Myiasis, unspecified (ICD10)	0	0
B88.0	Other acariasis (ICD10)	0	0
B88.1	Tungiasis [sandflea infestation] (ICD10)	0	0
B88.2	Other arthropod infestations (ICD10)	0	0
B88.3	External hirudiniasis (ICD10)	0	0
B88.8	Other specified infestations (ICD10)	0	0
B88.9	Infestation, unspecified (ICD10)	0	0
H60.3	Other infective otitis externa (ICD10)	0	0

L00	Staphylococcal scalded skin syndrome	0	0
L01.0	Impetigo [any organism] [any site] (ICD10)	0	0
L02.0	Cutaneous abscess, furuncle and carbuncle of face (ICD10)	0	0
L02.1	Cutaneous abscess, furuncle and carbuncle of neck (ICD10)	0	0
L02.2	Cutaneous abscess, furuncle and carbuncle of trunk (ICD10)	0	0
L02.3	Cutaneous abscess, furuncle and carbuncle of buttock (ICD10)	0	0
L02.4	Cutaneous abscess, furuncle and carbuncle of limb (ICD10)	0	0
L02.8	Cutaneous abscess, furuncle and carbuncle of other sites (ICD10)	0	0
L02.9	Cutaneous abscess, furuncle and carbuncle, unspecified (ICD10)	0	0
L03.01	Cellulitis of finger (ICD10)	0	0
L03.02	Cellulitis of toe (ICD10)	0	0
L03.10	Cellulitis of upper limb (ICD10)	0	0
L03.11	Cellulitis of lower limb (ICD10)	0	0
L03.2	Cellulitis of face (ICD10)	0	0
L03.3	Cellulitis of trunk (ICD10)	0	0
L03.8	Cellulitis of other sites (ICD10)	0	0
L03.9	Cellulitis, unspecified (ICD10)	0	0
L05.0	Pilonidal cyst with abscess	0	0
L05.9	Pilonidal cyst without abscess (ICD10)	0	0
L08.0	Pyoderma (ICD10)	0	0
L08.1	Erythrasma	0	0
L08.8	Other specified local infections of skin and subcutaneous tissue	0	0
L08.9	Local infection of the skin and subcutaneous tissue, unspecified	0	0
L13.0	Dermatitis herpetiformis	0	0
L30.3	Infective dermatitis	0	0
L88	Pyoderma gangrenosum	0	0
M60.09	Infective myositis, multiple sites	0	0
M63.0	Myositis in bacterial diseases classified elsewhere	0	0
M63.1	Myositis in protozoal and parasitic infections classified elsewhere	0	0
M63.2	Myositis in other infectious diseases classified elsewhere	0	0
M65.0	Abscess of tendon sheath	0	0
M65.1	Other infective (teno)synovitis	0	0
M68.0	Synovitis and tenosynovitis in bacterial diseases classified elsewhere	0	0
M71.0	Abscess of bursa	0	0
M71.1	Other infective bursitis	0	0
O91.00	Infection of nipple associated with childbirth, without mention of attachment difficulty	0	0
O91.10	Abscess of breast associated with childbirth, without mention of attachment difficulty	0	0
P39.0	Neonatal infective mastitis	0	0
P39.4	Neonatal skin infection	0	0
B95.6	Aureus staphylococcus	0	0
B87.4	Aural myiasis	0	0
B36.9	Superficial mycosis	0	0
A71	Trachoma	0	0
A43.1	Cutaneous nocardiosis	0	0
Viral			
A80.1	Acute paralytic poliomyelitis, wild virus, imported	0	0
A80.2	Acute paralytic poliomyelitis, wild virus, indigenous	0	0
A80.3	Acute paralytic poliomyelitis, other and unspecified (ICD10)	0	0
A80.4	Acute nonparalytic poliomyelitis (ICD10)	0	0
A80.9	Acute poliomyelitis, unspecified (ICD10)	0	0
A81.8	Other atypical virus infections of central nervous system (ICD10)	0	0
A81.9	Atypical virus infection of central nervous system, unspecified	0	0
A82.0	Sylvatic rabies	0	0
A82.1	Urban rabies	0	0
A82.9	Rabies, unspecified (ICD10)	0	0
A83.0	Japanese encephalitis (ICD10)	0	0
A83.1	Western equine encephalitis (ICD10)	0	0
A83.2	Eastern equine encephalitis (ICD10)	0	0
A83.3	St Louis encephalitis (ICD10)	0	0
A83.4	Australian encephalitis (ICD10)	0	0

A83.5	California encephalitis (ICD10)	0	0
A83.6	Rocio virus disease	0	0
A83.8	Other mosquito-borne viral encephalitis (ICD10)	0	0
A83.9	Mosquito-borne viral encephalitis, unspecified (ICD10)	0	0
A84.0	Far Eastern tick-borne encephalitis [Russian spring-summer encephalitis] (ICD10)	0	0
A84.1	Central European tick-borne encephalitis (ICD10)	0	0
A84.8	Other tick-borne viral encephalitis (ICD10)	0	0
A84.9	Tick-borne viral encephalitis, unspecified (ICD10)	0	0
A85.0	Enteroviral encephalitis	0	0
A85.1	Adenoviral encephalitis	0	0
A85.2	Arthropod-borne viral encephalitis, unspecified (ICD10)	0	0
A85.8	Other specified viral encephalitis	0	0
A86	Unspecified viral encephalitis	0	0
A87.0	Enteroviral meningitis (ICD10)	0	1
A87.1	Adenoviral meningitis (ICD10)	0	1
A87.2	Lymphocytic choriomeningitis (ICD10)	0	0
A87.8	Other viral meningitis	0	0
A87.9	Viral meningitis, unspecified	0	0
A88.0	Enteroviral exanthematous fever	0	0
A88.8	Other specified viral infections of central nervous system (ICD10)	0	0
A89	Unspecified viral infection of central nervous system (ICD10)	0	0
A90	Dengue fever [classical dengue] (ICD10)	0	0
A91	Dengue haemorrhagic fever	0	0
A92.0	Chikungunya virus disease	0	0
A92.1	O'nyong-nyong fever	0	0
A92.2	Venezuelan equine fever (ICD10)	0	0
A92.3	West Nile virus infection	0	0
A92.4	Rift Valley fever	0	0
A92.8	Other specified mosquito-borne viral fevers (ICD10)	0	0
A92.9	Mosquito-borne viral fever, unspecified	0	0
A93.0	Oropouche virus disease	0	0
A93.2	Colorado tick fever (ICD10)	0	0
A93.8	Other specified arthropod-borne viral fevers (ICD10)	0	0
A94	Unspecified arthropod-borne viral fever (ICD10)	0	0
A95.0	Sylvatic yellow fever (ICD10)	0	0
A95.1	Urban yellow fever (ICD10)	0	0
A95.9	Yellow fever, unspecified (ICD10)	0	0
A96.0	Junin haemorrhagic fever	0	0
A96.1	Machupo haemorrhagic fever	0	0
A96.2	Lassa fever	0	0
A96.8	Other arenaviral hemorrhagic fevers (ICD10)	0	0
A96.9	Arenaviral haemorrhagic fever, unspecified	0	0
A98.0	Crimean-Congo hemorrhagic fever (ICD10)	0	0
A98.1	Omsk hemorrhagic fever (ICD10)	0	0
A98.2	Kyasanur Forest disease (ICD10)	0	0
A98.3	Marburg virus disease	0	0
A98.4	Ebola virus disease	0	0
A98.5	Hemorrhagic fever with renal syndrome (ICD10)	0	0
A98.8	Other specified viral haemorrhagic fevers	0	0
A99	Unspecified viral haemorrhagic fever	0	0
B00.1	Herpesviral vesicular dermatitis (ICD10)	0	0
B00.2	Herpesviral gingivostomatitis and pharyngotonsillitis (ICD10)	0	0
B00.3	Herpesviral meningitis (ICD10)	0	1
B00.4	Herpesviral encephalitis (ICD10)	0	1
B00.5	Herpesviral ocular disease (ICD10)	0	1
B00.7	Disseminated herpesviral disease (ICD10)	0	0
B00.8	Other forms of herpesviral infection (ICD10)	0	0
B00.9	Herpesviral infection, unspecified (ICD10)	0	0
B01.0	Varicella meningitis	0	0
B01.1	Varicella encephalitis (G05.1*) (ICD10)	0	1
B01.8	Varicella with other complications (ICD10)	0	0
B01.9	Varicella without complication (ICD10)	0	0

B02.0	Zoster encephalitis	0	0
B02.1	Zoster meningitis (ICD10)	0	1
B02.2	Zoster with other nervous system involvement (ICD10)	0	1
B02.3	Zoster ocular disease (ICD10)	0	1
B02.7	Disseminated zoster	0	0
B02.8	Zoster with other complications (ICD10)	0	0
B02.9	Zoster without complications (ICD10)	0	0
B03	Smallpox (ICD10)	0	0
B04	Monkeypox	0	0
B05.0	Measles complicated by encephalitis (ICD10)	0	1
B05.1	Measles complicated by meningitis	0	0
B05.4	Measles with intestinal complications	0	0
B05.8	Measles with other complications (ICD10)	0	0
B05.9	Measles without complication (ICD10)	0	0
B06.0	Rubella with neurological complications (ICD10)	0	1
B06.8	Rubella with other complications (ICD10)	0	0
B06.9	Rubella without complication (ICD10)	0	0
B07	Viral warts (ICD10)	0	0
B08.0	Other orthopoxvirus infections (ICD10)	0	0
B08.1	Molluscum contagiosum (ICD10)	0	0
B08.2	Exanthema subitum [sixth disease]	0	0
B08.3	Erythema infectiosum [fifth disease] (ICD10)	0	0
B08.4	Enteroviral vesicular stomatitis with exanthem (ICD10)	0	0
B08.5	Enteroviral vesicular pharyngitis (ICD10)	0	0
B08.8	Other specified viral infections characterized by skin and mucous membrane lesions (ICD10)	0	0
B09	Unspecified viral infection characterized by skin and mucous membrane lesions (ICD10)	0	0
B15.0	Hepatitis A with hepatic coma (ICD10)	0	0
B15.9	Hepatitis A without hepatic coma (ICD10)	0	0
B16.0	Acute hepatitis B with delta-agent with hepatic coma (ICD10)	0	0
B16.1	Acute hepatitis B with delta-agent without hepatic coma (ICD10)	0	0
B16.2	Acute hepatitis B without delta-agent with hepatic coma (ICD10)	0	0
B16.9	Acute hepatitis B without delta-agent and without hepatic coma	0	0
B17.0	Acute delta-(super) infection of hepatitis B carrier (ICD10)	0	0
B17.1	Acute hepatitis C (ICD10)	0	0
B17.2	Acute hepatitis E (ICD10)	0	0
B17.8	Other specified acute viral hepatitis (ICD10)	0	0
B17.9	Acute viral hepatitis, unspecified	0	0
B18.0	Chronic viral hepatitis B with delta-agent (ICD10)	0	0
B18.1	Chronic viral hepatitis B without delta-agent (ICD10)	0	0
B18.2	Chronic viral hepatitis C (ICD10)	0	0
B18.8	Other chronic viral hepatitis	0	0
B18.9	Chronic viral hepatitis, unspecified	0	0
B19.0	Unspecified viral hepatitis with hepatic coma (ICD10)	0	0
B19.9	Unspecified viral hepatitis without hepatic coma (ICD10)	0	0
B20.0	HIV disease resulting in mycobacterial infection	0	0
B20.1	HIV disease resulting in other bacterial infections	0	0
B20.2	HIV disease resulting in cytomegaloviral disease	0	0
B20.3	HIV disease resulting in other viral infections	0	0
B20.4	HIV disease resulting in candidiasis	0	0
B20.5	HIV disease resulting in other mycoses	0	0
B20.6	HIV disease resulting in Pneumocystis jirovecii pneumonia	0	0
B20.7	HIV disease resulting in multiple infections	0	0
B20.8	HIV disease resulting in other infectious and parasitic diseases	0	0
B20.9	HIV disease resulting in unspecified infectious or parasitic	0	0
B21.0	HIV disease resulting in Kaposi sarcoma	0	0
B21.1	HIV disease resulting in Burkitt lymphoma	0	0
B21.2	HIV disease resulting in other types of non-Hodgkin lymphoma	0	0
B21.3	HIV disease resulting in other malignant neoplasms of lymphoid, haematopoietic and related tissue	0	0
B21.7	HIV disease resulting in multiple malignant neoplasms	0	0
B21.8	HIV disease resulting in other malignant neoplasms	0	0

