



**HAL**  
open science

# Critical Periods in Cognitive and Socioemotional Development: Evidence from Weather Shocks in Indonesia

Duncan Webb

► **To cite this version:**

Duncan Webb. Critical Periods in Cognitive and Socioemotional Development: Evidence from Weather Shocks in Indonesia. Economics and Finance. 2019. dumas-02407565

**HAL Id: dumas-02407565**

**<https://dumas.ccsd.cnrs.fr/dumas-02407565>**

Submitted on 12 Dec 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**MASTER THESIS N° 2019 – 04**

**Critical Periods in Cognitive and  
Socioemotional Development  
Evidence from Weather Shocks in Indonesia**

**Duncan Webb**

**JEL Codes: I15, I24, I31, I3, O1, O12, O15, O53, Q54**

**Keywords:**



# Critical Periods in Cognitive and Socioemotional Development

Evidence from Weather Shocks in Indonesia

**Duncan Webb**

Supervisor: **Karen Macours**

Referee: **Luc Behaghel**

Master Thesis

Public Policy & Development



**PARIS SCHOOL OF ECONOMICS**  
ÉCOLE D'ÉCONOMIE DE PARIS

Paris School of Economics

June 1, 2019

## Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Conceptual Framework</b>	<b>6</b>
<b>3</b>	<b>Data and Measurement System</b>	<b>8</b>
3.1	IFIS Data	8
3.2	District, Weather, and Crop Data	9
3.3	Measurement of skills	9
3.4	Samples and summary statistics	11
<b>4</b>	<b>Empirical Strategy</b>	<b>13</b>
<b>5</b>	<b>Main Results</b>	<b>15</b>
5.1	Cognitive Skills	15
5.2	Socioemotional Outcomes and Mental Health	22
<b>6</b>	<b>Mechanisms</b>	<b>29</b>
6.1	Household expenditure	29
6.2	Anthropometrics	30
6.3	Mediation Analysis	34
6.3.1	Methodology	36
6.3.2	Results	39
<b>7</b>	<b>Discussion</b>	<b>43</b>
	<b>Appendices</b>	<b>45</b>
<b>A</b>	<b>Data Appendix</b>	<b>45</b>
A.1	IFIS Data	45
A.2	District, Weather, and Crop Data	46
<b>B</b>	<b>Exploratory Factor Analysis</b>	<b>51</b>
B.1	Correcting for acquiescence bias (socioemotional only)	51
B.2	Age Standardisation	52
B.3	Determining the number of factors	52
B.4	Estimating the factor analysis model	53
B.5	Generating factor scores	57
<b>C</b>	<b>Multiple Testing Adjustment</b>	<b>62</b>
<b>D</b>	<b>Sample Selection</b>	<b>63</b>
<b>E</b>	<b>Other Tables and Figures</b>	<b>66</b>

### **Abstract**

A large literature now points towards the importance of early life circumstance in determining long-run human capital and wellbeing outcomes. This literature often justifies a focus on the very early years by citing the first 1000 days of life as a ‘critical period’ for child development, but this notion has rarely been directly tested, and it is still unclear whether it applies to socioemotional as well as cognitive development. Using an empirical setting in which children are potentially subject to shocks in every year of their childhood, I estimate the impact of early life weather shocks on adult cognitive and socioemotional outcomes for individuals born in rural Indonesia between 1988 and 2000. I show evidence suggesting there is a strong critical period for these shocks at age 2 for cognitive development, but no similar critical period for socioemotional development. I further demonstrate that the impacts of the shocks are likely to be taking place through household agricultural income and nutrition channels, and may be mitigated by compensatory parental investments.

**JEL Codes:** *I15, I24, I31, I3, O1, O12, O15, O53, Q54*

### **Acknowledgements**

I’d like to thank my supervisor Karen Macours for all her fantastic advice and support. I’d also like to thank Liam Wren-Lewis, Denis Cogneau, Sylvie Lambert, and Jérémie Gignoux for helpful comments and guidance. All errors are my own.

## 1 Introduction

There are an estimated 219 million children under the age of 5 worldwide who do not reach their developmental potential due to risk factors associated with low economic status, such as disease, nutrition, and inadequate parental investment (Grantham-McGregor et al., 2007). Furthermore, the link between early life socioeconomic circumstance and psychological development appears to be not only associative but causal. Such a causal link would constitute a powerful reinforcing relationship between economic wellbeing and human capital formation, thus providing an important explanation of the intergenerational persistence of inequality (Cunha and Heckman, 2009), both within and between countries. Moreover, understanding such a link could unlock powerful policy tools that could be used to improve the plight of those in poverty and reduce inequality worldwide.

The literature has now amassed a wide range of evidence pointing towards the importance of early life circumstance on adult health and human capital. Research in public health and economics has documented the exceptionally robust association between early life indicators of nutrition and growth and a range of adult outcomes including health, height, education, economic productivity, and cognitive ability (Crookston et al., 2011; Pollitt et al., 2006; Kuklina et al., 2006; Adair et al., 2013; Glewwe and King, 2001; Sudfeld et al., 2015; Leroy and Frongillo, 2019). Further work originating in development economics has made use of plausibly exogenous shocks to show that even relatively mild shocks in the first few years of life may have long run impacts on a similar span of adult outcomes (see extensive survey in Almond et al., 2018).

Many of the human capital outcomes that economists care most about, such as educational and labour market success, are very well predicted by a low-dimensional set of psychological traits, typically divided into *cognitive* and *socioemotional* traits (Heckman et al., 2006). Adult cognitive ability has been shown to be strongly affected by a range of shocks or interventions in early life, including natural disasters (Baez et al., 2010), disease exposure (Venkataramani, 2012), and cash transfer programmes (Macours et al., 2012). And tests of the fetal origins hypothesis (Barker, 1995) have shown that shocks in utero can have an equally strong effect (Black et al., 2018; Nilsson, 2017). Evidence on socioemotional traits is sparser, but there are now results showing that early life and in utero circumstance may have an important effect on mental health (Adhvaryu and Fenske, 2016; Pasha et al., 2018; Singhal, 2018; Persson and Rossin-Slater, 2018), personality (Krutikova and Lilleør, 2015; Evans and English, 2002), and trust in adulthood (BenYishay, 2013). This complements claims made about early childhood interventions that suggest that large long-run benefits of such programmes may in fact be mostly explained by positive impacts upon socioemotional and behavioural traits rather than cognitive ability (Heckman et al., 2013).

There are, however, a number of unanswered questions in this area. In particular, we know relatively little about how sensitivity to shocks and interventions varies over different ages of childhood. The literature frequently makes reference to the first 1000 days of life as a “critical period” of development, that is, a period during which experiences are crucial for normal neurological development (Knudsen, 2004; Cunha and Heckman, 2009). The evidence cited above does seem to suggest that effects on adult cognitive ability are particularly large during this time window. On the other hand, the critical period hypothesis excludes the possibility of large effects later in childhood and has rarely been tested for directly in the economics literature. And Thompson and Nelson (2001) claim that the emphasis on this critical development period risks ignoring important changes that occur in the later childhood years, as well as the continuing plasticity of the adult brain. Perhaps, they suggest, the first 1000 days should be seen as a “sensitive” period rather than a critical one.

The sparse evidence we have on this topic suggests that the first 3 years are more likely to be critical for cognitive rather than socioemotional development. In particular, Barham et al. (2013) examine the effect of conditional cash transfers received between in utero and age 2 as compared to ages 2 to 5 in Nicaragua on cognitive and health outcomes. They show that boys exposed to the program between in utero and age 2 have cognitive scores that are approximately 0.15 standard deviations

higher than boys exposed to the program between 2 and 5. This appears to confirm the critical period hypothesis in this context for cognitive outcomes. And [Heckman et al. \(2013\)](#) claim that the elasticity of substitution of cognitive skills with respect to investment declines sharply with age, whereas socioemotional skills remain malleable for longer (the elasticity of substitution may even increase with age). Such a picture is consistent with biological evidence showing that the pre-frontal cortex, which governs “higher-order” cognitive function and socio-emotional behaviour, develops later than the areas of the brain that govern visual, spatial, and language capabilities (see [Grantham-McGregor et al., 2007](#) or [Sapolsky, 2017](#) for more detail). But the evidence provided in [Heckman et al. \(2013\)](#) is based on a calibrated model of the human capital production function (see [Section 2](#)) that relies on a number of parametric assumptions, rather than exogenous variation resulting from a shock or an intervention. If any of these assumptions are false, this would give reason to doubt their exact empirical results.

In this paper, therefore, I take advantage of an empirical setting in which individuals are subject to repeated exogenous shocks throughout their entire childhood and use it to carry out a fine-grained test for the existence of critical and sensitive periods in both cognitive and socio-emotional development. In particular, I examine the reduced-form effects of exogenous weather shocks at every age from between in utero and age 15 on the adult psychological outcomes of individuals in rural Indonesia. I match historical weather data to survey data on individuals from the Indonesia Family Life Survey (IFLS) born in rural areas between 1988 and 2000 and examine the effects of childhood weather shocks on adult cognitive ability and socioemotional measures. These weather shocks are likely to have an effect on children of rural households via an impact on agricultural income that in turn affects household economic welfare and possibly childhood nutrition and other investments.

A key advantage of my empirical strategy for testing for critical periods is the use of weather shocks as an exogenous source of variation. Each child is subject to multiple weather shocks over their life, and moreover these shocks are not correlated across periods. Together this means that each shock provides independent variation that can be used to compare effects of similar shocks at different ages of childhood, thus testing for critical periods. This can be contrasted to much of the previous literature, which is often only able to look at single shocks or interventions (e.g. [Heckman et al., 2013](#)), or at shocks that are correlated over time (e.g. [Adhvaryu and Fenske, 2016](#)), resulting in limited ability to test for the timing of effects.

The IFLS data also provides unique opportunities for testing the notion of critical periods. It is a large, high-quality 5-wave panel survey that includes a total of over 30,000 individuals, is representative of 83% of the Indonesian population, and contains detailed outcome data on a wide range of socioeconomic and human capital measures. The richness of the data offers at least 3 key advantages over other datasets when testing for critical periods. First, all adult respondents are asked for a detailed migration history. This, along with very low attrition rates on migrants, provides an unusual opportunity to reconstruct the location of individuals for all years of their life since birth in order to examine the effects of weather shocks up until 15 years of age. Second, IFLS provides an uncommonly broad range of outcome data that includes measures of mental health, personality, and wellbeing, along with state-of-the-art measures of cognitive ability. This allows me to test for critical periods in all of these domains. And third, the panel nature of the survey means that there are contemporaneous data on many individuals from during their childhood years. I am able to use these data to examine the mechanisms through which weather shocks may be operating, including via household income, nutrition, and other measures of parental investments.

Indonesia is an ideal setting for testing the timing of the effects of weather shocks. First, it is a geographically large country (east to west, it spans 5,100 kilometres) with diverse patterns of weather across provinces and districts, meaning that there is likely to be significant variation within a given year and across years of the weather patterns experienced by individuals ([Frederick et al., 2011](#); [Maccini and Yang, 2009](#)). Second, Indonesia is a lower-middle-income country with a large rural population and large agricultural sector ([World Bank, 2018](#)). This population are likely to experience

weather shocks as agricultural income shocks that affect children through changes in nutrition and other investment channels,<sup>1</sup> allowing me to test for critical periods at various stages of childhood.

My results on cognitive ability are consistent with the notion of a critical period for income- or nutrition-driven shocks during the first 1000 days. In particular, I find a strong positive effect of weather shocks at age 2 on adult cognitive outcomes, but no statistically significant effect at any other age or in utero. These results are largely driven by effects on boys, in line with previous literature showing that boys may be more vulnerable to disruptions to cognitive development in utero and in very early childhood (Eriksson et al., 2010). My results on socioemotional outcomes provide no evidence for the existence of a critical period for income- or nutrition-driven shocks in this domain. For mental health and personality, I find no statistically significant effect of weather shocks on adult outcomes. However for wellbeing, I find some indication that weather shocks during early life (in utero to age 15) have some effect on adult outcomes; this effect appears to be concentrated on girls and is spread across multiple ages and in utero. I interpret this as evidence against the notion of a critical period for income and nutrition shocks on this aspect of socioemotional development.

Theoretical and empirical work in this field has pointed towards the importance of responses of parental investments to shocks and interventions (Heckman and Cunha, 2007; Fryer et al., 2015), which can compensate for or reinforce direct effects. The main results I report are “reduced-form” estimates that include the effects of resulting investment responses, which are informative in their own right. However, in order to further understand the mechanisms behind the effects of weather shocks in my sample, I show a set of supplementary results. First, I indicate that my results are consistent with the effects of a contemporaneous agricultural income shock that affects household consumption and subsequently childhood anthropometrics (a reasonable proxy for health). Then I run a mediation analysis exercise in which I show possible evidence of compensatory investment responses which may dampen the direct impact of weather shocks on adult cognitive ability. However, a number of empirical challenges means that we should interpret this latter result with caution.

The results found in this paper have potentially important policy implications. Specifically, I have been able to isolate the impact of weather shocks on cognitive development to the year after the second birthday. By contrast, no such period of sensitivity appears to exist for socio-emotional development. To the extent that the effects of such weather shocks occur through economic and health-related channels that can be manipulated by policy, this implies that early-childhood interventions focusing on cognitive development should ideally be targeted the year after the second birthday, which is possibly a “critical period” for such development. On the other hand, the timing of income and nutritional programmes aimed at improving socioemotional outcomes may be less crucial to strong effects. I also show that cognitive development is more sensitive to these shocks at age 2 for boys, whereas socioemotional development is malleable at various ages for girls. Such results indicate that returns to income- or nutrition-focused early childhood interventions may lead to higher cognitive returns for boys and higher wellbeing returns for girls. The existence of differential impacts by gender could usefully inform the optimal targeting of such interventions.

The rest of the paper is organised as follows. Section 2 describes a theoretical grounding for the results based on Heckman’s human capital production function framework, including definitions for critical and sensitive periods. Section 3 describes the data sources used and the latent factor model used to measure psychological outcomes. Section 4 outlines the empirical strategy, while Section 5 reports the results on cognitive and socioemotional outcomes. Section 6 reports the results on potential mechanisms, including effects on household expenditure and child anthropometrics, along with the mediation analysis described above. Section 7 concludes with a discussion of the results.

---

<sup>1</sup>For example, Levine and Yang (2014) find that rainfall shocks are positively associated with deviations in district-level rice output from district-level mean, and rice is the main harvested crop across much of Indonesia (see Appendix Figure 13). They interpret this as justification for interpreting higher rainfall as a positive contemporaneous shock to local economic conditions in Indonesia.



## 2 Conceptual Framework

I describe here the theoretical framework on which my empirical results are founded, which allows me to give a precise definition of the concepts of “critical” and “sensitive” periods. The framework is based on the human capital production function model seen in [Heckman and Cunha \(2007\)](#) and [Cunha et al. \(2010\)](#) in which the timing of investment into human capital during childhood matters.<sup>2</sup> Denote  $\theta_t$  as a vector of a child’s stock of skills (including, for example, cognitive and socioemotional skills). In this model, skills evolve over the course of childhood in a way that depends on parental characteristics and human capital ( $h^P$ ), the child’s existing stock of skills, and parental investments in each period ( $I_t$ ). Given my empirical setting, I also add a variable  $\mu_t$  which here denotes exogenous weather shocks to investment in period  $t$ .<sup>3</sup> We can therefore think of  $I_t$  as denoting the component of investment resulting from a variety of exogenous and endogenous factors, whereas  $\mu_t$  captures the specific exogenous component of investment that is driven by weather shocks.<sup>4</sup> These shocks are not correlated over periods.<sup>5</sup>

The evolution of skills over time can thus be written as:

$$\theta_{t+1} = f_t(h^P, \theta_t, (I_t + \mu_t)) \quad (1)$$

Note that all characters in boldface denote vectors. I assume for simplicity that investment and shocks are unidimensional scalars.

In my empirical analysis, I allow for 17 distinct time periods to affect adult skills, ranging from 2 years before birth up to age 15.<sup>6</sup> I therefore assume that adult skills settle by age 16, and will not be affected by any changes thereafter. Whereas equation 1 is written in recursive form, we can substitute in each equation for  $\theta_{15}, \theta_{14}, \theta_{13}$ , etc. to get an expression for the adult stock of skills as a function of initial endowments and all past investments and shocks:

$$h \equiv \theta_{16} = m\left(h^P, \theta_1, (I_{-2} + \mu_{-2}), (I_{-1} + \mu_{-1}), \dots, (I_{15} + \mu_{15})\right)$$

Following [Heckman and Cunha \(2007\)](#) I specify a flexible CES functional form for the investment technology, denoted  $g(\cdot)$ , which allows for elasticity of substitution between investment and shocks in each period  $\sigma = 1/(1 - \phi)$  to vary:

$$h = m\left(h^P, \theta_1, g(\mathbf{I}, \boldsymbol{\mu})\right)$$

where  $g(\mathbf{I}, \boldsymbol{\mu}) = \left[ \sum_{t=-2}^{T=15} \gamma_t (I_t + \mu_t)^\phi \right]^{1/\phi}$

Here, the CES shares  $\gamma_t \in [0, 1]$  can be thought of as “skill-multipliers” that describe the productivity of investment in period  $t$  via its direct effect and via the increased productivity of future investments. And  $\phi \in (-\infty, 1]$  describes the degree of complementarity or substitutability between early and late investments in producing skills.

<sup>2</sup>Full details of the model, including exact timings, examples of the parental investment maximisation programme, and other detailed definitions, can be found in the Heckman papers.

<sup>3</sup>This is similar to the adaptation seen in [Currie and Almond \(2011\)](#); [Almond et al. \(2018\)](#).

<sup>4</sup>For example, a positive weather shock may exogenously improve household income and lead to better nutritional investment in a child in period  $t$ .

<sup>5</sup>As I describe in Section 4, this assumption is born out in the data: weather shocks have a very low serial correlation across time periods.

<sup>6</sup>The lower bound is based on evidence cited in the Introduction that suggests that psychological traits are affected by shocks in utero. The upper bound is based loosely on the discussion in [Cunha and Heckman \(2009, p. 331\)](#), which suggests that socioemotional skills in particular are “malleable until later ages”.

This model implies that for a specific skill  $k$  (e.g. cognitive ability), adult skills  $h_k$  are:

$$h_k = m_k \left( \mathbf{h}^P, \boldsymbol{\theta}_1, g(\mathbf{I}, \boldsymbol{\mu}) \right)$$

I am now in a position to define precisely the notion of critical and sensitive periods. Because I am focusing on the effects of weather shocks in this paper, I define the notions relative to the exogenous shock  $\mu_t$  instead of the endogenous investment  $I_t$ . Compared to [Heckman and Cunha \(2007\)](#), I use more finely-grained time periods in my model set up, and therefore adapt their definition by allowing for the time intervals that possibly constitute critical and sensitive periods to last more than one model-period  $t$ .<sup>7</sup> In particular, define  $C$  as a time interval lasting from  $t_{min}$  to  $t_{max}$ :

$$C = [t_{min}, t_{max}]$$

Then  $C$  is a “critical period” for skill  $k$  if and only if:

$$\begin{aligned} \frac{\partial m_k(\cdot)}{\partial \mu_t} &> 0 \quad \forall t \in C \\ \frac{\partial m_k(\cdot)}{\partial \mu_t} &= 0 \quad \forall t \notin C \end{aligned}$$

Intuitively,  $C$  is a critical period when *only* shocks within the time interval have an effect on adult skills  $h_k$ , while shocks outside of the time interval have no effect on adult skills  $h_k$ .

Defined analogously,  $C$  is a “sensitive period” for skill  $k$  if and only if:

$$\frac{\partial m_k(\cdot)}{\partial \mu_t} > \frac{\partial m_k(\cdot)}{\partial \mu_s} \quad \forall t, s : t \in C \text{ and } s \notin C$$

In other words,  $C$  is a sensitive period when shocks have a bigger effect on  $h_k$  in all the periods within  $C$  than in all the periods outside of  $C$ .

Obtaining empirical causal estimates of  $\partial m_k(\cdot)/\partial \mu_t$  (what I will call the “direct effect”) for all the different time periods may be challenging. First, it may be hard to isolate exogenous shocks for every period. The weather shocks used in this paper are the solution to such a challenge. Second, parents may respond to early shocks by endogenously changing investment in later periods. So rather than estimating the direct effect, we may only directly observed a “reduced form” effect ( $dm_k(\cdot)/d\mu_t$ ) that includes these investment responses. Such an effect is nevertheless interesting in its own right: it is of policy relevance to know the effects of early childhood shocks while incorporating compensatory or reinforcing parental responses.

To reflect this, investment in  $t + 1$  can be written as a function of all previous shocks and investments.<sup>8</sup>

$$I_{t+1} = z_{t+1}(I_t, I_{t-1}, \dots, I_{-2}, \mu_t, \mu_{t-1}, \dots, \mu_{-2})$$

This implies that in general the full reduced form estimates of the impact of a shock  $\mu_t$  will not be equal to the partial equilibrium effects, that is:

$$\frac{dm_k(\cdot)}{d\mu_t} \neq \frac{\partial m_k(\cdot)}{\partial \mu_t}$$

To illustrate this, consider a simplified model which has only two childhood periods  $t = \{1, 2\}$ , and where  $I_1$  has already been chosen by parents to be  $\bar{I}_1$ . Then we can rewrite adult skills  $h_k$  as:<sup>9</sup>

$$h_k = m_k \left( \mathbf{h}^P, \boldsymbol{\theta}_1, g(\mu_1, \mu_2, \bar{I}_1, I_2(\mu_1)) \right)$$

<sup>7</sup>For example, in this model, ages 0 to 5 could constitute one “critical period” even though this covers 6 periods in the model.

<sup>8</sup>The functional form of  $z_{t+1}(\cdot)$  will depend on the exact optimisation programme faced by the parents. See the different suggestions in [Heckman and Cunha \(2007\)](#) or [Currie and Almond \(2011\)](#) and [Almond et al. \(2018\)](#).

<sup>9</sup>I rewrite  $I_2(\bar{I}_1, \mu_1)$  as  $I_2(\mu_1)$  since  $I_1$  is fixed.