B21.9	HIV disease resulting in unspecified malignant neoplasm	0	0
B22.0	HIV disease resulting in encephalopathy	0	0
B22.1	HIV disease resulting in lymphoid interstitial pneumonitis	0	0
B22.2	HIV disease resulting in wasting syndrome	0	0
B22.7	HIV disease resulting in multiple diseases classified elsewhere	0	0
B23.0	Acute HIV infection syndrome	0	0
B23.1	HIV disease resulting in (persistent) generalized lymphadenopathy	0	0
B23.2	HIV disease resulting in haematological and immunological abnormalities, not elsewhere classified	0	0
B23.8	HIV disease resulting in other specified conditions	0	0
B24	Unspecified human immunodeficiency virus [HIV] disease (ICD10)	0	0
B25.0	Cytomegaloviral pneumonitis	0	0
B25.1	Cytomegaloviral hepatitis	0	0
B25.2	Cytomegaloviral pancreatitis	0	0
B25.8	Other cytomegaloviral diseases	0	0
B25.9	Cytomegaloviral disease, unspecified (ICD10)	0	0
B26.0	Mumps orchitis (ICD10)	0	1
B26.1	Mumps meningitis (ICD10)	0	1
B26.2	Mumps encephalitis (ICD10)	0	1
B26.3	Mumps pancreatitis (ICD10)	0	1
B26.8	Mumps with other complications (ICD10)	0	0
B26.8	(no description found)	0	1
B26.9	Mumps without complication (ICD10)	0	0
B27.0	Gammaherpesviral mononucleosis	0	0
B27.1	Cytomegaloviral mononucleosis	0	0
B27.8	Other infectious mononucleosis	0	0
B27.9	Infectious mononucleosis, unspecified (ICD10)	0	0
B30.0	Keratoconjunctivitis due to adenovirus (ICD10)	0	1
B30.1	Conjunctivitis due to adenovirus (ICD10)	0	1
B30.2	(no description found)	0	1
B30.3	Acute epidemic hemorrhagic conjunctivitis (enteroviral)	0	1
B30.8	Other viral conjunctivitis (ICD10)	0	1
B30.9	Viral conjunctivitis, unspecified	0	0
B33.0	Epidemic myalgia (ICD10)	0	0
B33.1	Ross River disease	0	0
B33.2	Viral carditis	0	0
B33.3	Retrovirus infections, not elsewhere classified (ICD10)	0	0
B33.4	Hantavirus (cardio-) pulmonary syndrome	0	0
B33.8	Other specified viral diseases (ICD10)	0	0
B34.1	Enterovirus infection, unspecified (ICD10)	0	0
B34.2	Coronavirus infection, unspecified site	0	0
B34.3	Parvovirus infection, unspecified site	0	0
B34.4	Papovavirus infection, unspecified (ICD10)	0	0
B34.8	Other viral infections of unspecified site (ICD10)	0	0
B34.9	Viral infection, unspecified (ICD10)	0	0
B97.0	Adenovirus as the cause of diseases classified to other chapters	0	0
B97.1	Enterovirus as the cause of diseases classified to other chapters	0	0
B97.2	Coronavirus as the cause of diseases classified to other chapters	0	0
B97.3	Retrovirus as the cause of diseases classified to other chapters	0	0
B97.4	Respiratory syncytial virus as the cause of diseases classified to other chapters	0	0
B97.5	Reovirus as the cause of diseases classified to other chapters	0	0
B97.6	Parvovirus as the cause of diseases classified to other chapters	0	0
B97.7	Papillomavirus as the cause of diseases classified to other chapters	0	0
B97.8	Other viral agents as the cause of diseases classified to other chapters	0	0
G02.0	Meningitis in viral diseases classified elsewhere (ICD10)	1	0
G05.1	Encephalitis, myelitis and encephalomyelitis in viral diseases classified elsewhere (ICD10)	1	0
H19.1	Herpesviral keratitis and keratoconjunctivitis	0	0
I41.1	Myocarditis in viral diseases classified elsewhere	0	0
J10.8	Influenza with other manifestations, influenza virus identified	0	0
J11	Influenza, virus not identified (ICD10)	0	0
J11.8	Influenza with other manifestations, virus not identified (ICD10)	0	0
M01.4	Rubella arthritis	0	0

M01.50	Arthritis in other viral diseases classified elsewhere, multiple sites (ICD10)	1	0
M01.51	Arthritis in other viral diseases classified elsewhere, shoulder region (ICD10)	1	0
M01.52	Arthritis in other viral diseases classified elsewhere, upper arm (ICD10)	1	0
M01.53	Arthritis in other viral diseases classified elsewhere, forearm (ICD10)	1	0
M01.54	Arthritis in other viral diseases classified elsewhere, hand (ICD10)	1	0
M01.55	Arthritis in other viral diseases classified elsewhere, pelvic region and thigh (ICD10)	1	0
M01.56	Arthritis in other viral diseases classified elsewhere, lower leg (ICD10)	1	0
M01.57	Arthritis in other viral diseases classified elsewhere, ankle and foot (ICD10)	1	0
M01.58	Arthritis in other viral diseases classified elsewhere, other site (ICD10)	1	0
M01.59	Arthritis in other viral diseases classified elsewhere, site unspecified (ICD10)	1	0
O98.4	Viral hepatitis complicating pregnancy, childbirth and the puerperium	0	0
O98.5	Other viral diseases complicating pregnancy, childbirth and the puerperium (ICD10)	0	0
P35.0	Congenital rubella syndrome	0	0
P35.1	Congenital cytomegalovirus infection	0	0
P35.2	Congenital herpesviral [herpes simples] infection	0	0
P35.3	Congenital viral hepatitis	0	0
P35.8	Other congenital viral diseases	0	0
P35.9	Congenital viral disease, unspecified	0	0
Z21	Asymptomatic human immunodeficiency virus [HIV] infection status	0	0
Genitourinary			
A36.8	(no description found)	0	1
A51.0	Primary genital syphilis (ICD10)	0	0
A51.1	Primary anal syphilis (ICD10)	0	0
A51.2	Primary syphilis of other sites (ICD10)	0	0
A51.3	Secondary syphilis of skin and mucous membranes (ICD10)	0	0
A51.3	(no description found)	0	1
A51.4	Other secondary syphilis (ICD10)	0	0
A51.5	Early syphilis, latent (ICD10)	0	0
A51.9	Early syphilis, unspecified	0	0
A52.9	Late syphilis, unspecified (ICD10)	0	0
A53.0	Latent syphilis, unspecified as early or late (ICD10)	0	0
A53.9	Syphilis, unspecified (ICD10)	0	0
A54.0	Gonococcal infection of lower genitourinary tract without periurethral or accessory gland abscess (ICD10)	0	0
A54.1	Gonococcal infection of lower genitourinary tract with periurethral and accessory gland abscess (ICD10)	0	0
A54.2	Gonococcal pelviperitonitis and other gonococcal genitourinary infections (ICD10)	0	1
A54.3	Gonococcal infection of eye (ICD10)	0	0
A54.3	(no description found)	0	1
A54.4	Gonococcal infection of musculoskeletal system (ICD10)	0	1
A54.5	Gonococcal pharyngitis (ICD10)	0	0
A54.6	Gonococcal infection of anus and rectum (ICD10)	0	0
A54.8	Other gonococcal infections (ICD10)	0	0
A54.8	(no description found)	0	1
A54.9	Gonococcal infection, unspecified	0	0
A55	Chlamydial lymphogranuloma (venereum) (ICD10)	0	0
A56.0	Chlamydial infection of lower genitourinary tract (ICD10)	0	0
A56.1	Chlamydial infection of pelviperitoneum and other genitourinary organs (ICD10)	0	0
A56.2	Chlamydial infection of genitourinary tract, unspecified (ICD10)	0	0
A56.3	Chlamydial infection of anus and rectum (ICD10)	0	0
A56.4	Chlamydial infection of pharynx (ICD10)	0	0
A56.8	Sexually transmitted chlamydial infection of other sites (ICD10)	0	0
A57	Chancroid (ICD10)	0	0
A58	Granuloma inguinale (ICD10)	0	0
A59.0	Urogenital trichomoniasis (ICD10)	0	0
A59.0	(no description found)	0	1
A60.0	Herpesviral infection of genitalia and urogenital tract (ICD10)	0	0
A60.1	Herpesviral infection of perianal skin and rectum	0	0
A60.9	Anogenital herpesviral infection, unpsecified	0	0

A63.0	Anogenital (venereal) warts (ICD10)	0	0
A63.8	Other specified predominantly sexually transmitted diseases (ICD10)	0	0
A64	Unspecified sexually transmitted disease (ICD10)	0	0
B37.3	Candidiasis of vulva and vagina (ICD10)	0	1
B37.4	Candidiasis of other urogenital sites (ICD10)	0	1
N13.6	Pyonephrosis	0	0
N15.1	Renal and perinephric abscess (ICD10)	0	0
N30.0	Acute cystitis (ICD10)	0	0
N30.8	Other cystitis, abscess of bladder (ICD10)	0	0
N34.0	Urethral abscess (ICD10)	0	0
N35.1	Postinfective urethral stricture, not elsewhere classified (ICD10)	0	0
N39.0	Urinary tract infection, site not specified (ICD10)	0	0
N41.0	Acute prostatitis (ICD10)	0	0
N41.2	Abscess of prostate (ICD10)	0	0
N43.1	Infected hydrocele (ICD10)	0	0
N45.0	Orchitis, epididymitis and epididymo-orchitis with abscess (ICD10)	0	0
N45.9	Orchitis, epididymitis and epididymo-orchits without abscess	0	0
N51.2	Balanitis in diseases classified elsewhere	0	0
N70.0	Acute salpingitis and oophoritis (ICD10)	0	0
N73.0	Acute parametritis and pelvic cellulitis (ICD10)	0	0
N73.1	Chronic parametritis and pelvic cellulitis (ICD10)	0	0
N73.2	Unspecified parametritis and pelvic cellulitis	0	0
N73.3	Female acute pelvic peritonitis (ICD10)	0	0
N73.4	Female chronic pelvic peritonitis (ICD10)	0	0
N73.9	Female pelvic inflammatory disease, unspecified (ICD10)	0	0
N74.2	Female syphilitic pelvic inflammatory disease	0	0
N74.3	Female gonococcal pelvic inflammatory disease	0	0
N74.4	Female chlamydial pelvic inflammatory disease	0	0
N75.0	Cyst of Bartholins gland (ICD10)	0	0
N75.1	Abscess of Bartholins gland (ICD10)	0	0
N76.0	Acute vaginitis (ICD10)	0	0
N76.4	Abscess of vulva (ICD10)	0	0
O03.0	Spontaneous abortion, incomplete, complicated by genital tract and pelvic infection (ICD10)	0	0
O03.5	Spontaneous abortion, complete or unspecified, complicated by genital tract and pelvic infection (ICD10)	0	0
O08.0	Genital tract and pelvic infection following ectopic and molar pregnancy (ICD10)	0	0
O26.4	Herpes gestationis	0	0
O86.2	Urinary tract infection following delivery	0	0
O98.1	Syphilis complicating pregnancy, childbirth and the puerperium	0	0
O98.2	Gonorrhea complicating pregnancy, childbirth and the puerperium	0	0
P39.3	Neonatal urinary tract infection	0	0
A50	Congenital syphilis	0	0

Chapter 6: Conclusions

6.1 Main findings and contributions

In this thesis I have analysed the impact of two perinatal conditions on children's health at birth and their subsequent development. Below I reflect on the implications of my findings for our understanding of how the perinatal environment affects children's development, drawing on all four empirical chapters. I then turn to implications regarding the two specific conditions – heat exposure and benefits policy.

6.1.1 Implications for research on perinatal conditions

In the introduction, I set out three questions I aimed to contribute to through this thesis. Here I revisit each, with some broad conclusions.

How do different perinatal risks impact children in childhood?

I have studied the impacts of two changes to the perinatal environment: heat exposure and an income shock. Importantly, I study both in the same context and dataset, using consistent health outcome measures, which allows for some comparison between the two.

Both exposures increased hospital admissions in childhood, but they did so in different ways. The income shock affected admissions for infection, which I link to the immune programming hypothesis – as described further in Chapter 5. Heat exposure increased overall admissions in early childhood, but not for any specific type of diagnosis: instead, the number of diagnoses per admission increased. I take this to mean that while prenatal heat exposure did not cause conditions acute enough to require hospitalisation, it may have acted as a complicating factor for other unrelated conditions.

It may seem obvious that different prenatal exposures have different effects. But this is frequently forgotten. Research has often taken the approach of connecting the links from different studies to form a chain (i.e. prenatal heat exposure affects birth outcomes, birth

outcomes affect later outcomes). This 'links in a chain' approach is partly for practical reasons: the type of longitudinal data I use, with which I can study specific exposure and their effects throughout childhood, are not widely available. It has allowed me to analyse whether two exposures with similar effects on health at birth have the same health effects in childhood, and I find that they do not. As linked administrative data becomes more readily available for research, we will be able to gain a much richer understanding of the impacts of perinatal exposures on children's development, and how they vary.

Birth outcomes as the 'tip of the iceberg'

A related question is how much we can learn from analysis of birth outcomes alone. Other researchers such as Romero, Dey and Fisher (2014) in public health, and Conti et al. (2020) in economics, have long argued that measures of heath at birth such as birthweight and preterm birth are not a single 'outcome' but symptoms with many potential explanations. Despite this, the conceptualisation of birthweight and preterm birth as outcomes in their own right is persistent, in both research and policy discourse.

Proponents often argue that while birth outcomes are not perfect predictors of later life outcomes, these measures still have some predictive power. Fortunately their predictive power has declined over time, as Goisis, Özcan and Myrskylä (2017) show for the UK – likely due to advances in obstetric and neonatal care.

Providing comprehensive obstetric and neonatal care to small or preterm babies is very costly, and therefore even if birth outcomes had no bearing on later life health and development, it is still worthwhile to study how these outcomes, and accompanying healthcare costs, can be avoided. However, as I show in Chapters 4 and 5, these measures are poor indicators of who is affected by a specific prenatal exposure and by how much. For the prenatal exposures I study, I find that both reduce birthweight. However, the impacts on birthweight and other commonly used birth outcomes do little to explain longer-term outcomes.

I see two implications of these findings.

First, when studying the impacts of prenatal exposures, we should look beyond measures of health at birth, where possible. For instance, in Chapter 5, I find that a prediction of the impact of an income shock, based on its impact on birthweight, would capture only 7 percent of its full effect on childhood hospital admissions (excluding the birth admission). In this case, the policy's initial impact was just the tip of the iceberg. In other cases, it may be that some exposures affect birth outcomes but that their effects fade over time – this appears to be the case with moderate heat exposure: I found that while both high and moderate heat in utero reduced birthweight, it was only temperatures of 30 and above which consistently affected outcomes observed later in childhood (Table 1).

	Birthweight	**Childhood hospital admissions**	**School test scores**
Warm weather exposure (<30)	↓	No significant impact	No significant impact
Hot weather exposure (30+)	↓	↑	↓
Introduction of income management	↓	↑	Not studied

Table 1: summary of findings across Chapters 4 and 5

Second, where measures of health at birth are the only measures available – as is the case in many contexts where long-term or linked data are unavailable – my findings show that we can do better than focusing on birthweight and preterm birth alone. I find in Chapter 5 that using an expanded set of measures of health at birth (including Apgar scores, admissions to special care nurseries and obstetric complications) could predict around 30 percent of the impact of the policy shock on health outcomes in childhood. This is a marked improvement from 7 percent, but still, most of the effect remains unexplained. Further research could revisit this question in other populations, with the aim of identifying additional and alternative measures.

What have we learnt about pathways of early origins of human capital?

In the Introduction I set out three potential pathways that may link pre- and perinatal exposures to children's health and development: specific health conditions, developmental cascades, and developmental programming. What have we learned about which, if any, of these pathways may explain the impacts of the exposures I study?

The analysis in Chapter 4 provides most insight on this, as I was able to analyse both health and cognitive outcomes together. A first point to make is that while the patterns I see may be more consistent with one pathway, I cannot rule out other pathways. This is because although the data I study are rich, they do not tell us everything.

The data show that prenatal heat exposure did not increase hospital diagnoses for congenital conditions or any acute health condition. This is not conclusive evidence that heat exposure does not cause such conditions at all: it is possible that heat exposure causes conditions that are routinely diagnosed outside of hospital. But the effects on hospital admissions that I find are not explained by admissions for congenital conditions.

In addition, I do not find clear evidence of the interplay between health and cognitive outcomes that would reflect the 'developmental cascades' pathway; again, while this may partially explain the findings, it does not fully explain them.

Because health and cognitive development are so intertwined in early childhood, a lack of comprehensive healthcare data beyond hospital admissions limits how far I can take this analysis. However, I can say that the impacts of prenatal heat exposure on the health and cognitive outcomes that I measure appear independent of each other. This fact points towards fetal programming as the main causal pathway. That is, heat exposure may affect both health and cognitive outcomes, but the impacts on the cognitive outcomes I measure are not mediated through the health outcomes I can measure.

Similarly, in Chapter 5, the pattern of results I find points towards a specific type of fetal (and neonatal) programming: immune programming. Therefore, it looks like a substantial amount of the impacts of both exposures comes through developmental programming.

The important conclusion, reinforcing the 'tip of the iceberg' point above, is that I do not find a linear trajectory from perinatal exposure to poor health at birth, to poor health in childhood, to poorer performance at school. Instead, I find that prenatal heat exposure affects hospital admissions and school test scores even for children with no health concerns or complications and birth, and that a perinatal income shock can affect the risk of severe infection and asthma even for children whose health at birth was not compromised.

6.1.2 Heat exposure and climate change

While the effects of prenatal heat exposure that I find are small, they are important in contexts with high levels of heat exposure. The impact of typical variation in heat exposure in utero in the Northern Territory on children's school test scores equates to around thirty percent of the size of intensive early childhood interventions. While we generally discuss research on heat exposure in the context of predicted effects of climate change, it is important to remember that these findings reflect a cohort born from 2000 onwards, who are currently young adults. Climate change will likely accentuate such risks for future cohorts, but almost all children in the Northern Territory – as well as in many other parts of the world with similar climates – already experience at least some extreme heat exposure in utero.

The fact that I find these effects in Northern Territory, where residents have adapted to high levels of heat exposure, and air conditioning is present in most households, is somewhat surprising. Contrary to suggestions in previous research by, for example, Isen, Rossin-Slater and Walker (2017) for the US, and Chen et al. (2020) for China, it does not appear that increasing prevalence of air conditioning has eliminated the risks of prenatal heat exposure, at least in this context. We do not know precisely why this is: no

comprehensive data are available on use patterns of air conditioning. It may be that we need to better understand the risks of heat exposure and communicate them, alongside recommendations for how to best use such adaptations. It may also be an economic issue: even if households have an air conditioning unit and are aware of optimal use patterns, that does not mean they can afford the energy costs.

At face value, my findings have concerning implications. If heat exposure in utero leads to poorer health and cognitive development, this means children exposed in utero can expect lower human capital and higher healthcare costs (Kidd et al. 2023; Fishman, Carrillo, and Russ 2019; Yali Zhang et al. 2022). This is particularly concerning given projections that climate change will increase global temperatures and that this is just one of many mechanisms through which climate change is likely to affect human capital (Caruso, de Marcos, and Noy 2024).

However, my main takeaway from my analysis is that the effects of heat exposure, and their long-term costs, can be avoided. The impacts of seasonal risk in general, and heat exposure in particular, vary across climate zone and Aboriginal and non-Aboriginal populations. This suggests that the effects I find are not driven by an unavoidable biological mechanism linking heat exposure to these outcomes. It is not the case that a given level of exposure leads mechanically to a given outcome. Instead, there are likely multiple, interrelated, risks. This makes it much harder to research the impacts of heat exposure: the effects vary across contexts, and we cannot pinpoint a specific exposure period or threshold that consistently triggers the same effects. But this fact also gives hope: institutions, behaviours and the built environment can influence how prenatal heat exposure affects children's development. More research is needed to better understand these complex relationships and how they differ across contexts – this could involve both administrative data analysis similar to mine but in different contexts, and cohort studies, collecting new data on individual-level exposure to institutional, behavioural and build environment conditions affecting pregnant people, and linking this to outcomes for their children.

Such contextual differences are clear from international comparisons: for instance, Masselot et al. (2023) show that the mortality risk of a cold winter's day in London, UK is substantially larger than the mortality risk from a cold winter's day in Stockholm, Sweden – despite a cold day in Stockholm being objectively colder.

The fact that countries like Sweden have such small health effects of extreme cold weather exposure means they must be doing something right – with appropriate policy settings, we can reduce the health impacts of heat exposure as well. This requires treating the risks of hot summer temperatures the same way that countries with cold weather treat winter. I see three elements of such a strategy:

First, pregnant people should be provided with health advice specific to their climate; a challenge is that the research is lagging here in refining what that advice should be. In Chapter 2, I estimate that babies exposed to the hottest months of the year in utero have a 2 percentage point higher risk of preterm birth. In Chapter 3, I found that using a more complete heat exposure metric, the true risk may be closer to 4.5 percentage points. We do not yet have a good sense of what level and duration of heat exposure is most harmful (Baharav et al. 2023). But my tentative conclusion, based on my analysis of outcomes in childhood, is that prenatal exposure to temperatures above 30 degrees appear to pose the greatest risks of childhood hospitalisations and poorer school test performance. More research must be done estimating these impacts across other contexts to form a robust basis for more informed, context-specific health advice. In the meantime, advice should emphasise the importance of staying hydrated, avoiding extreme heat, as well as avoiding prolonged exposure even to hot temperatures that are seen in some contexts as relatively mild (e.g. in the 30-35 degree range).

Second, appropriate regulations to building standards and energy providers (Quilty et al. 2022) would help ensure all pregnant people have a well-insulated home that is safe during periods of extreme heat, and with reliable energy supply to power refrigeration, air conditioning and fans.

Third, while overall adoption of climate adaptations like air conditioning are high in countries like Australia, there are clearly socioeconomic inequalities in who has access. These inequalities should be addressed. Providing additional welfare payments to help low-income households pay for air conditioning or other adaptations during the hotter months of the year – along the lines of the UK's Winter Fuel Payment and Cold Weather Payment – could facilitate more equitable access.

6.1.3 Administration of benefits policy

Turning to my analysis of income management in Chapter 5, I have estimated the impact of the introduction of a policy which restricts what recipients can spend their money on, and where they can spend it. This was a 'conditionalisation' of existing cash transfers that recipients received.

Conditional cash transfers are more complex to administer, compared with unconditional cash transfers. This added complexity means two things: first, that administration is more costly (Margolies and Hoddinott 2015). In the case of income management, costs were particularly high during the policy rollout, but have remained high on an ongoing basis. Up to date data on costs are not public, but the Australian National Audit Office estimated that the administrative costs (i.e. excluding the value of the transfers themselves) were in the order of AU$6,600 per participant per year, 5 years after the rollout (Australian National Audit Office 2018). Second, the added complexity also means that it is more likely that something will go wrong in implementing the policy. This was indeed the case during the policy rollout, as I discuss in Chapter 5.

Beyond the effects of major implementation challenges that Chapter 5 analyses, we also see more subtle costs of income management on recipients, as highlighted in Marston et al.'s (2022) analysis of the policy's ongoing operation in other parts of Australia. Recipients must comply with restrictions over where they can spend their funds. They are unable to participate in the cash economy. They also experience time costs in requesting changes and exceptions to fund allocations, as well as the stress of relying on a single

electronic payment method which sometimes fails or is unavailable due to patchy connectivity.

I have emphasised the idea of seeing this policy through an 'early childhood lens'. Conditionality in the benefits system is designed to affect adults' behaviour: in this case, to influence what they spend their money on. However, such policies can have unintended effects on children.

It is generally understood that economic stress can negatively affect children's development contemporaneously (Masarik and Conger 2017). I find that, over and above these potential ongoing effects, the introduction of income management led to worse health for the cohort of children born around the time of the policy change.

In sum, income management led not only to an additional cost to public finances (in the order of AU$6,600 per participant) and inconvenience for recipients, but also worsened health for a cohort of children in communities that already faced a high burden of poor health. Despite having been in place for over 15 years, there is so far no evidence of any benefits from income management relative to unconditional cash transfers.

My analysis has broader implications for policies similar to income management. In terms of international comparisons, the closest are the UK's Healthy Start food vouchers system, and the USA's SNAP program – both of which provide means-tested funds to low-income households, with restrictions over what the money can be spent on.

Evaluations have found that these programs have improved outcomes for families and children (Bailey et al. 2020; Griffith, von Hinke, and Smith 2018). Globally, there has been much enthusiasm for conditional cash transfers in general, following their success in low and middle income countries. However, these 'successes' reflected the introduction of a new cash transfer, and should not be conflated with the impact of the conditionality. I find that when we analyse conditionality separately from the introduction of the payment, the conditionality itself may be harmful. Even if they are not harmful, it

is important to consider whether any benefits of the conditions imposed outweigh the additional costs of administering it.

Because my analysis focuses on children in remote Aboriginal communities, my estimates reflect the impact of an income shock on particularly disadvantaged groups with high pre-existing health risks within high-income countries. This differs from other studies analysing income shocks that affect a more general population (e.g. Borra et al. (2024)).

6.2 Limitations

Alongside its contributions, this thesis also has some limitations, and many open questions remain for future research.

A primary limitation is that I have not analysed a fully consistent set of measures across all four empirical chapters.

Where this relates to measurement of heat exposure, this reflects the way my understanding of the topic has evolved over time. In Chapter 2, I measured heat exposure primarily using trimester-average maximum and minimum temperatures, and indicators for extreme temperatures. As is likely evident, my understanding of how to conceptualise and measure heat exposure developed as I continued working on the topic, leading me to write Chapter 3, which focuses on measurement. The measures of heat exposure that I use in Chapter 2 are the same measures I criticise in Chapter 3.

Further, in Chapter 3, I recommend using measures that incorporate both moderate and extreme heat. Using moderate temperatures of 25-30 degrees as my omitted category, I found temperatures below this improved birth outcomes, indicating that temperatures of 25-30 degrees had some negative impact as well. In Chapter 4, I find that while those measures were important in explaining variation in birth outcomes, heat exposure below

30 degrees had no measured impact on childhood hospital admissions or school test scores, and I therefore used a simpler set of measures in the main analysis of that chapter.

A key limitation in interpreting these findings is that I do not know why we see these differences in the impacts of more extreme versus moderate heat. It may be that while moderate heat affects health at birth, these effects fade over time. Or it may be that there are longer-term effects of moderate heat, but these are not picked up on the institutional measures of health and cognitive development that I analyse.

A related issue is my use of comparable outcome measures. Across Chapters 4 and 5, I use one comparable outcome measure: hospital admissions in childhood. This allows me to compare outcomes qualitatively, though I stop short of comparing the magnitudes of effects, because I study a specific subpopulation in Chapter 5.

I would ideally have also measured the impact of both prenatal exposures on school test scores. The reason I cannot do this is because of the relatively small sample for analysis in Chapter 5. School attendance and test participation is particularly low in Aboriginal communities; therefore, I would be faced with both a decreased sample and potential selection bias. It may be possible to do this analysis, but it would require careful consideration, and was not something I could add on to my initial plan to focus on health outcomes.

My analysis also raises some unanswered questions, which I have not been able to fully explore with the data available to me. Of particular concern is the combination of humidity and air temperatures. We have good reasons to believe that humid heat affects health more than dry heat, and much climate research emphasises this (Vecellio et al. 2022; Rogers et al. 2021). But in Chapter 2 I find no impact of humidity on birth outcomes. In Chapter 3, I find no impact of wet bulb temperatures, a measure that combines air temperatures with humidity, on birth outcomes. As I describe in Chapter 3, further analysis, using both weather station observation and reanalysis data, and

interacting measures of heat, humidity and precipitation, may help to pinpoint whether certain combinations of these variables are particularly harmful.

Another limitation is that I am unable to pinpoint specific critical or sensitive periods within the perinatal exposure periods that I study. In all four empirical chapters I analyse exposures by trimester of pregnancy. However, I cannot draw any firm conclusions around whether any specific trimester matters most. This is partly because I do not find any consistent patterns. My main analysis in Chapter 4 suggests that third trimester heat exposure matters most for childhood health outcomes. But exposure to an income shock in any trimester, and even as a newborn, affects health outcomes. I find that heat exposure in any trimester can affect school test scores. In my heterogeneity analysis, I find that the magnitude and statistical significance of exposures in each trimester depend on the subpopulation studied. My takeaway is that any attempt to set out which period in pregnancy matters most for certain risks or outcomes must be grounded in an understanding of both biological processes and observations of parents' behaviour and attitudes. Analysis of large administrative datasets tells us that there is a causal impact of heat exposure or an income shock, when experienced in utero, and that families would benefit from avoiding these risks. But this type of analysis does not tell us much about how families can best avoid these risks: for instance, what material resources or healthcare should be provided in pregnancy to reduce the risks of income fluctuations on the baby, or how pregnant people can most efficiently reduce the risk of heat exposure with minimal disruption to their livelihood and way of life. Such analysis would require alternative research approaches, e.g. data collection on day-to-day exposures and experiences in pregnancy that increase risks at the individual level. This type of analysis is not possible with administrative records.

6.3 Future research

Throughout the thesis, I have discussed ideas for future research. But there are three specific avenues for further research which emerge taking account of the full thesis.

First, while I analyse two elements of the prenatal environment, more can be done to analyse the impacts of these and other prenatal exposures in other contexts. Further research that analyses multiple risks within a single population, using consistent measures, would be valuable in teaching us more about how different risks affect children, and how and whether the effects differ across contexts and cultures. In addition to heat exposure and policy shocks, such research could focus on other common exposures, such as access to and quality of prenatal care, disease prevalence, pollution and acute stressors.

Second, the continued research and policy attention on preterm birth and birthweight as predictors of longer-term outcomes reflects a practical necessity: if a policy intervention is provided in pregnancy, policymakers do not want to wait until the first cohort of affected children start school and undertake standardised tests to analyse its impact. Doing so could mean years of continuing a policy that may not be effective – or, alternatively, a lost opportunity to scale up a highly effective policy earlier. Policymakers need a 'sufficient statistic' (Chetty 2009) – or a set of such statistics – which provide a reliable indicator of the likely scale and direction of effects before outcomes have been realised.