If we assume for simplicity that  $m(\cdot)$  is additively separable in  $h^P$ ,  $\theta_1$ , and  $g(\cdot)$ , then the reduced-form effect of a change in  $\mu_1$  will be:

$$\begin{aligned}\frac{dm_k}{d\mu_1} &= \frac{dm_k}{dg} \times \frac{dg}{d\mu_1} \\ &= \frac{dm_k}{dg} \left[ \frac{\partial g}{\partial \mu_1} + \frac{\partial g}{\partial I_2} \cdot \frac{dI_2}{d\mu_1} \right]\end{aligned}$$

whereas the direct effect will be:

$$\frac{\partial m_k}{\partial \mu_1} = \frac{dm_k}{dg} \frac{\partial g}{\partial \mu_1}$$

So the total reduced-form effect can be written as the direct effect plus a term that is some positive function<sup>10</sup> of the investment change of  $I_2$  in response to the shock  $\mu_1$ :

$$\underbrace{\frac{dm_k}{d\mu_1}}_{\text{reduced form}} = \underbrace{\frac{\partial m_k}{\partial \mu_1}}_{\text{"direct" effect}} + \underbrace{\frac{dm_k}{dg} \cdot \frac{\partial g}{\partial I_2} \cdot I_2'(\mu_1)}_{\text{investment response effect}}$$

Thus, depending on whether the investment response to  $\mu_1$  is *compensating* ( $I_2'(\mu_1) < 0$ ) or *reinforcing* ( $I_2'(\mu_1) > 0$ ), the reduced form will understate or overstate the "direct" effect of the shock. The main results found in this paper identify the reduced form impacts of weather shocks in each period. Indeed, the coefficients of interest  $\beta_j$  in my main specification (see Section 4) can be thought of as identifying  $dm_k(\cdot)/d\mu_j$  for weather shocks at different ages. As noted, these coefficients are of direct policy relevance themselves. But it is also important to understand the mechanisms behind such reduced-form effects, and whether direct effects may be driven or dampened by later investment. Therefore in Section 6 I attempt to understand to what degree my results might be explained by different channels and investment responses.

### 3 Data and Measurement System

#### 3.1 IFLS Data

My main source of outcome variable data is the Indonesia Family Life Survey (IFLS) (RAND, 1994, 1997, 2000, 2008, 2015). This is a comprehensive longitudinal survey with 5 waves between 1993 and 2015 that contains detailed information on over 30,000 individuals. These individuals come from 13 of the 27 provinces in the country and are representative of approximately 83% of the Indonesian population. A wide range of questions are asked to each respondent, covering multiple areas of socioeconomic and human capital status. Particularly relevant to this paper, later waves include multiple questions on cognitive ability and other socio-emotional dimensions including mental health, subjective wellbeing, personality, positive and negative affect, and risk and time preferences. Cognitive questions include versions of Raven's progressive matrices (John and Raven, 2003) and number series tests, some of the most widely used measures in the literature. Subjective wellbeing and emotional "affect" are measured using cognitive evaluation questions such as the Cantrill Ladder, along with questions relating to positive and negative emotions the respondent felt in the last day. Personality questions are based on the "Big 5" personality traits (Rammstedt and John, 2007), and the mental health questions are based on the CES-D depression scale (Radloff, 1977). Both of these are extensively used in the psychology and economics literature. More details on the individual measures are found in Appendix Section A.1. In order to combine measures of traits into low-dimensional factors, I use a latent factor model in the style of (Heckman et al., 2013) to generate factor scores that are then used as the dependent variables in my main specification. This process is described further in Section 3.3 and in Appendix Section B.

<sup>10</sup>Skills are naturally assumed to be a positive function of investment, so  $dm_k/dg > 0$  and  $\partial g/\partial I_2 > 0$ .

The IFLS data include the measurements of height and weight that I use for results on anthropometrics, and measures of household consumption that I use to construct an overall measure of household expenditure. It further includes the multiple other demographic and socioeconomic variables that I use as controls in my models, such as gender, religion, and education levels. I also make use of detailed information on pregnancy history, child investment, and other behaviours of parents and children during individuals' early life to create the variables used in the mediation analysis.

The survey contains detailed information on an individual's location at birth and, for individuals older than 15, a full migration history tracking an individual's location at age 12 and every migration thereafter. I am able to combine this retrospective data with the contemporaneous data on location of respondents at each of the five IFLS waves (in 1993, 1997, 2000, 2007, and 2015) to construct a year-by-year history of each individual's location. I also use information on whether the individual was born in a rural household to restrict to only the rural population, under the assumption that these are the households for whom weather shocks are likely to have a significant effect on household welfare through changes in agricultural income.

More detail on the measures used in the IFLS data can be found in Section A.1 in the Appendix.

### 3.2 District, Weather, and Crop Data

The primary weather variable I use in the paper is the Standardised Precipitation Evapotranspiration Index (SPEI), a drought index developed by [Vicente-Serrano et al. \(2010\)](#). Although most of the economic literature on weather shocks focuses solely on precipitation or temperature, this neglects other features of the climate that can affect the growing cycle of a plant. In particular, the impact of rainfall on crop cycles depends on *potential evapotranspiration* (the ability of soil to retain water), which is in turn affected by multiple features of the environment, including temperature, latitude, windspeed, and the number of sunlight hours. SPEI is a rich index that combines all of these features. It has been shown to be a better predictor of crop yields and other ecological and hydrological variables than other climate indexes in a variety of climatic contexts [Beguería et al. \(2014\)](#).

I use monthly gridded data for SPEI from the Global SPEI database that is available in grids of 0.5 latitude by 0.5 longitude (approximately 50km by 50km at the equator) ([Beguería and Vicente Serrano, 2017](#)). The SPEI is a standardised index, so that a value of 1 represents weather conditions that are 1 standard deviation higher on the SPEI scale compared to the local historical average (1900-2015). I use GIS shapefile data on Indonesian districts to match each district to its overlapping grids and calculate district-level SPEI by taking a weighted average of these overlapping grids.

This monthly district SPEI is used to construct the main  $SPEI_{i(t+j)}$  variable used throughout the paper by calculating the average district SPEI over the growing season of the main crop of that district in that period. For the main specification, I match this shock back to individuals using the constructed year-by-year record of each individual's location described above. For example, the variable  $SPEI_{i(t+2)}$  for an individual  $i$  born in  $t$  would denote the average SPEI during the growing season of the main crop of the district where  $i$  spent the year of her life after her 2nd birthday.

I describe this process in more detail, along with the relevant sources for each dataset, in Appendix Section A.2.

### 3.3 Measurement of skills

The use of the psychological outcome measures seen in the IFLS as outcome variables presents at least two key empirical challenges. First, psychological measures tend to be "riddled with measurement error" ([Cunha and Heckman, 2009](#); [Heckman et al., 2013](#)). If such error is classical, it may lead to imprecise estimates of causal effects. If it is non-classical, then it may bias the estimated impact of weather shocks on psychological outcome variables. Moreover, such error may be even more severe in

low- and middle-income country contexts (Laajaj and Macours, 2017; Laajaj et al., 2018). Second, each psychological measure doesn't truly represent an independent variable. For example, the responses to two questions such as (i) "I see myself as someone who is outgoing and sociable" and (ii) "I see myself as someone who is talkative" are likely to be two proxy measures of a single psychological trait, in this case extraversion. We therefore require a method of condensing highly intercorrelated measures down to a lower-dimensional set of psychological skills.

In this paper I use the latent factor model seen in Cunha et al. (2010) to combine psychological measures.<sup>11</sup> This technique is commonly used in the psychometric literature, and it partially deals with the two problems described above by explicitly accounting for random (but not systematic) measurement error and by reducing observed measures down to a lower-dimensional set of latent factors. Suppose, for example, that for each individual  $i$  there are three different survey questions that relate to extraversion  $\{Y_i^1, Y_i^2, Y_i^3\}$ . The latent factor model assumes that these all measure the underlying factor  $F_i$  (e.g. extraversion) along with some measurement error  $\{u_i^1, u_i^2, u_i^3\}$ . So we can write:

$$\begin{cases} Y_i^1 &= \alpha_1 F_i + u_i^1 \\ Y_i^2 &= \alpha_2 F_i + u_i^2 \\ Y_i^3 &= \alpha_3 F_i + u_i^3 \end{cases}$$

We can then use a process of exploratory factor analysis to recover the latent factor score  $F_i$  from each of the measures  $\{Y_i^1, Y_i^2, Y_i^3\}$ , for example by using maximum likelihood estimation to fit the model.

The process and results used to generate factor scores is described in detail in Appendix Section B, and the results are summarised in Table 1. Running a factor analysis exercise on cognitive measures suggests that there is a single cognitive factor that is proxied by all the cognitive measures available in the data. Running a separate factor analysis exercise on socioemotional measures indicates that there are 3 latent factors being measured. The measures in each of the 3 latent factors align relatively closely with the survey module structure, meaning that each has an easily applicable interpretation (based on the measures from which it is composed)<sup>12</sup>. The first I interpret as "affect / wellbeing". This is mostly composed of measures from the Positive-Negative Affect module of the IFLS survey, which asks about which emotions the individual felt in the past week, and also includes standard life-evaluation questions such as the Cantrill Ladder. The second factor I interpret as "personality", and is composed of questions from the Big 5 personality test. The interpretation of this factor is somewhat less clear as it mixes questions which are meant to identify distinct personality traits such as extraversion and conscientiousness.<sup>13</sup> The third factor is exclusively composed of questions from the CES-D depression scale, and so can be clearly interpreted as "mental-health". A fuller discussion of the interpretation of each factor is found in Appendix Section B. Having generated the factors from the psychological measures, I then have factor scores for each individual which are used as dependent variables in the main results of the paper. Note that these factors are always to be interpreted "positively" (in line with previous literature which sees these traits as skills, see e.g. Cunha et al., 2010). For example, a higher mental health factor score will mean *better* mental health or *less* depression. Moreover, the scores are normalised to have a mean of 0 and a standard deviation of 1 so that they can be easily interpreted. For example, a mental health factor score of 1 can be interpreted as saying that the individual's mental health is one standard deviation above the sample average.

<sup>11</sup>For similar reasons, I also use a latent factor model to create the variables used in the mediation analysis seen in Section 6.3.

<sup>12</sup>The full list of questions associated with each socioemotional measure can be found in Appendix Table 13.

<sup>13</sup>This is in keeping with Laajaj and Macours (2017) and Laajaj et al. (2018). The five personality traits in the Big 5 are based on factor analyses of populations in upper-income countries. The authors show that these traits are less likely to show up in the factor structure in lower-income contexts, and surveys often point towards a single "personality" factor.

**Table 1: Psychological Latent Factors and their corresponding measures**

<i>Cognitive</i>	<i>“Affect / Wellbeing”</i>	<i>Socioemotional “Personality”</i>	<i>“Mental Health”</i>
Raven’s Matrices	Affect: Frustrated	Big 5: Agreeableness - Forgiving	CES-D: I was bothered by things
Maths - Written	Affect: Sad	Big 5: Agreeableness - Considerate	CES-D: I had trouble concentrating
Word Recall	Affect: Enthusiastic	Big 5: Agreeableness - Rude	CES-D: I felt depressed
Maths - Oral	Affect: Lonely	Big 5: Conscientiousness - Thorough	CES-D: I felt hopeful
Number Series	Affect: Content	Big 5: Conscientiousness - Lazy	CES-D: I felt fearful
	Affect: Worried	Big 5: Extroversion - Outgoing	CES-D: My sleep was restless
	Affect: Bored	Big 5: Neuroticism - Relaxed	CES-D: I was happy
	Affect: Happy	Big 5: Neuroticism - Worries	CES-D: I felt lonely
	Affect: Angry	Big 5: Neuroticism - Gets Nervous	CES-D: I could not get going
	Affect: Tired	Big 5: Openness - Original	
	Affect: Stressed	Big 5: Openness - Artistic	
	Wellbeing: Life Satisfaction		
	Wellbeing: Assessment situation		
	CES-D: I was happy		

*Notes:* Each column represents a single factor, containing a list of all the measures which are deemed to load on that factor. Full description of each of the variables, methodology for the creation of each factor, and factor loadings tables are all found in Appendix Section B.