This thesis suggests that preterm birth and birthweight are not sufficient statistics. To find such indicators, I have shown that we should look beyond standard measures of health at birth. I find there is some value in analysing multiple outcomes together, such as preterm birth, Apgar scores, and neonatal admissions. But it is likely that further analysis, taking a data-driven approach, could identify and verify other combinations of early life indicators that are more predictive of longer-term outcomes. This would be a complex task, and the appropriate indicators would likely vary depending on the specific outcome they aim to predict. But it would be immensely worthwhile. It would help to inform governments' cost-benefit analyses, as well as research efforts to improve such analyses

(e.g. Skarda, Asaria and Cookson's (2022) 'LifeSim' model for the UK). Beyond policy evaluation, identification of such indicators would allow us to better target remedial and preventative healthcare to reduce the impacts of these risks.

Finally, turning to heat exposure specifically: more research is needed to fully understand both the mechanisms through which heat exposure affects fetal development, and the long-term implications of these effects. Availability of administrative datasets with linked maternal and child health records could help to facilitate further analysis of the mechanisms: identifying maternal risk factors, health conditions experienced during pregnancy, and protective factors, and how these interact with heat exposure. This should be complemented with survey and observational data collected during pregnancy in cohort studies; while administrative records are valuable for this research, they do not give us the full picture of the institutional, behavioural and built environment factors that moderate the effect of heat exposure. Further analysis of additional outcomes in childhood, as well as the predictive power of childhood outcomes on adult outcomes, would also be immensely valuable, both as an addition to the literature on early origins of adult human capital in general, and specifically in understanding the implications of heat exposure on population health and productivity.

References

ABS. 2020. ‘Labour Force, Australia, Detailed - October 2020’. 6291.0.55.003. Canberra, Australia: Australian Bureau of Statistics. https://www.abs.gov.au/statistics/labour/employment-and-unemployment/labour-force-australia-detailed/oct-2020.

Adkins, Becky, Claude Leclerc, and Stuart Marshall-Clarke. 2004. ‘Neonatal Adaptive Immunity Comes of Age’. *Nature Reviews Immunology* 4 (7): 553–64. https://doi.org/10.1038/nri1394.

Ai, Hongshan, and Xiaoqing Tan. 2023. ‘The Effects of Exposure to High Temperatures during Pregnancy on Adolescent Mental Health: Evidence from China’. *China Economic Review* 80 (August):101991. https://doi.org/10.1016/j.chieco.2023.101991.

———. 2024. ‘The Impact of Early-Life Exposure to High Temperatures on Child Development: Evidence from China’. *Population and Environment* 46 (3): 18. https://doi.org/10.1007/s11111-024-00458-3.

Aizer, Anna. 2011. ‘Poverty, Violence and Health: The Impact of Domestic Violence During Pregnancy on Newborn Health’. *The Journal of Human Resources* 46 (3): 518–38. https://doi.org/10.1353/jhr.2011.0024.

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

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

Alderman, Harold, Jere R. Behrman, Victor Lavy, and Rekha Menon. 2001. ‘Child Health and School Enrollment: A Longitudinal Analysis’. *The Journal of Human Resources* 36 (1): 185–205. https://doi.org/10.2307/3069675.

Allotey, J, J Zamora, F Cheong-See, M Kalidindi, D Arroyo-Manzano, E Asztalos, Jam van der Post, et al. 2018. ‘Cognitive, Motor, Behavioural and Academic Performances of Children Born Preterm: A Meta-Analysis and Systematic Review Involving 64 061 Children’. *BJOG: An International Journal of Obstetrics & Gynaecology* 125 (1): 16–25. https://doi.org/10.1111/1471-0528.14832.

Almond, Douglas. 2006. ‘Is the 1918 Influenza Pandemic Over? Long-Term Effects of In Utero Influenza Exposure in the Post-1940 U.S. Population’. *Journal of Political Economy* 114 (4): 672–712. https://doi.org/10.1086/507154.

Almond, Douglas, and Janet Currie. 2011. 'Killing Me Softly: The Fetal Origins Hypothesis'. *Journal of Economic Perspectives* 25 (3): 153–72. https://doi.org/10.1257/jep.25.3.153.

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

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

Altman, Jon. 2017. 'Deepening Indigenous Poverty in the Northern Territory'. *Land Right News*, October 2017.

Amarante, Verónica, Marco Manacorda, Edward Miguel, and Andrea Vigorito. 2016. 'Do Cash Transfers Improve Birth Outcomes? Evidence from Matched Vital Statistics, Program, and Social Security Data'. *American Economic Journal: Economic Policy* 8 (2): 1–43. https://doi.org/10.1257/pol.20140344.

Anand, K. J. S., and Frank M. Scalzo. 2000. 'Can Adverse Neonatal Experiences Alter Brain Development and Subsequent Behavior?' *Neonatology* 77 (2): 69–82. https://doi.org/10.1159/000014197.

Andalón, Mabel, João Pedro Azevedo, Carlos Rodríguez-Castelán, Viviane Sanfelice, and Daniel Valderrama-González. 2016. 'Weather Shocks and Health at Birth in Colombia'. *World Development* 82 (June):69–82. https://doi.org/10.1016/j.worlddev.2016.01.015.

Anderson, Michael L. 2008. 'Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects'. *Journal of the American Statistical Association* 103 (484): 1481–95. https://doi.org/10.1198/016214508000000841.

Aquino, Taylor, Jennie E. Brand, and Florencia Torche. 2022. 'Unequal Effects of Disruptive Events'. *Sociology Compass* 16 (4): e12972. https://doi.org/10.1111/soc4.12972.

Attorney General's Department, Anne Poelina, and Ian Perdrisat. 2004. 'A Report of the Derby/West Kimberley Project : Working with Adolescents to Prevent Domestic Violence / Anne Poelina and Ian Perdrisat (Madjulla Incorporated).' Attorney General's Department. Canberra. https://trove.nla.gov.au/work/32848946?

Auger, Nathalie, William D. Fraser, Reg Sauve, -Bertrand Marianne Bilodeau, and Tom Kosatsky. 2017. 'Risk of Congenital Heart Defects after Ambient Heat Exposure Early in Pregnancy'. *Environmental Health Perspectives* 125 (1): 8–14. https://doi.org/10.1289/EHP171.

Australian Bureau of Statistics. 2008. ‘Environmental Issues: Energy Use and Conservation, Mar 2008’. c=AU; o=Commonwealth of Australia; ou=Australian Bureau of Statistics. https://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/4602.0.55.001Mar%202008?OpenDocument.

———. 2014. ‘Environmental Issues: Energy Use and Conservation, Mar 2014’. 4602.0.55.001. c=AU; o=Commonwealth of Australia; ou=Australian Bureau of Statistics. https://www.abs.gov.au/AUSSTATS/abs@.nsf/allprimarymainfeatures/A38DDA7F40718E43CA25750E00112A97?opendocument.

———. 2017. ‘Aboriginal and Torres Strait Islander Peoples Profile - Northern Territory’. 2002.0. 2016 Census of Population and Housing. https://quickstats.censusdata.abs.gov.au/census_services/getproduct/census/2016/communityprofile/7?opendocument.

———. 2021. ‘Migration, Australia’. https://www.abs.gov.au/statistics/people/population/migration-australia/latest-release.

———. 2022a. ‘Northern Territory: Aboriginal and Torres Strait Islander Population Summary | Australian Bureau of Statistics’. Canberra, ACT. https://www.abs.gov.au/articles/northern-territory-aboriginal-and-torres-strait-islander-population-summary.

———. 2022b. ‘Snapshot of Northern Territory’. 28 June 2022. https://www.abs.gov.au/articles/snapshot-nt-2021.

———. 2023. ‘Estimates of Aboriginal and Torres Strait Islander Australians’. ABS. https://www.abs.gov.au/statistics/people/aboriginal-and-torres-strait-islander-peoples/estimates-aboriginal-and-torres-strait-islander-australians/latest-release.

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.

———. 2020. ‘Antenatal Care Use and Outcomes for Aboriginal and Torres Strait Islander Mothers and Their Babies 2016-2017’. https://doi.org/10.25816/5Z3C-QC06.

———. 2022. ‘Australia’s Children’. Australian Government. https://www.aihw.gov.au/reports/children-youth/australias-children.

Australian National Audit Office. 2018. ‘The Implementation and Performance of the Cashless Debit Card Trial’. Performance Audit 1 2018-19. Canberra, Australia: The Auditor General. https://www.anao.gov.au/work/performance-audit/implementation-and-performance-cashless-debit-card-trial.

Baharav, Yuval, Lilly Nichols, Anya Wahal, Owen Gow, Kurt Shickman, Maya Edwards, and Katie Huffling. 2023. 'The Impact of Extreme Heat Exposure on Pregnant People and Neonates: A State of the Science Review'. *Journal of Midwifery & Women's Health* 68 (3): 324–32. https://doi.org/10.1111/jmwh.13502.

Bailey, Martha, Hilary Hoynes, Maya Rossin-Slater, and Reed Walker. 2020. 'Is the Social Safety Net a Long-Term Investment? Large-Scale Evidence from the Food Stamps Program'. w26942. Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w26942.

Bakhtsiyarava, Maryia, Kathryn Grace, and Raphael J. Nawrotzki. 2018. 'Climate, Birth Weight, and Agricultural Livelihoods in Kenya and Mali'. *American Journal of Public Health* 108 (Suppl 2): S144–50. https://doi.org/10.2105/AJPH.2017.304128.

Baldwin, Jane W., Tarik Benmarhnia, Kristie L. Ebi, Ollie Jay, Nicholas J. Lutsko, and Jennifer K. Vanos. 2023. 'Humidity's Role in Heat-Related Health Outcomes: A Heated Debate'. *Environmental Health Perspectives* 131 (5): 055001. https://doi.org/10.1289/EHP11807.

Baranowska-Rataj, Anna, Kieron Barclay, Joan Costa-Font, Mikko Myrskylä, and Berkay Özcan. 2023. 'Preterm Birth and Educational Disadvantage: Heterogeneous Effects'. *Population Studies* 77 (3): 459–74. https://doi.org/10.1080/00324728.2022.2080247.

Barker, D. J. 1998. 'In Utero Programming of Chronic Disease'. *Clinical Science (London, England: 1979)* 95 (2): 115–28.

Barreca, Alan, and Jessamyn Schaller. 2019. 'The Impact of High Ambient Temperatures on Delivery Timing and Gestational Lengths'. *Nature Climate Change*, December, 1–6. https://doi.org/10.1038/s41558-019-0632-4.

Bar-Zeev, Sarah J., Sue G. Kruske, Lesley M. Barclay, Naor H. Bar-Zeev, Jonathan R. Carapetis, and Sue V. Kildea. 2012. 'Use of Health Services by Remote Dwelling Aboriginal Infants in Tropical Northern Australia: A Retrospective Cohort Study'. *BMC Pediatrics* 12 (1): 19. https://doi.org/10.1186/1471-2431-12-19.

Beltran, Alyssa J., Jun Wu, and Olivier Laurent. 2014. 'Associations of Meteorology with Adverse Pregnancy Outcomes: A Systematic Review of Preeclampsia, Preterm Birth and Birth Weight'. *International Journal of Environmental Research and Public Health* 11 (1): 91–172. https://doi.org/10.3390/ijerph110100091.

Benjamini, Yoav, Abba M. Krieger, and Daniel Yekutieli. 2006. 'Adaptive Linear Step-up Procedures That Control the False Discovery Rate'. *Biometrika* 93 (3): 491–507. https://doi.org/10.1093/biomet/93.3.491.

Bermick, Jennifer, and Matthew Schaller. 2022. 'Epigenetic Regulation of Pediatric and Neonatal Immune Responses'. *Pediatric Research* 91 (2): 297–327. https://doi.org/10.1038/s41390-021-01630-3.

Bharadwaj, Prashant, Juan Pedro Eberhard, and Christopher A. Neilson. 2018. 'Health at Birth, Parental Investments, and Academic Outcomes'. *Journal of Labor Economics* 36 (2): 349–94. https://doi.org/10.1086/695616.

Bharadwaj, Prashant, Petter Lundborg, and Dan-Olof Rooth. 2018. 'Birth Weight in the Long Run'. *Journal of Human Resources* 53 (1): 189–231. https://doi.org/10.3368/jhr.53.1.0715-7235R.

Billings, Stephen B., and Kevin T. Schnepel. 2018. 'Life after Lead: Effects of Early Interventions for Children Exposed to Lead'. *American Economic Journal: Applied Economics* 10 (3): 315–44. https://doi.org/10.1257/app.20160056.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2016. 'Does Grief Transfer across Generations? Bereavements during Pregnancy and Child Outcomes'. *American Economic Journal: Applied Economics* 8 (1): 193–223. https://doi.org/10.1257/app.20140262.

Boardman, Jason D., Daniel A. Powers, Yolanda C. Padilla, and Robert A. Hummer. 2002. 'Low Birth Weight, Social Factors, and Developmental Outcomes among Children in the United States'. *Demography* 39 (2): 353–68. https://doi.org/10.1353/dem.2002.0015.

Bonthrone, Alexandra F, Ralica Dimitrova, Andrew Chew, Christopher J Kelly, Lucilio Cordero-Grande, Olivia Carney, Alexia Egloff, et al. 2021. 'Individualized Brain Development and Cognitive Outcome in Infants with Congenital Heart Disease'. *Brain Communications* 3 (2): fcab046. https://doi.org/10.1093/braincomms/fcab046.

Borra, Cristina, Ana Costa-Ramón, Libertad Gonzalez, and Almudena Sevilla. 2024. 'The Causal Effect of an Income Shock on Children's Human Capital'. *Journal of Labor Economics*, September. https://doi.org/10.1086/733052.

Borusyak, Kirill. 2023. 'DID_IMPUTATION: Stata Module to Perform Treatment Effect Estimation and Pre-Trend Testing in Event Studies'. *Statistical Software Components*, November. https://ideas.repec.org//c/boc/bocode/s458957.html.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess. 2024. 'Revisiting Event-Study Designs: Robust and Efficient Estimation'. *Review of Economic Studies* 91 (6): 3253–85. https://doi.org/10.1093/restud/rdae007.

Bourke, Claire D., James A. Berkley, and Andrew J. Prendergast. 2016. 'Immune Dysfunction as a Cause and Consequence of Malnutrition'. *Trends in Immunology* 37 (6): 386–98. https://doi.org/10.1016/j.it.2016.04.003.

Braga, Breno, Fredric Blavin, and Anuj Gangopadhyaya. 2020. 'The Long-Term Effects of Childhood Exposure to the Earned Income Tax Credit on Health Outcomes'. *Journal of Public Economics* 190 (October):104249. https://doi.org/10.1016/j.jpubeco.2020.104249.

Bratti, Massimiliano, Prince Boakye Frimpong, and Simone Russo. 2021. 'Prenatal Exposure to Heat Waves and Child Health in Sub-Saharan Africa'. IZA Discussion Paper 14424. https://docs.iza.org/dp14424.pdf.

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

Brimicombe, Chloe, Francesca Conway, Anayda Portela, Darshnika Lakhoo, Nathalie Roos, Chuansi Gao, Ijeoma Solarin, and Debra Jackson. 2024. 'A Scoping Review on Heat Indices Used to Measure the Effects of Heat on Maternal and Perinatal Health'. *BMJ Public Health* 2 (1). https://doi.org/10.1136/bmjph-2023-000308.

Brink, Nicholas, Darshnika P. Lakhoo, Ijeoma Solarin, Gloria Maimela, Peter von Dadelszen, Shane Norris, Matthew F. Chersich, et al. 2024. 'Impacts of Heat Exposure in Utero on Long-Term Health and Social Outcomes: A Systematic Review'. *BMC Pregnancy and Childbirth* 24 (1): 344. https://doi.org/10.1186/s12884-024-06512-0.

Bruckner, Tim A., Nhung T. H. Trinh, Nathalie Lelong, Kaveh Madani, Rémy Slama, Joanne Given, and Babak Khoshnood. 2024. 'Climate Change and Congenital Anomalies: A Population-Based Study of the Effect of Prolonged Extreme Heat Exposure on the Risk of Neural Tube Defects in France'. *Birth Defects Research* 116 (9): e2397. https://doi.org/10.1002/bdr2.2397.

Brunsdon, Chris, Jonathan Corcoran, Gary Higgs, and Andrew Ware. 2009. 'The Influence of Weather on Local Geographical Patterns of Police Calls for Service': *Environment and Planning B: Planning and Design*, January. https://doi.org/10.1068/b32133.

Callaway, Brantly, and Pedro H. C. Sant'Anna. 2020. 'Difference-in-Differences with Multiple Time Periods'. *Journal of Econometrics*, December. https://doi.org/10.1016/j.jeconom.2020.12.001.

Calvin, Katherine, Dipak Dasgupta, Gerhard Krinner, Aditi Mukherji, Peter W. Thorne, Christopher Trisos, José Romero, et al. 2023. 'IPCC, 2023: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (Eds.)]. IPCC, Geneva, Switzerland.' First. Intergovernmental Panel on Climate Change (IPCC). https://doi.org/10.59327/IPCC/AR6-9789291691647.

Caruso, Germán, Inés de Marcos, and Ilan Noy. 2024. 'Climate Changes Affect Human Capital'. *Economics of Disasters and Climate Change* 8 (1): 157–96. https://doi.org/10.1007/s41885-023-00140-2.

Carville, Kylie S., Deborah Lehmann, Gillian Hall, Hannah Moore, Peter Richmond, Nicholas de Klerk, and David Burgner. 2007. 'Infection Is the Major Component

of the Disease Burden in Aboriginal and Non-Aboriginal Australian Children: A Population-Based Study'. *The Pediatric Infectious Disease Journal* 26 (3): 210–16. https://doi.org/10.1097/01.inf.0000254148.09831.7f.

Chang, Grace, Marta Favara, and Rafael Novella. 2020. 'The Origins of Cognitive Skills and Non-Cognitive Skills: The Long-Term Effect of in-Utero Rainfall Shocks in India'. 13960. IZA Discussion Paper. IZA Institute of Labor Economics.

Charlot, Keyne, Juliette Millet, Florane Pasquier, Pauline Oustric, Graham Finlayson, Pascal Van Beers, Jonathan Monin, Fabien Sauvet, Pierre-Emmanuel Tardo-Dino, and Alexandra Malgoyre. 2022. 'The Impact of 16-h Heat Exposure on Appetite and Food Reward in Adults'. *Appetite* 177 (October):106144. https://doi.org/10.1016/j.appet.2022.106144.

Chen, Han-Yang, Sean C. Blackwell, and Suneet P. Chauhan. 2020. 'Association between Apgar Score at 5 Minutes and Adverse Outcomes among Low-Risk Pregnancies'. *The Journal of Maternal-Fetal & Neonatal Medicine* 0 (0): 1–8. https://doi.org/10.1080/14767058.2020.1754789.

Chen, Xi, Chih Ming Tan, Xiaobo Zhang, and Xin Zhang. 2020. 'The Effects of Prenatal Exposure to Temperature Extremes on Birth Outcomes: The Case of China'. *Journal of Population Economics*, March. https://doi.org/10.1007/s00148-020-00768-4.

Chersich, Matthew Francis, Minh Duc Pham, Ashtyn Areal, Marjan Mosalam Haghighi, Albert Manyuchi, Callum P. Swift, Bianca Wernecke, et al. 2020. 'Associations between High Temperatures in Pregnancy and Risk of Preterm Birth, Low Birth Weight, and Stillbirths: Systematic Review and Meta-Analysis'. *BMJ* 371 (November):m3811. https://doi.org/10.1136/bmj.m3811.

Chetty, Raj. 2009. 'Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods'. *Annual Review of Economics* 1 (Volume 1, 2009): 451–88. https://doi.org/10.1146/annurev.economics.050708.142910.

Cho, Hyunkuk. 2017. 'The Effects of Summer Heat on Academic Achievement: A Cohort Analysis'. *Journal of Environmental Economics and Management* 83 (May):185–96. https://doi.org/10.1016/j.jeem.2017.03.005.

Chodick, Gabriel, Shira Flash, Yonit Deoitch, and Varda Shalev. 2009. 'Seasonality in Birth Weight: Review of Global Patterns and Potential Causes'. *Human Biology* 81 (4): 463–77. https://doi.org/10.3378/027.081.0405.

Chorniy, Anna, Janet Currie, and Lyudmyla Sonchak. 2020. 'Does Prenatal WIC Participation Improve Child Outcomes?' *American Journal of Health Economics* 6 (2): 169–98. https://doi.org/10.1086/707832.

Chyn, Eric, Samantha Gold, and Justine Hastings. 2021. 'The Returns to Early-Life Interventions for Very Low Birth Weight Children'. *Journal of Health Economics* 75 (January):102400. https://doi.org/10.1016/j.jhealeco.2020.102400.

Cil, Gulcan, and Jiyoon Kim. 2022. 'Extreme Temperatures during Pregnancy and Adverse Birth Outcomes: Evidence from 2009 to 2018 U.S. National Birth Data'. *Health Economics* 31 (9): 1993–2024. https://doi.org/10.1002/hec.4559.

Clark, Dylan G., Elise H. Jackson, Corinne M. Hohl, and Kevin E. Liang. 2024. 'Extreme Heat Impacts on Acute Care: Examining Emergency Department Visits and Hospital Admissions during the 2021 British Columbia Heatwave'. *The Journal of Climate Change and Health* 17 (May):100310. https://doi.org/10.1016/j.joclim.2024.100310.

Clarke, Damian, Sonia Oreffice, and Climent Quintana-Domeque. 2019. 'The Demand for Season of Birth'. *Journal of Applied Econometrics* 34 (5): 707–23. https://doi.org/10.1002/jae.2711.

Cobb-Clark, Deborah A., Nathan Kettlewell, Stefanie Schurer, and Sven Silburn. 2021. 'The Effect of Quarantining Welfare on School Attendance in Indigenous Communities'. *Journal of Human Resources*, September, 1218. https://doi.org/10.3368/jhr.1218-9909R2.

Conti, Gabriella, Mark Hanson, Hazel M. Inskip, Sarah Crozier, Cyrus Cooper, and Keith Godfrey. 2020. 'Beyond Birth Weight: The Origins of Human Capital'. 13296. IZA Institute of Labor Economics. https://www.iza.org/publications/dp/13296.

Costa, Dora L. 2023. 'Overweight Grandsons and Grandfathers' Starvation Exposure'. *Journal of Health Economics* 91 (September):102796. https://doi.org/10.1016/j.jhealeco.2023.102796.

Cozzani, Marco, Siddartha Aradhya, and Alice Goisis. 2021. 'The Cognitive Development from Childhood to Adolescence of Low Birthweight Children Born after Medically Assisted Reproduction—a UK Longitudinal Cohort Study'. *International Journal of Epidemiology* 50 (5): 1514–23. https://doi.org/10.1093/ije/dyab009.

Crump, Casey. 2020. 'An Overview of Adult Health Outcomes after Preterm Birth'. *Early Human Development*, Neonatal Update Special Edition, 150 (November):105187. https://doi.org/10.1016/j.earlhumdev.2020.105187.

CSIRO, and The Bureau of Meteorology. 2022. 'State of The Climate 2022'. Commonwealth of Australia. http://www.bom.gov.au/state-of-the-climate/index.shtml.

Cunha, Flavio, and James Heckman. 2007. 'The Technology of Skill Formation' 97 (2).

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

Currie, Janet, and Hannes Schwandt. 2013. 'Within-Mother Analysis of Seasonal Patterns in Health at Birth'. *Proceedings of the National Academy of Sciences* 110 (30): 12265–70. https://doi.org/10.1073/pnas.1307582110.

Dadvand, Payam, Xavier Basagaña, Claudio Sartini, Francesc Figueras, Martine Vrijheid, Audrey de Nazelle, Jordi Sunyer, and Mark J. Nieuwenhuijsen. 2011. 'Climate Extremes and the Length of Gestation'. *Environmental Health Perspectives* 119 (10): 1449–53. https://doi.org/10.1289/ehp.1003241.

Davis, Robert E., Glenn R. McGregor, and Kyle B. Enfield. 2016. 'Humidity: A Review and Primer on Atmospheric Moisture and Human Health'. *Environmental Research* 144 (January):106–16. https://doi.org/10.1016/j.envres.2015.10.014.

De Gendre, Alexandra, John Lynch, Aurélie Meunier, Rhiannon Pilkington, and Stefanie Schurer. 2021. 'Child Health and Parental Responses to an Unconditional Cash Transfer at Birth'. IZA Discussion Paper Series 14693. IZA Institute of Labor Economics. https://www.ssrn.com/abstract=3917308.

de Chaisemartin, Clément, and Xavier D'Haultfœuille. 2023. 'Two-Way Fixed Effects and Differences-in-Differences with Heterogeneous Treatment Effects: A Survey'. *The Econometrics Journal* 26 (3): C1–30. https://doi.org/10.1093/ectj/utac017.

Defeyter, Margaret Anne, Marion Hetherington, Mike McKean, and Andrew Forsey. 2022. 'The Bungled Digitisation of Healthy Start Is Hampering Low Income Families' Access to Healthy Food'. *BMJ* 377 (June):o1462. https://doi.org/10.1136/bmj.o1462.

Dell, Melissa, Benjamin F. Jones, and Benjamin A. Olken. 2014. 'What Do We Learn from the Weather? The New Climate-Economy Literature'. *Journal of Economic Literature* 52 (3): 740–98. https://doi.org/10.1257/jel.52.3.740.

Deschênes, Olivier, Michael Greenstone, and Jonathan Guryan. 2009. 'Climate Change and Birth Weight'. *American Economic Review* 99 (2): 211–17. https://doi.org/10.1257/aer.99.2.211.

DiPietro, J. A., and K. M. Voegtline. 2017. 'The Gestational Foundation of Sex Differences in Development and Vulnerability'. *Neuroscience*, Early Adversity and Brain Development, 342 (February):4–20. https://doi.org/10.1016/j.neuroscience.2015.07.068.

Dobbins, Timothy A., Elizabeth A. Sullivan, Christine L. Roberts, and Judy M. Simpson. 2012. 'Australian National Birthweight Percentiles by Sex and Gestational Age, 1998–2007'. *The Medical Journal of Australia* 197 (5): 291–94. https://doi.org/10.5694/mja11.11331.

Dorélien, Audrey. 2016. 'Birth Seasonality in Sub-Saharan Africa'. *Demographic Research* 34 (27): 761–96. https://doi.org/10.4054/DemRes.2016.34.27.

———. 2019. 'The Effects of In Utero Exposure to Influenza on Birth and Infant Outcomes in the US'. *Population and Development Review* 45 (3): 489–523. https://doi.org/10.1111/padr.12232.

Doyle, Mary-Alice. 2023. 'Seasonal Patterns in Newborns' Health: Quantifying the Roles of Climate, Communicable Disease, Economic and Social Factors'. *Economics & Human Biology* 51 (December):101287. https://doi.org/10.1016/j.ehb.2023.101287.

Doyle, Mary-Alice, Stefanie Schurer, and Sven Silburn. 2022. 'Unintended Consequences of Welfare Reform: Evidence from Birthweight of Aboriginal Children in Australia'. *Journal of Health Economics* 84 (July):102618. https://doi.org/10.1016/j.jhealeco.2022.102618.

Doyle, Orla. 2024. 'Can Early Intervention Have a Sustained Effect on Human Capital?' *Journal of Human Resources* 59 (5): 1599–1636. https://doi.org/10.3368/jhr.0321-11557R1.

Ebi, Kristie L., Anthony Capon, Peter Berry, Carolyn Broderick, Richard de Dear, George Havenith, Yasushi Honda, et al. 2021. 'Hot Weather and Heat Extremes: Health Risks'. *The Lancet* 398 (10301): 698–708. https://doi.org/10.1016/S0140-6736(21)01208-3.

Eshaghnia, Sadegh, James J Heckman, and Rasmus Landersø. 2024. 'The Impact of the Level and Timing of Parental Resources on Child Development and Intergenerational Mobility'. Discussion Paper Series. https://www.iza.org/en/publications/dp/16784.

Fishman, Ram, Paul Carrillo, and Jason Russ. 2019. 'Long-Term Impacts of Exposure to High Temperatures on Human Capital and Economic Productivity'. *Journal of Environmental Economics and Management* 93 (January):221–38. https://doi.org/10.1016/j.jeem.2018.10.001.

Foster, Tim, and Nina L. Hall. 2021. 'Housing Conditions and Health in Indigenous Australian Communities: Current Status and Recent Trends'. *International Journal of Environmental Health Research* 31 (3): 325–43. https://doi.org/10.1080/09603123.2019.1657074.

Gardner, John. 2021. 'Two-Stage Differences in Differences'. Department of Economics, University of Mississippi. https://jrgcmu.github.io/2sdd_current.pdf.