### 3.4 Samples and summary statistics

In order to make full use of the wealth of the IFLS data in order to understand both the effects of weather shocks and the mechanisms through which they can affect adult psychological outcomes, I run analyses on a number of different samples with different units of observation. For clarity, these samples are summarised in 2. All samples are restricted to rural only, under the assumption that urban households are unlikely to be affected by weather shocks, while rural households are likely to be impacted by these shocks through changes in agricultural incomes. The “main sample” used in Sections 5.1 and 5.2 includes all the individuals born after 1988 in the IFLS data for whom there is either socio-emotional or cognitive outcome data available. Since the first wave of the IFLS data was in 1993, focusing on individuals born after 1988 means that I will have some data available on individuals during their childhood (such as measures of early life investment in the child), which can then be used in a mediation analysis.<sup>14 15</sup> The mediation samples (one for each mediator) are the subsets of the main sample that have non-missing values for the relevant mediator (see more discussion in Section 6.3.1 on the methodology for mediation analysis). Both the main and the mediation samples are at the level of the individual.<sup>16</sup> For the results on household expenditure (Section 6.1), I do not restrict to households containing individuals in the main sample but instead use all available rural households in the IFLS data. I also use a repeated cross-sectional specification in which one observation is a (household × wave). For the results on anthropometrics (Section 6.2) I also do not restrict to the main sample, and instead increase sample size by using all individuals who were measured at least once between the ages of 0 and 5 years old in rural households<sup>17</sup>. This sample is also treated as a repeated cross-section where one observation is (individual × wave).

Table 3 presents summary statistics for individuals in the main sample. As noted above, the factor scores for each of the psychological variables is normalised so that by construction, the mean is 0 and the standard deviation is 1. Note that since Affect/Wellbeing and Personality were only measured in the 5th wave of IFLS (in 2015), any members interviewed only in IFLS4 (in 2007) have missing

<sup>14</sup>The main sample does not include individuals born after 2000 because the last wave of the IFLS survey was in 2015 and individuals were only asked questions on psychological outcomes if they were at least 15 years old.

<sup>15</sup>It also makes it more likely that my definition of a weather shock (which is based on the growing season for the main crop in terms of harvested area, based on data from 2000) is most relevant to agricultural yield.

<sup>16</sup>This was appropriate since many psychological variables are only available in the most recent wave, IFLS5 (2015). In cases where individuals had psychological variables measured in both IFLS4 and IFLS5, I generated combined factor scores by taking the unweighted mean of the factor score from both waves.

<sup>17</sup>Again since the first IFLS wave was in 1993, this means that the oldest members of this sample were born in 1988.

**Table 2:** *Guide to different IFLS samples used*

<i>Sample</i>	<i>N</i>	<i>Level of observation</i>	<i>Year of birth for sample members</i>	<i>Rural</i>	<i>Other restrictions</i>
Main Sample	5903	Individual	1988–2000	Born Rural	Only individuals with psych variables
Mediation Samples	3242 to 5528	Individual	1988–2000	Born Rural	Only individuals with psych & relevant mediation variable
Household Expenditure	12618	Household x Wave	-	Rural Households	-
Anthropometrics	9823	Individual x Wave	1988–2015	Rural Households	Only 0–5 year olds

values for these variables. Therefore N is only 4672 for these factors.<sup>18</sup> Roughly half of the sample (47%) were born during the growing season of the main crop in their district. The main weather shock variables used in Results section 5.1 and 5.2 are the SPEI growing season from  $t - 2$  to  $t + 15$  (denoting the weather conditions 2 years before birth up to 15 years after birth). Since the SPEI is a standardised index we expect the mean to be approximately 0. As we can see, the mean of SPEI growing season is above 0 for all years after birth, suggesting that on average, the weather conditions experienced by members of the main sample after birth were slightly more favourable than the historical average from 1900 to 2015. The standard deviation for each SPEI shock variable is consistently between 0.43 and 0.46, suggesting substantial variation between individuals in the weather conditions experienced at every age. As noted above, the main sample is restricted to only include individuals born after 1988 and the median individual is born in 1992. 54% of the main sample are female. The vast majority of the sample (90%) are Muslim while 5% are Hindu, 4% Protestant, and 1% some other religion. The highest level of education reached by a parent of sample members is typically elementary (52%), while 19% have a parent that reached Junior High School, 20% have a parent that reached Senior High, and 8% have a parent that reached university.

<sup>18</sup>Individuals who were born after 1988 were only tracked after moving household if they were part of the original sample from the first IFLS wave (1993), but will be missing if they were added in a later wave and moved households since IFLS4.

**Table 3: Summary Statistics - Main Sample**

Statistic	N	Mean	St. Dev.	Min	Median	Max
Factor Score - Cognitive	5815	0.00	1.00	-4.00	0.02	2.91
Factor Score - Affect/Wellbeing	4672	0.00	1.00	-4.45	0.13	2.18
Factor Score - Personality	4672	0.00	1.00	-3.94	0.03	3.15
Factor Score - Mental Health	5418	0.00	1.00	-4.29	0.12	1.77
Born during the growing season (=1)	5892	0.47	0.50	0.00	0.00	1.00
SPEI growing season in t - 2	5892	-0.04	0.46	-1.76	-0.01	1.54
SPEI growing season in t - 1	5892	-0.003	0.47	-1.78	0.03	1.57
SPEI growing season in t	5892	0.02	0.46	-1.72	0.05	1.54
SPEI growing season in t + 1	5892	0.02	0.45	-1.76	0.07	1.34
SPEI growing season in t + 2	5892	0.03	0.45	-1.72	0.08	1.34
SPEI growing season in t + 3	5892	0.07	0.45	-1.69	0.12	1.34
SPEI growing season in t + 4	5892	0.05	0.45	-1.60	0.10	1.34
SPEI growing season in t + 5	5892	0.07	0.48	-1.60	0.11	1.60
SPEI growing season in t + 6	5892	0.07	0.48	-1.51	0.12	1.60
SPEI growing season in t + 7	5892	0.07	0.46	-1.60	0.10	1.60
SPEI growing season in t + 8	5892	0.08	0.48	-1.52	0.10	1.60
SPEI growing season in t + 9	5892	0.11	0.46	-1.52	0.12	1.60
SPEI growing season in t + 10	5892	0.16	0.43	-1.52	0.17	1.91
SPEI growing season in t + 11	5892	0.14	0.43	-1.52	0.13	1.91
SPEI growing season in t + 12	5892	0.15	0.44	-1.52	0.15	1.91
SPEI growing season in t + 13	5892	0.16	0.45	-1.52	0.17	1.91
SPEI growing season in t + 14	5892	0.13	0.45	-1.52	0.14	1.91
SPEI growing season in t + 15	5892	0.13	0.45	-1.64	0.15	1.91
Female (=1)	5868	0.54	0.50	0.00	1.00	1.00
Year of Birth	5903	1993.07	3.56	1988	1992	2000
Highest level of parent education - Elementary (=1)	5609	0.52	0.50	0.00	1.00	1.00
Highest level of parent education - Junior High (=1)	5609	0.19	0.39	0.00	0.00	1.00
Highest level of parent education - Senior High (=1)	5609	0.20	0.40	0.00	0.00	1.00
Highest level of parent education - University (=1)	5609	0.08	0.28	0.00	0.00	1.00
Religion - Muslim (=1)	5903	0.90	0.30	0	1	1
Religion - Protestant (=1)	5903	0.04	0.19	0	0	1
Religion - Hindu (=1)	5903	0.05	0.22	0	0	1
Religion - Other (=1)	5903	0.01	0.11	0	0	1

## 4 Empirical Strategy

To test for the effect of early-life weather shocks on adult psychological outcomes, I estimate the following equation:

$$Y_{irt} = \alpha + \sum_{j=-2}^{J=15} \beta_j SPEI_{i(t+j)} + \pi G_{ir} + \mathbf{X}'_i \gamma + \delta_r + \mu_t + \varepsilon_{irt}$$

Here,  $Y_{irt}$  denotes the adult psychological outcome variable of interest for an individual  $i$  born in region  $r$  at time  $t$ .<sup>19</sup> In Section 5.1, I use the factor score for cognitive skills, while in Section 5.2 I use the three socio-emotional factor measures: affect/wellbeing, personality, and mental health. My weather shock variable is  $SPEI_{i(t+j)}$ . This is a shock variable based on the SPEI climate indicator as described in Section 3, and matched back to individuals so that  $SPEI_{i(t+j)}$  describes the average SPEI over the growing season of the main crop in the district where  $i$  was located in the year after their  $j$ th birthday, according to the constructed location history I described in Section 3.1.<sup>20</sup>  $G_i$  is an indicator variable that takes the value 1 when the individual  $i$  was born during the growing season of their

<sup>19</sup>What I here call “region” corresponds to an Indonesian *kabupaten*. There are individuals from 187 regions in the main sample.

<sup>20</sup>For years before birth, the district of birth is used. More details on the construction of this variable is in Appendix Section A.2.



birth district  $r$ .  $X_i$  is a vector of individual-level controls, that (in the preferred specification) includes dummies for the highest level of education achieved by a parent of  $i$ , dummies for religion, and a female dummy.  $\delta_r$  and  $\mu_t$  are birth-region and birth-year fixed effects, while  $\alpha$  denotes the intercept.

The coefficients of interest in this specification are  $\beta_j$ ,  $j \in \{-2, -1, \dots, 15\}$ . These allow me to test for the effects of weather shocks at different stages of life on adult psychological outcomes. In particular, I can look at the different effects of weather conditions before birth ( $j \in \{-2, -1\}$ ), at birth ( $j = 0$ ), in early childhood ( $j \in \{1, \dots, 5\}$ ) and later childhood ( $j \in \{6, \dots, 15\}$ ). The key advantage of my empirical setting when trying to make such distinctions is that weather shocks are unlikely to be serially correlated over multiple years.<sup>21</sup> To illustrate, the empirical correlation in the main sample between  $SPEI_{i(t+s)}$  and  $SPEI_{i(t+s-1)}$  (that is, the correlation between a shock and the 1 period lag) is only 0.076 (95% confidence interval: [0.071, 0.081]). This means that the weather shocks provide significant variation that is independent year on year and allows me to make fine-grained distinctions between effects in each year.

The birth-year fixed effects control for any unobserved factors that are common to all individuals in the sample born in the same year, such as macroeconomic shocks. The birth-region fixed effects control for any time-invariant unobserved factors that are common across all individuals born in the same region. My identification strategy then relies on the claim that conditional on these birth-year and birth-region fixed effects, the SPEI shocks are exogenous.