Garfinkel, Irwin, Laurel Sariscsany, Elizabeth Ananat, Sophie Collyer, Robert P. Hartley, Buyi Wang, and Christopher Wimer. 2022. 'The Benefits and Costs of a Child Allowance'. *Journal of Benefit-Cost Analysis* 13 (3): 335–62. https://doi.org/10.1017/bca.2022.15.

Georgieff, Michael K., Katya E. Brunette, and Phu V. Tran. 2015. 'Early Life Nutrition and Neural Plasticity'. *Development and Psychopathology* 27 (2): 411–23. https://doi.org/10.1017/S0954579415000061.

Gluckman, P. D., M. A. Hanson, and T. Buklijas. 2010. 'A Conceptual Framework for the Developmental Origins of Health and Disease'. *Journal of Developmental Origins of Health and Disease* 1 (1): 6–18. https://doi.org/10.1017/S2040174409990171.

Goisis, Alice, Berkay Özcan, and Mikko Myrskylä. 2017. 'Decline in the Negative Association between Low Birth Weight and Cognitive Ability'. *Proceedings of the National Academy of Sciences* 114 (1): 84–88. https://doi.org/10.1073/pnas.1605544114.

Goldenberg, Robert L, Jennifer F Culhane, Jay D Iams, and Roberto Romero. 2008. 'Epidemiology and Causes of Preterm Birth'. *Lancet (London, England)* 371 (9606): 75–84. https://doi.org/10.1016/S0140-6736(08)60074-4.

González, Libertad, and Sofia Trommlerová. 2022. 'Cash Transfers before Pregnancy and Infant Health'. *Journal of Health Economics* 83 (May):102622. https://doi.org/10.1016/j.jhealeco.2022.102622.

Gooley, Cameron. 2021. 'Northern Australia's Decade-Long Syphilis Outbreak Prompts Calls for a National Response'. *Australian Broadcasting Corporation*, 14 January 2021. https://www.abc.net.au/news/2021-01-15/decade-long-syphilis-outbreak-in-northern-australia/13058166.

Govers, Coen, Philip C. Calder, Huub F. J. Savelkoul, Ruud Albers, and R. J. Joost van Neerven. 2022. 'Ingestion, Immunity, and Infection: Nutrition and Viral Respiratory Tract Infections'. *Frontiers in Immunology* 13. https://www.frontiersin.org/articles/10.3389/fimmu.2022.841532.

Grace, Kathryn, Frank Davenport, Heidi Hanson, Christopher Funk, and Shraddhanand Shukla. 2015. 'Linking Climate Change and Health Outcomes: Examining the Relationship between Temperature, Precipitation and Birth Weight in Africa'. *Global Environmental Change* 35 (November):125–37. https://doi.org/10.1016/j.gloenvcha.2015.06.010.

Grassly, Nicholas C, and Christophe Fraser. 2006. 'Seasonal Infectious Disease Epidemiology'. *Proceedings of the Royal Society B: Biological Sciences* 273 (1600): 2541–50. https://doi.org/10.1098/rspb.2006.3604.

Griffith, Rachel, Stephanie von Hinke, and Sarah Smith. 2018. 'Getting a Healthy Start: The Effectiveness of Targeted Benefits for Improving Dietary Choices'. *Journal of Health Economics* 58 (March):176–87. https://doi.org/10.1016/j.jhealeco.2018.02.009.

Grossman, Michael. 1972. 'On the Concept of Health Capital and the Demand for Health'. *Journal of Political Economy*, March. https://doi.org/10.1086/259880.

Gunnsteinsson, Snaebjorn, Teresa Molina, Achyuta Adhvaryu, Parul Christian, Alain Labrique, Jonathan Sugimoto, Abu Ahmed Shamim, and Keith P. West. 2022. 'Protecting Infants from Natural Disasters: The Case of Vitamin A Supplementation and a Tornado in Bangladesh'. *Journal of Development Economics* 158 (September):102914. https://doi.org/10.1016/j.jdeveco.2022.102914.

Guthridge, Steven, Lin Li, Sven Silburn, Shu Qin Li, John McKenzie, and John Lynch. 2016. 'Early Influences on Developmental Outcomes among Children, at Age 5,

in Australia's Northern Territory'. *Early Childhood Research Quarterly*, International research utilizing the Early Development Instrument (EDI) as a measure of early child development, 35 (April):124–34. https://doi.org/10.1016/j.ecresq.2015.12.008.

Hawkins, Amelia A., Christopher A. Hollrah, Sarah Miller, Laura R. Wherry, Gloria Aldana, and Mitchell D. Wong. 2023. 'The Long-Term Effects of Income for At-Risk Infants: Evidence from Supplemental Security Income'. Working Paper. Working Paper Series. National Bureau of Economic Research. https://doi.org/10.3386/w31746.

He, Siyi, Tom Kosatsky, Audrey Smargiassi, Marianne Bilodeau-Bertrand, and Nathalie Auger. 2018. 'Heat and Pregnancy-Related Emergencies: Risk of Placental Abruption during Hot Weather'. *Environment International* 111 (February):295–300. https://doi.org/10.1016/j.envint.2017.11.004.

Heckman, James J. 2007. 'The Economics, Technology, and Neuroscience of Human Capability Formation'. *Proceedings of the National Academy of Sciences of the United States of America* 104 (33): 13250. https://doi.org/10.1073/pnas.0701362104.

Heinze, Georg, Christine Wallisch, and Daniela Dunkler. 2018. 'Variable Selection – A Review and Recommendations for the Practicing Statistician'. *Biometrical Journal* 60 (3): 431–49. https://doi.org/10.1002/bimj.201700067.

HIGH Horizons, and World Health Organization. 2024. 'Conceptual Framework on Extreme Heat and Maternal, Newborn and Child Health'. https://www.high-horizons.eu/conceptual-framework-on-extreme-heat-and-maternal-newborn-and-child-health/.

Hoynes, Hilary, Diane Whitmore Schanzenbach, and Douglas Almond. 2016. 'Long-Run Impacts of Childhood Access to the Safety Net'. *American Economic Review* 106 (4): 903–34. https://doi.org/10.1257/aer.20130375.

Hu, Zihan, and Teng Li. 2019. 'Too Hot to Handle: The Effects of High Temperatures during Pregnancy on Adult Welfare Outcomes'. *Journal of Environmental Economics and Management* 94 (March):236–53. https://doi.org/10.1016/j.jeem.2019.01.006.

Imai, Kosuke, and In Song Kim. 2020. 'On the Use of Two-Way Fixed Effects Regression Models for Causal Inference with Panel Data'. *Political Analysis*, 1–11. https://doi.org/10.1017/pan.2020.33.

Independent Hospital Pricing Authority. 2014. 'Australian Public Hospitals Cost Report 2013-2014 Round 18'. National Hospital Cost Data Collection. https://www.ihpa.gov.au/sites/default/files/publications/nhcdc-round18.pdf.

Isen, Adam, Maya Rossin-Slater, and Reed Walker. 2017. 'Relationship between Season of Birth, Temperature Exposure, and Later Life Wellbeing'. *Proceedings of the*

National Academy of Sciences 114 (51): 13447–52. https://doi.org/10.1073/pnas.1702436114.

Jacob, Josephine, Moritz Lehne, Andrea Mischker, Normen Klinger, Claudia Zickermann, and Jochen Walker. 2017. 'Cost Effects of Preterm Birth: A Comparison of Health Care Costs Associated with Early Preterm, Late Preterm, and Full-Term Birth in the First 3 Years after Birth'. *The European Journal of Health Economics* 18 (8): 1041–46. https://doi.org/10.1007/s10198-016-0850-x.

Johnson, Jay S, Jacob M Maskal, Alan W Duttlinger, Kouassi R Kpodo, Betty R McConn, Christopher J Byrd, Brian T Richert, et al. 2020. 'In Utero Heat Stress Alters the Postnatal Innate Immune Response of Pigs'. *Journal of Animal Science* 98 (12): skaa356. https://doi.org/10.1093/jas/skaa356.

KC, Ashish, Sujeena Maharjan, Omkar Basnet, Honey Malla, Rejina Gurung, Sunil Mani Pokharel, Gyanu Kumari Ghimire, Masoud Vaezghasemi, and Julia Schröders. 2024. 'Development, Validation and Reliability of Scales and Items for Heat Wave Risk Assessment of Pregnant Women'. *International Journal of Biometeorology*, August. https://doi.org/10.1007/s00484-024-02738-x.

Kelly, Michelle M., and Patricia B. Griffith. 2020. 'Umbrella Review of School Age Health Outcomes of Preterm Birth Survivors'. *Journal of Pediatric Health Care* 34 (5): e59–76. https://doi.org/10.1016/j.pedhc.2020.05.007.

Khosla, Radhika, Anant Jani, and Rafael Perera. 2021. 'Health Risks of Extreme Heat'. *BMJ* 375 (October):n2438. https://doi.org/10.1136/bmj.n2438.

Kidd, Sean A., Jessica Gong, Alessandro Massazza, Mariya Bezgrebelna, Yali Zhang, and Shakoor Hajat. 2023. 'Climate Change and Its Implications for Developing Brains – In Utero to Youth: A Scoping Review'. *The Journal of Climate Change and Health* 13 (September):100258. https://doi.org/10.1016/j.joclim.2023.100258.

Kim, Jiyoon, Ajin Lee, and Maya Rossin-Slater. 2021. 'What to Expect When It Gets Hotter: The Impacts of Prenatal Exposure to Extreme Temperature on Maternal Health'. *American Journal of Health Economics* 7 (3): 44.

Kim, Satbyul Estella, Masahiro Hashizume, Ben Armstrong, Antonio Gasparrini, Kazutaka Oka, Yasuaki Hijioka, -Cabrera Ana M. Vicedo, and Yasushi Honda. 2023. 'Mortality Risk of Hot Nights: A Nationwide Population-Based Retrospective Study in Japan'. *Environmental Health Perspectives* 131 (5): 057005. https://doi.org/10.1289/EHP11444.

Knotz, Carlo Michael. 2018. 'A Rising Workfare State? Unemployment Benefit Conditionality in 21 OECD Countries, 1980–2012'. *Journal of International and Comparative Social Policy* 34 (2): 91–108. https://doi.org/10.1080/21699763.2018.1472136.

Kottek, Markus, Jürgen Grieser, Christoph Beck, Bruno Rudolf, and Franz Rubel. 2006. 'World Map of the Köppen-Geiger Climate Classification Updated'.

Meteorologische Zeitschrift, July, 259–63. https://doi.org/10.1127/0941-2948/2006/0130.

Kourtis, Athena P., Jennifer S. Read, and Denise J. Jamieson. 2014. 'Pregnancy and Infection'. *New England Journal of Medicine* 370 (23): 2211–18.

Kyriopoulos, Ilias, Zlatko Nikoloski, and Elias Mossialos. 2019. 'Does Economic Recession Impact Newborn Health? Evidence from Greece'. *Social Science & Medicine* 237 (September):112451. https://doi.org/10.1016/j.socscimed.2019.112451.

LaPointe, Sarah, Lauren E. Beagle, Xiaping Zheng, Vijaya Kancherla, Abby Mutic, Howard H. Chang, and Audrey J. Gaskins. 2024. 'Associations between Exposure to Extreme Ambient Heat and Neural Tube Defects in Georgia, USA: A Population-Based Case-Control Study'. *Environmental Research* 261 (November):119756. https://doi.org/10.1016/j.envres.2024.119756.

Leonard, Dympna, Danielle Aquino, Nyssa Hadgraft, Fintan Thompson, and Julia V. Marley. 2017. 'Poor Nutrition from First Foods: A Cross-Sectional Study of Complementary Feeding of Infants and Young Children in Six Remote Aboriginal Communities across Northern Australia'. *Nutrition & Dietetics* 74 (5): 436–45. https://doi.org/10.1111/1747-0080.12386.

Leung, Michael. 2023. 'Invited Perspective: Climate Change and Reproductive Health—the Perils of Oversimplification'. *Environmental Health Perspectives* 131 (12): 121307. https://doi.org/10.1289/EHP13725.

Li, Zhihui, Lincoln Chen, Mingqiang Li, and Jessica Cohen. 2018. 'Prenatal Exposure to Sand and Dust Storms and Children's Cognitive Function in China: A Quasi-Experimental Study'. *The Lancet Planetary Health* 2 (5): e214–22. https://doi.org/10.1016/S2542-5196(18)30068-8.

Liu, Buyun, Guifeng Xu, Yangbo Sun, Xiu Qiu, Kelli K. Ryckman, Yongfu Yu, Linda G. Snetselaar, and Wei Bao. 2020. 'Maternal Cigarette Smoking before and during Pregnancy and the Risk of Preterm Birth: A Dose–Response Analysis of 25 Million Mother–Infant Pairs'. *PLoS Medicine* 17 (8): e1003158. https://doi.org/10.1371/journal.pmed.1003158.

Longmore, Danielle K., Elizabeth L. M. Barr, Alyce N. Wilson, Federica Barzi, Marie Kirkwood, Alison Simmonds, I-Lynn Lee, et al. 2020. 'Associations of Gestational Diabetes and Type 2 Diabetes during Pregnancy with Breastfeeding at Hospital Discharge and up to 6 Months: The PANDORA Study'. *Diabetologia* 63 (12): 2571–81. https://doi.org/10.1007/s00125-020-05271-9.

Lovett, Nicholas, and Yuhan Xue. 2017. 'Have Electronic Benefits Cards Improved Food Access for Food Stamp Recipients?' *Journal of Economic Studies* 44 (6): 958–75. https://doi.org/10.1108/JES-10-2016-0193.

Maccini, Sharon, and Dean Yang. 2009. 'Under the Weather: Health, Schooling, and Economic Consequences of Early-Life Rainfall'. *American Economic Review* 99 (3): 1006–26. https://doi.org/10.1257/aer.99.3.1006.

MAL-ED Network Investigators, Am Acosta, Rr De Burga, Cb Chavez, Jt Flores, Mp Olortegui, Sr Pinedo, et al. 2018. 'Early Childhood Cognitive Development Is Affected by Interactions among Illness, Diet, Enteropathogens and the Home Environment: Findings from the MAL-ED Birth Cohort Study'. *BMJ Global Health* 3 (4): e000752. https://doi.org/10.1136/bmjgh-2018-000752.

Mallett, Lea H., and Ruth A. Etzel. 2018. 'Flooding: What Is the Impact on Pregnancy and Child Health?' *Disasters* 42 (3): 432–58. https://doi.org/10.1111/disa.12256.