One notable threat to this identification strategy would be if there are endogenous reactions to weather shocks that confound the causal effect of weather shocks on adult psychology. Suppose, for example, that some families are able to migrate within a given year to areas with favourable weather, and families of higher cognitive ability are more likely to do so. In this case the positive correlation between  $SPEI_{i(t+j)}$  and adult cognitive ability would confound the causal effect of a shock with this selection effect. This seems unlikely; although migration is common in the sample, it is implausible that rural households are able to react sufficiently quickly to seasonal weather patterns. Another possibility is selective mortality. For example, if extreme negative weather shocks (such as drought) lead to the death of children, and these deaths are concentrated upon those children who would otherwise have developed poorly, then any positive relationship between weather shocks and adult psychological outcomes would be biased downwards. I test for such an effect in Appendix Section D, and show that while there may be a positive impact of SPEI shocks in utero on infant mortality, the magnitude is small and appears not to affect the probability of inclusion in my main sample. This implies that selection bias will be close to 0, although it cannot be ruled out entirely. Absent such selection bias, the coefficients on  $\beta_j$  will be unbiased estimates of the impact of early-life weather shocks on adult psychology.

As noted in the introduction, one key advantage of the IFLS data is that it includes full migration histories and thus allows me to construct a year-by-year record of the location of each individual (and then match each individual to a shock). I make use of retrospective survey questions that ask about location of birth and subsequent migrations. There is likely to be non-negligible measurement error in the retrospective migration histories given by respondents; it may be difficult for respondents to remember the year in which they migrated if it was a long time ago, especially if they were young at the time. If such measurement error is classical, it will lead to attenuation bias on the coefficients estimated in the main specification.<sup>22</sup> However, since my main sample was born between 1988 and

<sup>21</sup>Compare this to the empirical strategy employed in [Adhvaryu and Fenske \(2016\)](#). Since cocoa prices in Ghana are highly autocorrelated year-on-year, they do not have enough independent variation in their shock variable to be able to distinguish the timings of different effects, resulting on very high standard errors on estimates when including shocks from multiple years in one regression.

<sup>22</sup>Note too, that attenuation bias may vary by age. First, birth locations and years are likely to be more accurate than the year of migration, so attenuation bias may be smaller for years close to birth than in later years. On the other hand, more recent migrations may be remembered more accurately than more distant ones, which would imply that later years are less attenuated. The net effect is ambiguous.

2000, many individuals were children in IFLS households during IFLS survey-wave years (1993, 1997, 2000, 2007, 2015). Therefore, to construct the year-by-year record I combine the retrospective questions with the contemporaneous data on current location of individuals in each of the survey years. This is likely to mitigate at least some of the measurement error stemming from retrospective migration history.

To test for the existence of critical periods in early childhood development, I need be able to answer two questions. First, I need to test whether there is an overall effect of early childhood weather shocks on adult psychological outcomes (regardless of the timing). To do this, I make use of F-tests that test the joint null hypothesis  $H_0 : \beta_{-2} = \dots = \beta_{15} = 0$ . A significant result on this test provides evidence for an overall effect of early life weather shocks on adult outcomes.

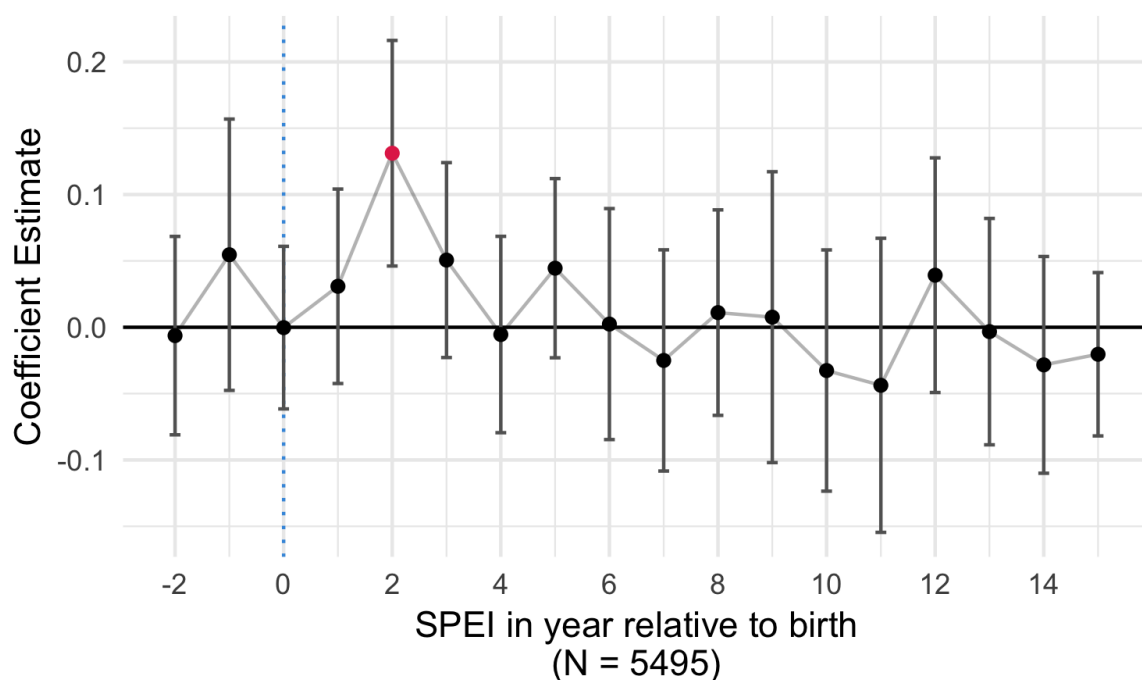
Second, I need to test whether the effect can be isolated to any specific periods. To do this, I will need to test the individual coefficients on specific years of life, i.e. test  $H_j : \beta_j = 0$ . However, given that there are 17 such coefficients to test, this means I will be testing the significance of 17 hypothesis tests simultaneously. Without making further adjustments, such multiple testing will dramatically increase the likelihood of detecting false positives. I account for this by calculating the [Benjamini and Hochberg \(1995\)](#) q-values for the set of  $\beta_j$  coefficients in my main specification (following [Anderson, 2008](#) in a similar context). These q-values are p-values adjusted to control for the false discovery rate (FDR), which describes the expected proportion of false rejections out of all rejections. (More detail on the process and the assumptions behind it are found in Appendix Section C.) Intuitively, a q-value of 0.05 indicates that we can reject the null hypothesis if we are willing to accept a false discovery rate (the average probability that any significant result is a false positive) of 5%. Using the q-value will lead to more conservative inference than the standard approach and reduce the probability of Type I errors. The advantage of this is that any significant effects of weather shocks at a certain age can be confidently asserted to be true effects rather than artefacts of the testing procedure, and thus allow me to test for the existence of critical periods in early childhood.

## 5 Main Results

### 5.1 Cognitive Skills

Table 4 shows the results of running the main specification with cognitive factor z-score as the outcome variable  $Y_{irt}$ , and a corresponding summary of results is shown in Figure 1. By controlling for birth-year and birth-district fixed effects, I account for unobservable differences across birth-year cohorts and birth-districts that could be correlated with both early-life weather patterns and adult cognitive skills. This allows me to interpret the coefficient estimates on SPEI growing season as causal estimates. In column (1), we see a very strong and statistically significant positive coefficient on SPEI growing season in the second period after birth (0.117), whereas none of the other coefficients on SPEI growing season are statistically different from 0. Adding individual-level controls for parental education, sex, and religion (column 2) leaves this coefficient largely unchanged at 0.131, and the p-value of the coefficient remains below 0.01, implying that the effect is robust to the addition of controls. Because of the concerns with testing multiple hypotheses described in Section 4, I show the FDR-controlled q-value in the final column, calculated using the procedure from [Benjamini and Hochberg \(1995\)](#) described in Appendix Section C. The q-value for SPEI growing season in  $t + 2$  is 0.046. This implies that the chance that the significant coefficient observed here is a false positive (given the fact that we have run 15 hypothesis tests) is below 5%, and so we can be confident that the true effect is different from 0. I also show the F-statistic and corresponding p-value for the joint hypothesis test of  $H_0 : \beta_{-2} = \dots = \beta_{15} = 0$ . The p-values of 0.078 in column (1) and 0.028 and (2) mean that we can reject the null at the 10% and 5% levels respectively. These results confirm the interpretation of the FDR procedure by suggesting that at least one of the coefficients  $\beta_j$  is significantly different from 0.

**Figure 1:** Effect of SPEI shock at different stages of childhood cognitive factor scores



Notes: Each point corresponds to a coefficient from column (2) of Table 4, so represents the effect of an SPEI shock in the specific year relative to birth on cognitive factor z-score in adulthood. All coefficients control for birth-year fixed effects, birth-district fixed effects, and individual-level controls (dummies for highest level of parental education, sex, and religion). Error bars are 95% confidence intervals (not adjusted for multiple testing). Points highlighted in red denote coefficients that have a q-value less than 0.05, implying that they are significantly different from 0 after adjusting for multiple testing.

The magnitude of the coefficient in  $t + 2$  is relatively large given the ‘mild’ nature of the shock. It implies that the average effect of increasing the SPEI during growing season from -0.5 to 0.5 (corresponding roughly to a change from somewhat unfavourable to mildly favourable conditions for agriculture) leads to an average increase of 0.13 standard deviations in adult cognitive score. This is comparable to the effects of interventions specifically designed to improve early life circumstance. For example, it is of the same order of magnitude as the 0.26 standard deviation effect on cognitive scores of a Colombian early childhood psychosocial programme described in [Attanasio et al. \(2014\)](#).

The results presented here can be interpreted as strong evidence in favour of the notion that there is a “critical period” in early childhood for cognitive development. Specifically, the fact that there is a large positive effect of weather shocks that exists only two years after birth suggests that *only* shocks in this time interval have an effect on adult cognitive skills, whereas shocks outside of the time interval have no such effect. The fact that an early-life critical period for cognitive development exists in the first place is in keeping with previous literature in economics ([Heckman and Mosso, 2014](#); [Barham et al., 2013](#)). The lack of effect later in life is also consistent with biological evidence suggesting that areas of the brain associated with spatial reasoning (important for the Raven’s matrices tests used as measures of cognitive ability in the IFLS data) develop in the first 2-3 years of childhood ([Grantham-McGregor et al., 2007](#)).

On the other hand, such literature typically cites the *entire* first 1000 days of childhood as a critical period. In light of this, it is surprising that the coefficients at age 0 and 1 are close to 0, suggesting that the critical period may be isolated to the year after the second birthday in this empirical setting. There are two possible interpretations for this result. First, we could take the reduced-form evidence at face value, indicating that shocks at age 2 are indeed critical for cognitive development in a way that shocks at ages 0 and 1 are not.<sup>23</sup> Brain areas associated with “higher” cognitive function (e.g. the

<sup>23</sup>In terms of the discussion in Section 2, this amounts to assuming that  $dm_{\text{cog}}(\cdot)/d\mu_t = \partial m_{\text{cog}}(\cdot)/\partial \mu_t$ .

pre-frontal cortex) and language and speech production tend to develop more after the first year of life (Grantham-McGregor et al., 2007). The null effects at age 0 and 1 may therefore be consistent with biological evidence if we suspect that the cognitive factor score may actually be detecting differences language or higher cognitive ability<sup>24</sup>. Nevertheless, the absence of impacts before age 2 is surprising. An alternative explanation therefore may be that shock impacts are dampened by compensating parental investment responses. For example, if negative weather shocks lead to poor harvest and poor nutrition at ages 0 and 1, but parents make up for this in later “good years” by investing in better nutrition, then this would dampen any reduced form effect of a shock in the first two years of life. The evidence presented in Section 6.3 would be consistent with such a story: although the coefficients are too imprecisely estimated to be confident, there may be an inverse relationship between early life SPEI and early life investment.

Much of the existing literature on early childhood shocks and interventions points to effects that vary by gender (see, for example, Maccini and Yang, 2009 or Heckman et al., 2013). Table 5, along with Figures 2a and 2b, present the results from running my main specification on boys and girls separately. Here we see a very large positive coefficient of 0.170 for boys that is strongly significant even after adjusting for multiple testing, with a q-value of 0.003. On the other hand, when restricting the sample to girls, the coefficient on  $t + 2$  is much lower in magnitude (0.077) and is only significant at the 10% level. Running a model equivalent to the main specification but including interaction terms between  $female_i$  and each of the SPEI growing season coefficients shows that the coefficient on  $(female_i \times SPEI_{i(t+2)})$  has a p-value of 0.08 (not shown), suggesting that the gender-difference is likely to be more than just an artefact of the sample. Taken together this evidence suggests that the positive coefficient on  $t + 2$  seen in Table 1 is driven mostly by an effect on the cognitive development of boys rather than girls. This is in accordance with existing evidence in the early childhood and biological literature that suggests that cognitive development in boys may be more vulnerable to shocks in early childhood and in utero than in girls (Eriksson et al., 2010). While many papers rightly point towards the importance of policies that focus on protecting young girls from shocks in developing countries (often because families may smooth consumption at the expense of girls, see e.g. Rose, 1999), this consideration provides a countervailing reason that could justify the prioritisation of young boys.

It is unclear from the linear specification seen in Table 1 what ‘type’ of weather shock is driving the positive relationship between  $SPEI_{i(t+2)}$  and cognitive factor score. *A priori* there may even be a highly non-linear relationship between SPEI and cognitive outcomes. For example, it could be the case that more rainfall and higher SPEI is typically good for crop harvest and therefore cognitive development, but that extremely heavy rainfall and very high SPEI values causes flood that hamper cognitive development. To address this concern, Figure 3 shows a non-parametric kernel regression of the shock variable ( $SPEI_{i(t+2)}$ ) on the residuals of cognitive factor score after conditioning for birth-year and birth-district fixed effects along with all the controls and other  $SPEI_{i(t+j)}$  variables found in column (2) of Table 1. The Figure depicts a relationship between the two that is well approximated by a linear function, with increases in  $SPEI_{i(t+2)}$  being on average associated with increases in cognitive factor score across almost the entire range of SPEI. This is consistent with a story in which there is a linear relationship between SPEI during growing season and crop yields, which translates into a linear relationship with cognitive scores in adulthood. There is some indication that the linear relationship may weaken at very low levels of SPEI, broadly implying that very extreme droughts have similar effects to significant but less severe ones. This would also be compatible with an explanation through agricultural yield in which crops fail at an SPEI growing season of -1 (so going below this has no additional negative effect), although absent detailed crop yield data this cannot be verified. Overall however, the assumption of a linear relationship between SPEI shocks and psychological outcomes that the main specification in this paper depends on appears to be verified by Figure 3.

<sup>24</sup>Indeed, the cognitive measures from the IFLS do include “word recall” tasks which may reflect language ability and “number series” tasks which may pick up elements of higher cognitive function. See A.1 for more details.

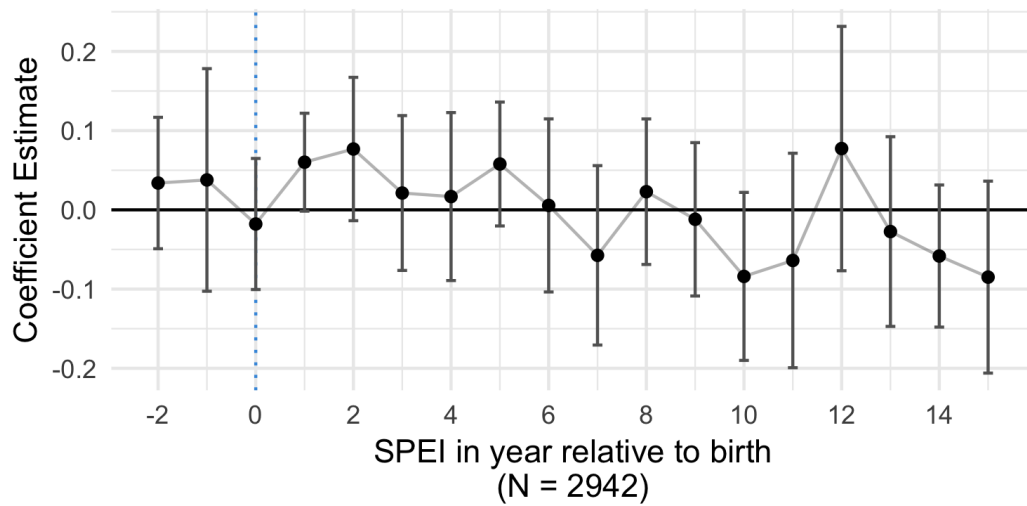
**Table 4:** *Effect of SPEI growing season on cognitive factor score of individual born in t*

	(1)	(2)	q value of (2)
SPEI growing season in t - 2	-0.009 (0.038)	-0.006 (0.038)	0.993
SPEI growing season in t - 1	0.062 (0.042)	0.055 (0.052)	0.911
SPEI growing season in t	0.027 (0.029)	-0.000 (0.031)	0.993
SPEI growing season in t + 1	0.035 (0.034)	0.031 (0.037)	0.911
SPEI growing season in t + 2	0.117*** (0.037)	0.131*** (0.043)	0.046
SPEI growing season in t + 3	0.027 (0.028)	0.051 (0.037)	0.911
SPEI growing season in t + 4	-0.011 (0.038)	-0.005 (0.038)	0.993
SPEI growing season in t + 5	0.060* (0.033)	0.045 (0.034)	0.911
SPEI growing season in t + 6	0.006 (0.043)	0.002 (0.044)	0.993
SPEI growing season in t + 7	-0.004 (0.036)	-0.025 (0.043)	0.911
SPEI growing season in t + 8	0.008 (0.032)	0.011 (0.040)	0.993
SPEI growing season in t + 9	0.015 (0.050)	0.008 (0.056)	0.993
SPEI growing season in t + 10	-0.037 (0.039)	-0.033 (0.046)	0.911
SPEI growing season in t + 11	-0.047 (0.049)	-0.044 (0.057)	0.911
SPEI growing season in t + 12	0.031 (0.041)	0.039 (0.045)	0.911
SPEI growing season in t + 13	-0.008 (0.042)	-0.003 (0.044)	0.993
SPEI growing season in t + 14	-0.023 (0.035)	-0.028 (0.042)	0.911
SPEI growing season in t + 15	-0.024 (0.031)	-0.020 (0.031)	0.911
Born during the growing season (=1)	0.053*** (0.020)	0.041* (0.021)	-
Birth Year Fixed Effects	Yes	Yes	
Birth District Fixed Effects	Yes	Yes	
Additional Controls	No	Yes	
F statistic	1.547	1.803	
p-value for F-test	0.078	0.028	
Adj. R <sup>2</sup>	0.129	0.182	
N	5804	5495	

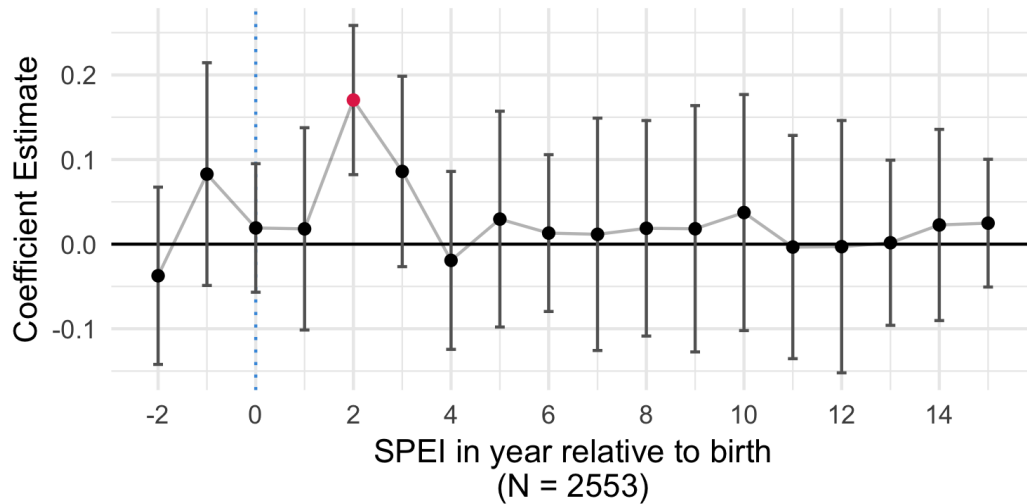
Notes: \*\*\* significant at the 1% level; \*\* significant at the 5% level; \* significant at the 10% level. Standard errors are reported in parentheses and are two-way clustered by birth year and birth district (Cameron et al., 2011). Cognitive factor score is internally standardised, so the sample mean is 0 by construction. Additional controls include dummies for the education level of the most-educated parent, dummies for religion, and a dummy for gender. Both models use the individual-level attrition-corrected weights provided in the IFLS data. The q-value is the FDR-adjusted p-value that is calculated according to the process in Benjamini and Hochberg (1995). F-statistic and F-test are for the joint hypothesis test  $H_0 : \beta_{-2} = \dots = \beta_{15} = 0$ , i.e. all the coefficients on SPEI growing season are 0.

**Figure 2: Effect of SPEI shock by gender**

**(a) Cog shock impact on females**



**(b) Cog shock impact on males**



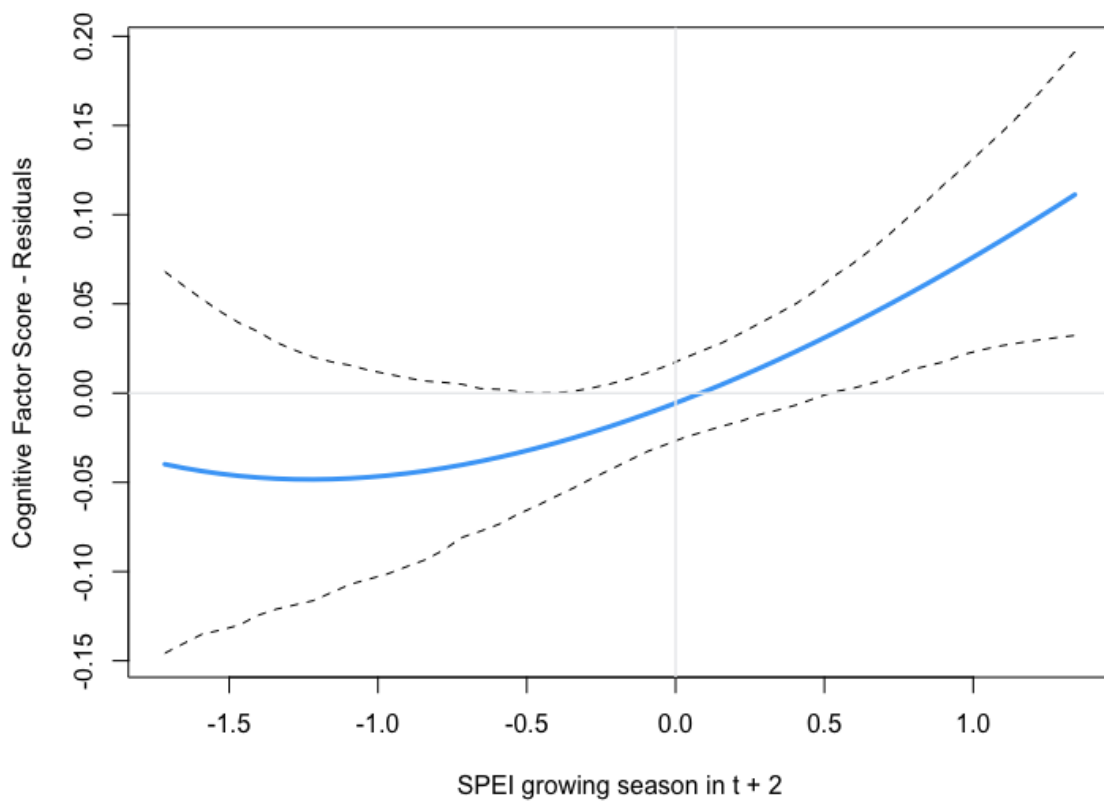
Notes: Each point corresponds to a coefficient from columns (1) and (2) of Table 5 respectively, so represents the effect of an SPEI shock in the specific year relative to birth on cognitive factor z-score in adulthood for each gender. All coefficients control for birth-year fixed effects, birth-district fixed effects, and individual-level controls (dummies for highest level of parental education, sex, and religion). Error bars are 95% confidence intervals (not adjusted for multiple testing). Points highlighted in red denote coefficients that have a q-value less than 0.05, implying that they are significantly different from 0 after adjusting for multiple testing.