Margolies, Amy, and John Hoddinott. 2015. 'Costing Alternative Transfer Modalities'. *Journal of Development Effectiveness* 7 (1): 1–16. https://doi.org/10.1080/19439342.2014.984745.

Marques, Andrea Horvath, Thomas G. O'Connor, Christine Roth, Ezra Susser, and Anne-Lise Bjørke-Monsen. 2013. 'The Influence of Maternal Prenatal and Early Childhood Nutrition and Maternal Prenatal Stress on Offspring Immune System Development and Neurodevelopmental Disorders'. *Frontiers in Neuroscience* 7. https://doi.org/10.3389/fnins.2013.00120.

Marston, Greg, Louise Humpage, Michelle Peterie, Philip Mendes, Shelley Bielefeld, and Zoe Staines. 2022. *Compulsory Income Management in Australia and New Zealand: More Harm than Good?* Policy Press. https://research.monash.edu/en/publications/compulsory-income-management-in-australia-and-new-zealand-more-ha.

Martin-Bassols, Nicolau, Sonja C. de New, Michael A. Shields, and David W. Johnston. 2024. 'Effect of In Utero Exposure to Air Pollution on Adulthood Hospitalizations'. *Journal of Urban Health* 101 (1): 92–108. https://doi.org/10.1007/s11524-023-00803-1.

Masarik, April S, and Rand D Conger. 2017. 'Stress and Child Development: A Review of the Family Stress Model'. *Current Opinion in Psychology*, Relationships and stress, 13 (February):85–90. https://doi.org/10.1016/j.copsyc.2016.05.008.

Masselot, Pierre, Malcolm Mistry, Jacopo Vanoli, Rochelle Schneider, Tamara Iungman, David Garcia-Leon, Juan-Carlos Ciscar, et al. 2023. 'Excess Mortality Attributed to Heat and Cold: A Health Impact Assessment Study in 854 Cities in Europe'. *The Lancet Planetary Health* 7 (4): e271–81. https://doi.org/10.1016/S2542-5196(23)00023-2.

Masten, Ann S., and Dante Cicchetti. 2010. 'Developmental Cascades'. *Development and Psychopathology* 22 (3): 491–95. https://doi.org/10.1017/S0954579410000222.

Mathew, Supriya, Deepika Mathur, Anne B. Chang, Elizabeth McDonald, Gurmeet R. Singh, Darfiana Nur, and Rolf Gerritsen. 2017. 'Examining the Effects of Ambient Temperature on Pre-Term Birth in Central Australia'. *International*

Journal of Environmental Research and Public Health 14 (2). https://doi.org/10.3390/ijerph14020147.

McDade, Thomas W, Melinda A Beck, Christopher Kuzawa, and Linda S Adair. 2001. 'Prenatal Undernutrition, Postnatal Environments, and Antibody Response to Vaccination in Adolescence'. *The American Journal of Clinical Nutrition* 74 (4): 543–48. https://doi.org/10.1093/ajcn/74.4.543.

McDade, Thomas W, and Christopher Kuzawa. 2004. 'Fetal Programming of Immune Function'. In *Fetal Nutrition and Adult Disease: Programming of Chronic Disease Through Fetal Exposure to Undernutrition*. CABI.

McGrath, John J., Adrian G. Barnett, and Darryl W. Eyles. 2005. 'The Association between Birth Weight, Season of Birth and Latitude'. *Annals of Human Biology* 32 (5): 547–59. https://doi.org/10.1080/03014460500154699.

McGregor, Glenn R., and Jennifer K. Vanos. 2018. 'Heat: A Primer for Public Health Researchers'. *Public Health* 161 (August):138–46. https://doi.org/10.1016/j.puhe.2017.11.005.

Melville, Jacqueline M., and Timothy J. M. Moss. 2013. 'The Immune Consequences of Preterm Birth'. *Frontiers in Neuroscience* 7. https://doi.org/10.3389/fnins.2013.00079.

Memmott, Paul, Nina Lansbury, Carroll Go-Sam, Daphne Nash, Andrew Martin Redmond, Samuel Barnes, Patrick (Pepy) Simpson, and Patricia Narrurlu Frank. 2022. 'Aboriginal Social Housing in Remote Australia: Crowded, Unrepaired and Raising the Risk of Infectious Diseases'. *Global Discourse* 12 (2): 255–84. https://doi.org/10.1332/204378921X16315375796362.

Miller, Jessica E, Geoffrey C Hammond, Tobias Strunk, Hannah C Moore, Helen Leonard, Kim W Carter, Zulfiqar Bhutta, Fiona Stanley, Nicholas de Klerk, and David P Burgner. 2016. 'Association of Gestational Age and Growth Measures at Birth with Infection-Related Admissions to Hospital throughout Childhood: A Population-Based, Data-Linkage Study from Western Australia'. *The Lancet Infectious Diseases* 16 (8): 952–61. https://doi.org/10.1016/S1473-3099(16)00150-X.

Moore, Kerryn A., Julie A. Simpson, Jacher Wiladphaingern, Aung Myat Min, Mupawjay Pimanpanarak, Moo Kho Paw, Jathee Raksuansak, et al. 2017. 'Influence of the Number and Timing of Malaria Episodes during Pregnancy on Prematurity and Small-for-Gestational-Age in an Area of Low Transmission'. *BMC Medicine* 15 (1): 117. https://doi.org/10.1186/s12916-017-0877-6.

Moore, Sophie E., Andrew C. Collinson, Pa Tamba N'Gom, Richard Aspinall, and Andrew M. Prentice. 2006. 'Early Immunological Development and Mortality from Infectious Disease in Later Life'. *Proceedings of the Nutrition Society* 65 (3): 311–18. https://doi.org/10.1079/PNS2006503.

Mora, Camilo, Tristan McKenzie, Isabella M. Gaw, Jacqueline M. Dean, Hannah von Hammerstein, Tabatha A. Knudson, Renee O. Setter, et al. 2022. 'Over Half of Known Human Pathogenic Diseases Can Be Aggravated by Climate Change'. *Nature Climate Change* 12 (9): 869–75. https://doi.org/10.1038/s41558-022-01426-1.

Moreira, Rafaela S., Lívia C. Magalhães, and Claudia R. L. Alves. 2014. 'Effect of Preterm Birth on Motor Development, Behavior, and School Performance of School-age Children: A Systematic Review'. *Jornal de Pediatria* 90 (2): 119–34. https://doi.org/10.1016/j.jped.2013.05.010.

Mrejen, Matias, Julian Perelman, and Danielle Carusi Machado. 2020. 'Environmental Disasters and Birth Outcomes: Impact of a Tailings Dam Breakage in Brazil'. *Social Science & Medicine* 250 (April):112868. https://doi.org/10.1016/j.socscimed.2020.112868.

Mullick, S., D. Watson-Jones, M. Beksinska, and D. Mabey. 2005. 'Sexually Transmitted Infections in Pregnancy: Prevalence, Impact on Pregnancy Outcomes, and Approach to Treatment in Developing Countries'. *Sexually Transmitted Infections* 81 (4): 294–302. https://doi.org/10.1136/sti.2002.004077.

Nairn, John R., and Robert J. B. Fawcett. 2015. 'The Excess Heat Factor: A Metric for Heatwave Intensity and Its Use in Classifying Heatwave Severity'. *International Journal of Environmental Research and Public Health* 12 (1): 227–53. https://doi.org/10.3390/ijerph120100227.

National Scientific Council on the Developing Child. 2023. 'Place Matters: The Environment We Create Shapes the Foundations of Healthy Development'. 16. Working Paper. Center on the Developing Child at Harvard University. https://ecdan.org/resource/hcdc-working-paper-16-media-kit-place-matters/.

Nawaro, Julia, Lorenzo Gianquintieri, Andrea Pagliosa, Giuseppe M. Sechi, and Enrico Gianluca Caiani. 2023. 'Heatwave Definition and Impact on Cardiovascular Health: A Systematic Review'. *Public Health Reviews* 44:1606266. https://doi.org/10.3389/phrs.2023.1606266.

NESP Earth Systems and Climate Change Hub. 2020. 'Climate Change in the Northern Territory: State of the Science and Climate Change Impacts'. Melbourne, Australia: NESP ESCC Hub.

Newman, Kira L., Kathryn Gustafson, Janet A. Englund, Amalia Magaret, Subarna Khatry, Steven C. LeClerq, James M. Tielsch, Joanne Katz, and Helen Y. Chu. 2019. 'Effect of Diarrheal Illness During Pregnancy on Adverse Birth Outcomes in Nepal'. *Open Forum Infectious Diseases* 6 (2). https://doi.org/10.1093/ofid/ofz011.

Newman, Lori, Mary Kamb, Sarah Hawkes, Gabriela Gomez, Lale Say, Armando Seuc, and Nathalie Broutet. 2013. 'Global Estimates of Syphilis in Pregnancy and Associated Adverse Outcomes: Analysis of Multinational Antenatal Surveillance

Data'. *PLOS Medicine* 10 (2): e1001396. https://doi.org/10.1371/journal.pmed.1001396.

Nielsen, Nete Munk, Anne Vinkel Hansen, Jacob Simonsen, and Anders Hviid. 2011. 'Prenatal Stress and Risk of Infectious Diseases in Offspring'. *American Journal of Epidemiology* 173 (9): 990–97. https://doi.org/10.1093/aje/kwq492.

Nunes, Marta C., Anushka R. Aqil, Saad B. Omer, and Shabir A. Madhi. 2016. 'The Effects of Influenza Vaccination during Pregnancy on Birth Outcomes: A Systematic Review and Meta-Analysis'. *American Journal of Perinatology* 33 (11): 1104–14. https://doi.org/10.1055/s-0036-1586101.

Nyadanu, Sylvester Dodzi, Jennifer Dunne, Gizachew A. Tessema, Ben Mullins, Bernard Kumi-Boateng, Michelle L. Bell, Bereket Duko, and Gavin Pereira. 2024. 'Maternal Exposure to Ambient Air Temperature and Adverse Birth Outcomes: An Umbrella Review of Systematic Reviews and Meta-Analyses'. *Science of The Total Environment* 917 (March):170236. https://doi.org/10.1016/j.scitotenv.2024.170236.

Nyadanu, Sylvester Dodzi, Gizachew Assefa Tessema, Ben Mullins, and Gavin Pereira. 2022. 'Maternal Acute Thermophysiological Stress and Stillbirth in Western Australia, 2000-2015: A Space-Time-Stratified Case-Crossover Analysis'. *The Science of the Total Environment* 836 (August):155750. https://doi.org/10.1016/j.scitotenv.2022.155750.

Oppermann, Elspeth, Matt Brearley, Lisa Law, James A. Smith, Alan Clough, and Kerstin Zander. 2017. 'Heat, Health, and Humidity in Australia's Monsoon Tropics: A Critical Review of the Problematization of "Heat" in a Changing Climate'. *WIREs Climate Change* 8 (4): e468. https://doi.org/10.1002/wcc.468.

Pacheco, Jimena, and Natascha Wagner. 2023. 'Long-Term Impacts of an Early Childhood Shock on Human Capital: Evidence from the 1999 Economic Crisis in Ecuador'. *Health Economics* 32 (11): 2460–76. https://doi.org/10.1002/hec.4742.

Park, R. Jisung. 2022. 'Hot Temperature and High-Stakes Performance'. *Journal of Human Resources* 57 (2): 400–434. https://doi.org/10.3368/jhr.57.2.0618-9535R3.

Park, R. Jisung, A. Patrick Behrer, and Joshua Goodman. 2021. 'Learning Is Inhibited by Heat Exposure, Both Internationally and within the United States'. *Nature Human Behaviour* 5 (1): 19–27. https://doi.org/10.1038/s41562-020-00959-9.

Park, R. Jisung, Joshua Goodman, Michael Hurwitz, and Jonathan Smith. 2020. 'Heat and Learning'. *American Economic Journal: Economic Policy* 12 (2): 306–39. https://doi.org/10.1257/pol.20180612.

Part, Chérie, Jean le Roux, Matthew Chersich, Shobna Sawry, Véronique Filippi, Nathalie Roos, Lee Fairlie, et al. 2022. 'Ambient Temperature during Pregnancy and Risk of Maternal Hypertensive Disorders: A Time-to-Event Study in

Johannesburg, South Africa'. *Environmental Research* 212 (September):113596. https://doi.org/10.1016/j.envres.2022.113596.

Peña, Pablo A. 2017. 'Creating Winners and Losers: Date of Birth, Relative Age in School, and Outcomes in Childhood and Adulthood'. *Economics of Education Review* 56 (February):152–76. https://doi.org/10.1016/j.econedurev.2016.12.001.

Peng, Peng, Jason Chow, Ni Yan, and Yuting Liu. 2024. 'The Developmental Mutualism of Language Skills and Behavioral Problems: The Time-Sensitive Mediating Role of Social Skills'. *Journal of Educational Psychology* 116 (6): 889–902. https://doi.org/10.1037/edu0000899.

Portet, Stéphanie. 2020. 'A Primer on Model Selection Using the Akaike Information Criterion'. *Infectious Disease Modelling* 5 (January):111–28. https://doi.org/10.1016/j.idm.2019.12.010.

Prasad, Marianne, Mirka Lumia, Maijaliisa Erkkola, Heli Tapanainen, Carina Kronberg-Kippilä, Jetta Tuokkola, Ulla Uusitalo, et al. 2010. 'Diet Composition of Pregnant Finnish Women: Changes over Time and across Seasons'. *Public Health Nutrition* 13 (6A): 939–46. https://doi.org/10.1017/S1368980010001138.

Quilty, Simon, Norman Frank Jupurrurla, Ross S Bailie, and Russell L Gruen. 2022. 'Climate, Housing, Energy and Indigenous Health: A Call to Action'. *Medical Journal of Australia* 217 (1): 9–12. https://doi.org/10.5694/mja2.51610.

Quilty, Simon, Norman Frank Jupurrurla, Aparna Lal, Veronica Matthews, Antonio Gasparrini, Pandora Hope, Matt Brearley, and Kris L. Ebi. 2023. 'The Relative Value of Sociocultural and Infrastructural Adaptations to Heat in a Very Hot Climate in Northern Australia: A Case Time Series of Heat-Associated Mortality'. *The Lancet Planetary Health* 7 (8): e684–93. https://doi.org/10.1016/S2542-5196(23)00138-9.