**Table 5:** *Effect of SPEI growing season on cognitive factor score of individual born in t by gender*

	(1) Female	q value of (1)	(2) Male	q value of (2)
SPEI growing season in t - 2	0.034 (0.042)	0.763	-0.037 (0.053)	0.974
SPEI growing season in t - 1	0.038 (0.072)	0.807	0.083 (0.067)	0.974
SPEI growing season in t	-0.018 (0.042)	0.807	0.019 (0.039)	0.974
SPEI growing season in t + 1	0.060* (0.032)	0.609	0.018 (0.061)	0.974
SPEI growing season in t + 2	0.077* (0.046)	0.609	0.170*** (0.045)	0.003
SPEI growing season in t + 3	0.021 (0.050)	0.807	0.086 (0.057)	0.974
SPEI growing season in t + 4	0.017 (0.054)	0.852	-0.019 (0.054)	0.974
SPEI growing season in t + 5	0.058 (0.040)	0.609	0.030 (0.065)	0.974
SPEI growing season in t + 6	0.006 (0.056)	0.921	0.013 (0.047)	0.974
SPEI growing season in t + 7	-0.057 (0.058)	0.71	0.012 (0.070)	0.974
SPEI growing season in t + 8	0.023 (0.047)	0.807	0.019 (0.065)	0.974
SPEI growing season in t + 9	-0.012 (0.049)	0.857	0.018 (0.074)	0.974
SPEI growing season in t + 10	-0.084 (0.054)	0.609	0.037 (0.071)	0.974
SPEI growing season in t + 11	-0.064 (0.069)	0.71	-0.003 (0.067)	0.974
SPEI growing season in t + 12	0.077 (0.079)	0.71	-0.003 (0.076)	0.974
SPEI growing season in t + 13	-0.027 (0.061)	0.807	0.002 (0.050)	0.974
SPEI growing season in t + 14	-0.058 (0.046)	0.609	0.023 (0.058)	0.974
SPEI growing season in t + 15	-0.085 (0.062)	0.609	0.025 (0.039)	0.974
Born during the growing season (=1)	0.068 (0.044)	-	0.018 (0.040)	-
Birth Year Fixed Effects	Yes		Yes	
Birth District Fixed Effects	Yes		Yes	
Additional Controls	Yes		Yes	
F statistic	0.842		2.562	
p-value for F-test	0.567		0.012	
Adj. R <sup>2</sup>	0.209		0.166	
N	2942		2553	

Notes: \*\*\* significant at the 1% level; \*\* significant at the 5% level; \* significant at the 10% level. Standard errors are reported in parentheses and are two-way clustered by birth year and birth district (Cameron et al., 2011). Cognitive factor score is internally standardised, so the overall sample mean (with both female and male) is 0 by construction. Additional controls are included in both regressions, and include dummies for the education level of the most-educated parent, dummies for religion, and a dummy for gender. Both models use the individual-level attrition-corrected weights provided in the IFLS data. The q-value is the FDR-adjusted p-value that is calculated according to the process in [Benjamini and Hochberg \(1995\)](#). F-statistic and F-test are for the joint hypothesis test  $H_0 : \beta_{-2} = \dots = \beta_{15} = 0$ , i.e. all the coefficients on SPEI growing season are 0.

**Figure 3:** Non-parametric kernel regression of  $SPEI_{t+2}$  on the residuals of cognitive factor score



*Notes:* The blue line denotes the estimates from a non-parametric kernel regression of  $SPEI_{i(t+2)}$  on the residuals of cognitive factor score after conditioning for birth-year fixed effects, birth-district fixed effects, and individual controls (parental education, gender, and religion). The specification is a local-linear regression that uses a Gaussian kernel and chooses a bandwidth of 0.95 using a process of least-squares cross-validation. The dashed lines show the bootstrapped 95% confidence intervals.



## 5.2 Socioemotional Outcomes and Mental Health

The early childhood literature indicates that cognitive skills are not the only important dimension that is both malleable in early childhood and important in adult life. Socioemotional skills are important to investigate for at least two reasons. First, they appear to have strong predictive power over success in various dimensions of adult life, including education and the labour market (Heckman et al., 2006). And second, we may be interested in some aspects of socioemotional “skills”, such as subjective well-being and mental health, as measures of overall welfare that are worth investigating in their own right. In order to investigate whether such psychological aspects are subject to an early life critical period in the context of rural Indonesia, I rerun the main specification described in Section 4 using each of the three measures of socio-emotional factors seen in Table 1 as a dependent variable.

Interpreting the results will require us to answer two questions: (i) Is there any effect of weather shocks in early life on adult socioemotional scores? and (ii) Can the effects of such weather shocks in early life be isolated to a “critical period”? The first question is important because if there is no such effect then we will not be able to answer the second question. And the second question is exactly the main research question of this paper.

Table 6 and Figure 4 show the results for the first factor, which can be interpreted as “affect/wellbeing”. No clear pattern emerges from looking at individual coefficients. In both specifications, the majority of coefficients on SPEI growing season are positive with an average coefficient of 0.26 in both cases. In the specification with only fixed effects and without additional controls (column 1), the coefficients on  $SPEI_{it}$ ,  $SPEI_{i(t+6)}$ ,  $SPEI_{i(t+10)}$ , and  $SPEI_{i(t+14)}$  are all significantly different from 0 at the 10% level. After adding controls for parental education, religion and gender, coefficient magnitudes remain roughly stable, but now  $SPEI_{i(t-1)}$ ,  $SPEI_{i(t+2)}$ ,  $SPEI_{i(t+10)}$ ,  $SPEI_{i(t+14)}$  are significant at the 10% level. Surprisingly,  $SPEI_{i(t+14)}$  has a strongly *negative* effect with a coefficient of -0.105, implying *prima facie* that positive weather shocks at this age lead to a decrease by approximately 0.1 standard deviations in affect/well-being factor score.

Given the lack of clear pattern in these results, it is important to make use of the statistics that amalgamate the results from all of the coefficients. First, the final column in Table 6, indicating the Benjamini and Hochberg (1995) q-value for each of the coefficients, shows that none of the individual coefficients can be rejected at the 5% level after adjusting for multiple testing. This indicates that the significant coefficients on specific years may be a result of random noise in the estimates rather than true effects in each year. On the other hand, the F-test of  $H_0 : \beta_{-2} = \dots = \beta_{15} = 0$  yields a p-value of 0.028 and 0.026, suggesting that at least one of the weather shock variables is having an effect on adult affect/wellbeing factor score.<sup>25</sup>

In Appendix Table 27 I examine whether the results on affect and wellbeing differ by gender. Here we see clear evidence that most of the significant effects in Table 6 are driven by girls. The coefficients for girls on  $SPEI_{i(t-1)}$  and  $SPEI_{i(t+2)}$  have q-values less than 0.1 and are strongly positive (0.144 and 0.163 respectively), while the coefficient for girls on  $SPEI_{i(t+14)}$  has a q-value of 0.033 and is strongly negative (-0.154). I interpret this as evidence against the existence of a critical period in this domain for girls. Specifically, since the effect of shocks is not isolated to single periods but rather is spread out over in utero, early childhood (age 2), and teenage years. By contrast, the F-test is not rejected and there are no significant q-values at the 5% level on any coefficients for boys, implying that I cannot make strong claims about the existence or lack of critical period for the effects of weather shocks on the affect/wellbeing of boys.

It is unclear *a priori* why the impact of early life weather shocks on adult affect/wellbeing, including

<sup>25</sup>One possible explanation for the disparity between the results of the multiple-testing correction and the F-test is related to correlation between the SPEI shocks. Although sample correlation between shocks year-on-year is low (0.07), multicollinearity between any of the SPEI variables will lead to higher standard errors on their coefficients. This will get passed on to the inflated q-values and reduce the power to reject null hypotheses when testing individual coefficients. The F-test, on the other hand, tests all coefficients jointly and so doesn't suffer from this problem.

the negative effect seen at age 14, should be isolated to girls in this context. Possibly, an additional set of channels is operating for girls and not boys, for example through changes in fertility or age of menarche. In particular, it may be the case that nutrition shocks in the early teenage years are associated with delayed fertility for females,<sup>26</sup> which could have a possible negative impact on adult wellbeing. Running a simple test of this proposition, in which I use a variation of my main specification with an indicator of whether a female respondent has ever given birth as an outcome variable,<sup>27</sup> shows no significant effect of shocks in teenage years on the probability to have given birth, implying that fertility is not a mechanism for the impact at age 14.

The ambiguity in the results on affect/wellbeing is evident; any interpretation of the data should be made with caution. We can however arrive at a tentative conclusions. First, the results on the F-test appear to show that weather shocks between  $t - 2$  and  $t + 15$  do have *some* effect on adult affect/wellbeing, although the timing and magnitude of such an effect is unclear given that it is possibly spread out over all of the first 15 years of life and in utero.<sup>28</sup> Second, the fact that none of the q-values in Table 6 can be rejected at the 5% level gives no clear evidence of a critical period for affect/wellbeing when considering both genders together. Third, breaking down the results by gender shows that the impact of early life weather shocks on adult affect/wellbeing is isolated to girls in this context. Since weather shocks on girls have effects of different magnitudes and even directions at *various* points throughout childhood, this seems to provide evidence against the existence of critical periods for girls. Taken together, the evidence presented here is stronger for girls than for boys, but generally runs counter to the notion of a socioemotional critical period. Broadly, my results are in line with Heckman's claims that socioemotional development is more malleable throughout later childhood than cognitive development (Cunha et al., 2010), although with the additional caveat that there is only robust evidence for any malleability at all for girls.

When looking at the other two measures of adult socioemotional outcomes, I find no evidence of the impact of early-life weather shocks. Table 7 and Figure 5 indicate that the SPEI shocks in childhood have no robust impact on personality scores in adulthood. Once controlling for fixed effects, only the coefficients from  $t + 5$  and  $t + 7$  are statistically significant at the 10% level, but their sign is the opposite of expected and the significance disappears once basic individual-level controls are added (column 2). In keeping with this, the F-test on  $H_0 : \beta_{-2} = \dots = \beta_{15} = 0$  yields a p-value of 0.511, implying no evidence of an effect of early life weather shocks on adult personality.