Randell, Heather, Clark Gray, and Kathryn Grace. 2020. 'Stunted from the Start: Early Life Weather Conditions and Child Undernutrition in Ethiopia'. *Social Science & Medicine* 261 (September):113234. https://doi.org/10.1016/j.socscimed.2020.113234.

Razaz, Neda, Sven Cnattingius, and K. S. Joseph. 2019. 'Association between Apgar Scores of 7 to 9 and Neonatal Mortality and Morbidity: Population Based Cohort Study of Term Infants in Sweden'. *BMJ* 365 (May):l1656. https://doi.org/10.1136/bmj.l1656.

Reader, Mary. 2023. 'The Infant Health Effects of Starting Universal Child Benefits in Pregnancy: Evidence from England and Wales'. *Journal of Health Economics* 89 (May):102751. https://doi.org/10.1016/j.jhealeco.2023.102751.

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

Rocha, Rudi, and Rodrigo R. Soares. 2015. 'Water Scarcity and Birth Outcomes in the Brazilian Semiarid'. *Journal of Development Economics* 112 (January):72–91. https://doi.org/10.1016/j.jdeveco.2014.10.003.

Rogers, Cassandra D. W., Mingfang Ting, Cuihua Li, Kai Kornhuber, Ethan D. Coffel, Radley M. Horton, Colin Raymond, and Deepti Singh. 2021. 'Recent Increases in Exposure to Extreme Humid-Heat Events Disproportionately Affect Populated Regions'. *Geophysical Research Letters* 48 (19): e2021GL094183. https://doi.org/10.1029/2021GL094183.

Romero, R., S. K. Dey, and S. J. Fisher. 2014. 'Preterm Labor: One Syndrome, Many Causes'. *Science* 345 (6198): 760–65. https://doi.org/10.1126/science.1251816.

Ross, Michael G., and Mina Desai. 2005. 'Gestational Programming: Population Survival Effects of Drought and Famine during Pregnancy'. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* 288 (1): R25–33. https://doi.org/10.1152/ajpregu.00418.2004.

Roth, Jonathan, Pedro H. C. Sant'Anna, Alyssa Bilinski, and John Poe. 2023. 'What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature'. *Journal of Econometrics* 235 (2): 2218–44. https://doi.org/10.1016/j.jeconom.2023.03.008.

Samuels, Louisa, Britt Nakstad, Nathalie Roos, Ana Bonell, Matthew Chersich, George Havenith, Stanley Luchters, et al. 2022. 'Physiological Mechanisms of the Impact of Heat during Pregnancy and the Clinical Implications: Review of the Evidence from an Expert Group Meeting'. *International Journal of Biometeorology* 66 (8): 1505–13. https://doi.org/10.1007/s00484-022-02301-6.

Schifano, Patrizia, Federica Asta, Payam Dadvand, Marina Davoli, Xavier Basagana, and Paola Michelozzi. 2016. 'Heat and Air Pollution Exposure as Triggers of Delivery: A Survival Analysis of Population-Based Pregnancy Cohorts in Rome and Barcelona'. *Environment International* 88 (March):153–59. https://doi.org/10.1016/j.envint.2015.12.013.

Schwandt, Hannes. 2019. 'The Lasting Legacy of Seasonal Influenza: In-Utero Exposure and Labor Market Outcomes'. *CEPR Discussion Paper*, no. 12653. https://doi.org/10.2139/ssrn.2981617.

Simeoni, Umberto, Catherine Yzydorczyk, Benazir Siddeek, and Mohamed Benahmed. 2014. 'Epigenetics and Neonatal Nutrition'. *Early Human Development*, Proceedings and Selected abstracts of the 5th International Conference on Clinical Neonatology, Torino, Italy, September 11–13, 2014, 90 (September):S23–24. https://doi.org/10.1016/S0378-3782(14)50007-2.

Skarda, Ieva, Miqdad Asaria, and Richard Cookson. 2022. 'Evaluating Childhood Policy Impacts on Lifetime Health, Wellbeing and Inequality: Lifecourse Distributional Economic Evaluation'. *Social Science & Medicine* 302 (June):114960. https://doi.org/10.1016/j.socscimed.2022.114960.

Skibiel, A. L., F. Peñagaricano, R. Amorín, B. M. Ahmed, G. E. Dahl, and J. Laporta. 2018. 'In Utero Heat Stress Alters the Offspring Epigenome'. *Scientific Reports* 8 (1): 14609. https://doi.org/10.1038/s41598-018-32975-1.

Spencer, Sarah. 2013. 'Perinatal Nutrition Programs Neuroimmune Function Long-Term: Mechanisms and Implications'. *Frontiers in Neuroscience* 7. https://www.frontiersin.org/articles/10.3389/fnins.2013.00144.

State of the Tropics. 2020. 'State of the Tropics 2020 Report'. Townsville, Australia: James Cook University. https://www.jcu.edu.au/__data/assets/pdf_file/0007/1146292/SOTT-Report-2020-Web-FINAL.pdf.

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

———. 2011b. 'The Influence of Season and Ambient Temperature on Birth Outcomes: A Review of the Epidemiological Literature'. *Environmental Research* 111 (3): 451–62. https://doi.org/10.1016/j.envres.2011.01.023.

———. 2012. 'Maternal Exposure to Ambient Temperature and the Risks of Preterm Birth and Stillbirth in Brisbane, Australia'. *American Journal of Epidemiology* 175 (2): 99–107. https://doi.org/10.1093/aje/kwr404.

Stull, Roland. 2011. 'Wet-Bulb Temperature from Relative Humidity and Air Temperature'. *Journal of Applied Meteorology and Climatology* 50 (11): 2267–69. https://doi.org/10.1175/JAMC-D-11-0143.1.

Su, Jiunn-Yih, Steven Guthridge, and Vincent He. 2023. 'Patterns of Mobility among Primary School Students in the Northern Territory'. Darwin: Menzies School of Health Research. https://www.menzies.edu.au/content/Document/Patterns%20of%20mobility%20among%20primary%20school%20students%20in%20the%20Northern%20Territory_July%202023.pdf.

Sun, Shengzhi, Keith R. Spangler, Kate R. Weinberger, Jeff D. Yanosky, Joseph M. Braun, and Gregory A. Wellenius. 2019. 'Ambient Temperature and Markers of Fetal Growth: A Retrospective Observational Study of 29 Million U.S. Singleton Births'. *Environmental Health Perspectives* 127 (6). https://doi.org/10.1289/EHP4648.

Tansey, Etain A., and Christopher D. Johnson. 2015. 'Recent Advances in Thermoregulation'. *Advances in Physiology Education* 39 (3): 139–48. https://doi.org/10.1152/advan.00126.2014.

Thavarajah, Harshithaa, Christopher Flatley, and Sailesh Kumar. 2018. 'The Relationship between the Five Minute Apgar Score, Mode of Birth and Neonatal Outcomes'.

The Journal of Maternal-Fetal & Neonatal Medicine 31 (10): 1335–41. https://doi.org/10.1080/14767058.2017.1315666.

The Preterm Birth Committee. 2024. 'Preterm Birth: Reducing Risks and Improving Lives, Report of Session 2024-25'. HL Paper 30. London: House of Lords. https://publications.parliament.uk/pa/ld5901/ldselect/ldpreterm/30/30.pdf.

Thiede, Brian C., and Clark Gray. 2020. 'Climate Exposures and Child Undernutrition: Evidence from Indonesia'. *Social Science & Medicine* 265 (November):113298. https://doi.org/10.1016/j.socscimed.2020.113298.

Tonkin, Emma, Dani Kennedy, Sarah Hanieh, Beverley-Ann Biggs, Therese Kearns, Veronica Gondarra, Roslyn Dhurrkay, and Julie Brimblecombe. 2020. 'Dietary Intake of Aboriginal Australian Children Aged 6–36 Months in a Remote Community: A Cross-Sectional Study'. *Nutrition Journal* 19 (1): 34. https://doi.org/10.1186/s12937-020-00550-y.

Torche, Florencia. 2019. 'Early-Life Circumstances and Their Effects Over the Life Course'. *Population Research and Policy Review* 38 (6): 771–82. https://doi.org/10.1007/s11113-019-09555-w.

Torche, Florencia, and Alejandro Corvalan. 2010. 'Seasonality of Birth Weight in Chile: Environmental and Socioeconomic Factors'. *Annals of Epidemiology* 20 (11): 818–26. https://doi.org/10.1016/j.annepidem.2010.08.005.

Torjesen, Ingrid. 2023. 'Number of UK Children Living in Destitution Almost Triples in Five Years'. *BMJ* 383 (October):p2478. https://doi.org/10.1136/bmj.p2478.

United Nations Environment Management Group. 2011. 'Global Drylands: A UN System-Wide Response'. United Nations. https://www.unep-wcmc.org/system/dataset_file_fields/files/000/000/091/original/Global-Drylands-FINAL-LR.pdf?1398440625.

US Census Bureau. 2024. 'Poverty in the United States: 2023'. P60-283. https://www.census.gov/library/publications/2024/demo/p60-283.html.

Vecellio, Daniel J., S. Tony Wolf, Rachel M. Cottle, and W. Larry Kenney. 2022. 'Evaluating the 35°C Wet-Bulb Temperature Adaptability Threshold for Young, Healthy Subjects (PSU HEAT Project)'. *Journal of Applied Physiology* 132 (2): 340–45. https://doi.org/10.1152/japplphysiol.00738.2021.

Veenendaal, M. V. E., R. C. Painter, S. R. de Rooij, P. M. M. Bossuyt, J. a. M. van der Post, P. D. Gluckman, M. A. Hanson, and T. J. Roseboom. 2013. 'Transgenerational Effects of Prenatal Exposure to the 1944-45 Dutch Famine'. *BJOG: An International Journal of Obstetrics and Gynaecology* 120 (5): 548–53. https://doi.org/10.1111/1471-0528.12136.

Venegas Marin, Sergio, Lara Schwarz, and Shwetlena Sabarwal. 2024. 'Impacts of Extreme Weather Events on Education Outcomes: A Review of Evidence'. *The*

World Bank Research Observer 39 (2): 177–226. https://doi.org/10.1093/wbro/lkae001.

Walson, Judd L., and James A. Berkley. 2018. 'The Impact of Malnutrition on Childhood Infections'. *Current Opinion in Infectious Diseases* 31 (3): 231–36. https://doi.org/10.1097/QCO.0000000000000448.

Watson, P. E., and B. W. McDonald. 2007. 'Seasonal Variation of Nutrient Intake in Pregnancy: Effects on Infant Measures and Possible Influence on Diseases Related to Season of Birth'. *European Journal of Clinical Nutrition* 61 (11): 1271–80. https://doi.org/10.1038/sj.ejcn.1602644.

Wellinghausen, Nele. 2001. 'Immunobiology of Gestational Zinc Deficiency'. *British Journal of Nutrition* 85 (S2): S81–86. https://doi.org/10.1049/BJN2000298.

Wiedmeier, Julia E, Lisa A Joss-Moore, Robert H Lane, and Josef Neu. 2011. 'Early Postnatal Nutrition and Programming of the Preterm Neonate'. *Nutrition Reviews* 69 (2): 76–82. https://doi.org/10.1111/j.1753-4887.2010.00370.x.

Wild, R, and P Anderson. 2007. 'Ampe Akelyernemane Meke Mekarle "Little Children Are Sacred"'. Northern Territory Board of Inquiry into the Protection of Aboriginal Children from Sexual Abuse. Darwin: Northern Territory Government.

Wilde, Joshua, Bénédicte H. Apouey, and Toni Jung. 2017. 'The Effect of Ambient Temperature Shocks during Conception and Early Pregnancy on Later Life Outcomes'. *European Economic Review* 97 (August):87–107. https://doi.org/10.1016/j.euroecorev.2017.05.003.

Williams, Corinne J, and Ann M Jacobs. 2009. 'The Impact of Otitis Media on Cognitive and Educational Outcomes'. *Medical Journal of Australia* 191 (S9): S69–72. https://doi.org/10.5694/j.1326-5377.2009.tb02931.x.

Wilson, Tom, Peter McDonald, and Jeromey Temple. 2020. 'The Geographical Patterns of Birth Seasonality in Australia'. *Demographic Research* 43 (40): 1185–98. https://doi.org/10.4054/DemRes.2020.43.40.

Wolf, Jennyfer, and Ben Armstrong. 2012. 'The Association of Season and Temperature with Adverse Pregnancy Outcome in Two German States, a Time-Series Analysis'. *PLOS ONE* 7 (7): e40228. https://doi.org/10.1371/journal.pone.0040228.

WONCA International Classification Committee. 2018. 'International Classification of Primary Care - 2nd Edition'. World Organization of Family Doctors. https://www.ehelse.no/kodeverk/icpc-2e--english-version.

Wondmagegn, Berhanu Y., Jianjun Xiang, Susan Williams, Dino Pisaniello, and Peng Bi. 2019. 'What Do We Know about the Healthcare Costs of Extreme Heat Exposure? A Comprehensive Literature Review'. *Science of The Total Environment* 657 (March):608–18. https://doi.org/10.1016/j.scitotenv.2018.11.479.

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

Yu, Peter, Marcia E Duncan, and Bill Gray. 2008. ‘Northern Territory Emergency Response: Report of the Review Board’. Canberra: Commonwealth of Australia. https://apo.org.au/node/551.

Zhang Xiaoxin, Wang Wenjing, Zhu Zhenghong, Wu Kaipu, Liang Liang, Li Yanqiu, Ma Wenjun, Huang Cunrui, and Wang Qiong. 2022. ‘Mediating role of sleep in association between maternal heat exposure during early pregnancy and risk of preterm birth’. *Journal of Environmental and Occupational Medicine* 39 (3): 281–88. https://doi.org/10.11836/JEOM21447.

Zhang, Yali, Shakoor Hajat, Liang Zhao, Huiqi Chen, Liangliang Cheng, Meng Ren, Kuiying Gu, John S. Ji, Wannian Liang, and Cunrui Huang. 2022. ‘The Burden of Heatwave-Related Preterm Births and Associated Human Capital Losses in China’. *Nature Communications* 13 (1): 7565. https://doi.org/10.1038/s41467-022-35008-8.

Zhang, Yunquan, Chuanhua Yu, and Lu Wang. 2017. ‘Temperature Exposure during Pregnancy and Birth Outcomes: An Updated Systematic Review of Epidemiological Evidence’. *Environmental Pollution* 225 (June):700–712. https://doi.org/10.1016/j.envpol.2017.02.066.