Similarly, Table 8 and Figure 6 show the impacts of SPEI shocks at different stages of childhood on adult mental health factor scores. Notably, there seems to be a strong positive coefficient on  $SPEI_{i(t-1)}$  that remains highly stable after adding additional controls, and is significant at the 5% level. The coefficient of 0.082 in column (1) would indicate that children who are exposed to 1 standard deviation higher SPEI conditions when in utero are expected to have mental health factor scores that are approximately 0.08 standard deviations higher in adulthood. Such a result would be in keeping with a recent literature that shows that exposure to shocks of various kinds when in utero can have an effect on adult mental health, typically mediated by nutrition, maternal health, or maternal stress (Adhvaryu and Fenske, 2016; Persson and Rossin-Slater, 2018; Singhal, 2018; Pasha et al., 2018). However, once accounting for multiple testing in Table 8 we cannot reject that the coefficient is different from 0. The q-value of the coefficient on  $SPEI_{i(t-1)}$  is 0.411, implying that there is a high 41% chance that this significant result is a false rejection after adjusting for multiple-testing. Moreover, the p-value for the F-test of all coefficients being 0 is 0.565, implying little evidence to reject the null that early life weather shocks have an effect on adult mental health.

<sup>26</sup>Such channels have been shown to operate as the mechanisms that explain some of the impacts of conditional cash transfers in Nicaragua (Barham et al., 2018).

<sup>27</sup>Since affect/wellbeing is only measured at one period in time (the 5th wave of IFLS in 2015), and I am controlling for birth-year fixed effects, I am effectively controlling for age fixed effects in this regression.

<sup>28</sup>An indication of the magnitude of effects is found in Appendix Figure 17, which shows the change in distribution of affect/wellbeing score as a result of an increase from -0.5 to 0.5 of *all* the values of SPEI.

Overall, I find no evidence for the existence of a critical period in the development of socioemotional outcomes in rural Indonesia. While weather shocks in early life appear to have no effect on adult personality or mental health outcomes, the results suggest similar weather shocks may have an influence on adult affect/wellbeing outcomes, with these effects being particularly concentrated on girls. The fact that it is difficult to distinguish the timing of this influence appears to reflect the fact that there is not a single critical period, but potentially multiple ages in which affect/wellbeing outcomes are weakly malleable, or malleable only for girls.

There are at least three plausible explanations for the overall observed lack of critical period for socioemotional outcomes. First, it could be due to weak overall impacts of weather shocks on socioemotional outcomes. If reduced form estimates are reflective of the direct impacts on socioemotional development, this would imply that such direct impacts are also small or 0. To some extent this would be in line with previous findings in the literature, which have tended to find much stronger and more robust effects of early childhood conditions on cognitive outcomes than on socioemotional ones (Almlund et al., 2011). The fact that strong impacts are observed for girls and not for boys sheds interesting new light on this claim. Second, it could be that parents make compensating investment responses to weather shocks, meaning that the effect of any critical or sensitive period is underestimated by the reduced-form coefficient. The evidence presented in Section 6.3 cannot rule out such a story, since although the coefficients are too imprecisely estimated to reject the null, there appears to be an inverse relationship between early life SPEI and early life investment. Moreover, if it is the case, as Heckman and coauthors have argued (Cunha and Heckman, 2009; Cunha et al., 2010), that socioemotional skills remain malleable for much longer than cognitive skills<sup>29</sup>, then there is a much longer time frame in which compensatory investments can take place for socioemotional outcomes than for cognitive outcomes. Such a consideration would provide at least a partial rationalisation of the results. Measurement challenges provide a third plausible explanation for the lack of observed critical period. Personality traits are typically harder to measure than cognitive ability (Laajaj and Macours, 2017); while there is an ‘objective’ answer about whether a respondent gets a cognitive test item right or wrong, no such impartial measure is available for socioemotional outcomes.<sup>30</sup> The latent factor model and the associated processes used when generating the factors (for example, correcting for acquiescence bias and accounting for measurement error in the generation of factors - see Appendix Section B for a full discussion) will mitigate such concerns and reduce attenuation bias to some degree.<sup>31</sup> Despite this, there is nevertheless likely to be persistent measurement error that could bias coefficients towards 0 and make it harder to detect the large effects of weather shocks during critical periods for socioemotional outcomes.

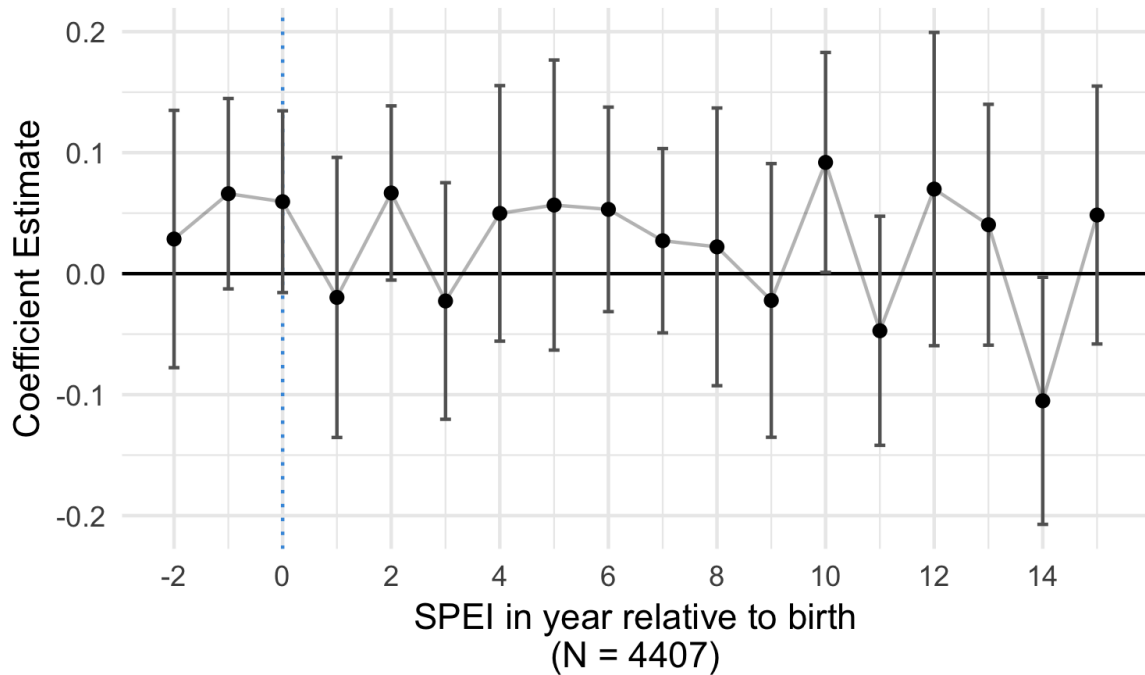
---

<sup>29</sup>Such a claim seems to comport with biological evidence showing that the prefrontal cortex, which governs “higher-order” cognitive function and socio-emotional behaviour, develops later than the areas of the brain that govern visual, spatial, and language capabilities (see Grantham-McGregor et al., 2007 or Sapolsky, 2017 for more detail).

<sup>30</sup>The problem of measurement error may be further compounded in low- or middle-income country contexts; Laajaj et al. (2018) show that noncognitive skill measures tend to perform worse on standard indexes of reliability and validity in lower-education and lower-income settings.

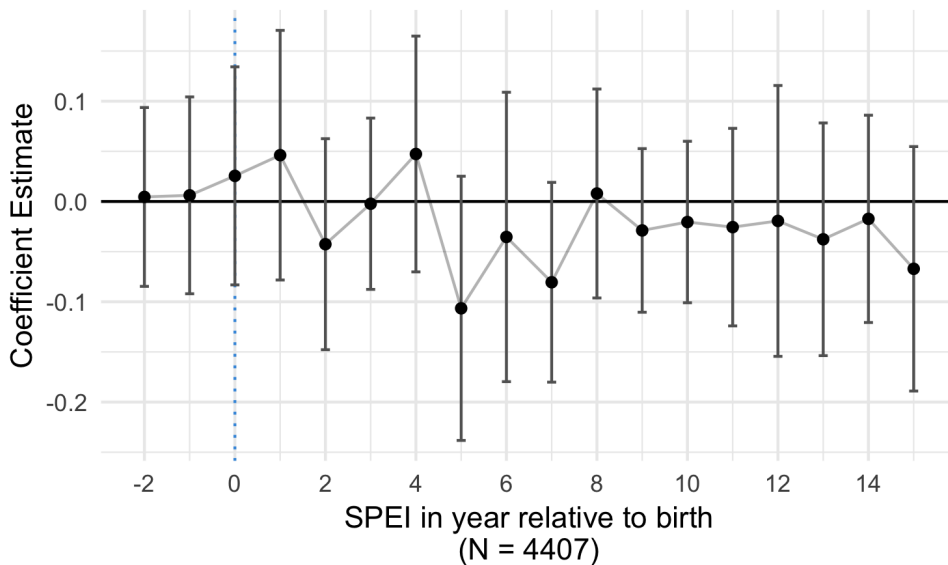
<sup>31</sup>Note too, that the factor analysis generates some of its own issues, particularly relating to the interpretation of factors. For example, it is unclear exactly what the “personality” factor used in Table 7 and Figure 5 is measuring. As per Table 17, this factor is made up of questions that relate to agreeableness, conscientiousness, extroversion, and openness, and combining these traits into one interpretation requires at least a stretch of the imagination. Similarly, the “affect / wellbeing” factor combines both cognitive evaluation and short-term affect questions, which are thought to be psychologically different phenomena (see e.g. Redelmeier et al., 2003; Kahneman and Deaton, 2010) This problem stems from the lack of a consistent factor structure behind personality traits measured in low- and middle-income countries, further documented in Laajaj et al. (2018). It could plausibly provide a fourth explanation for the lack of robust relationship between weather shocks and the socioemotional factors used here.

**Figure 4:** Effect of SPEI shocks on “affect/wellbeing” factor



Notes: Each point corresponds to a coefficient from column (2) of Table 6, so represents the effect of an SPEI shock in the specific year relative to birth on the “affect/wellbeing” factor z-score. All coefficients control for birth-year fixed effects, birth-district fixed effects, and individual-level controls (dummies for highest level of parental education, sex, and religion). Error bars are 95% confidence intervals (not adjusted for multiple testing). Points highlighted in red denote coefficients that have a q-value less than 0.05, implying that they are significantly different from 0 after adjusting for multiple testing.

**Figure 5:** Effect of SPEI shocks on “personality” factor



Notes: Each point corresponds to a coefficient from column (2) of Table 7, so represents the effect of an SPEI shock in the specific year relative to birth on the “personality” factor z-score. All coefficients control for birth-year fixed effects, birth-district fixed effects, and individual-level controls (dummies for highest level of parental education, sex, and religion). Error bars are 95% confidence intervals (not adjusted for multiple testing). Points highlighted in red denote coefficients that have a q-value less than 0.05, implying that they are significantly different from 0 after adjusting for multiple testing.

**Table 6:** Effect of SPEI growing season on “affect / wellbeing” factor score of individual born in  $t$ 

	(1)	(2)	q value of (2)
SPEI growing season in $t - 2$	0.025 (0.051)	0.029 (0.054)	0.74
SPEI growing season in $t - 1$	0.055 (0.039)	0.066* (0.040)	0.435
SPEI growing season in $t$	0.054* (0.032)	0.059 (0.038)	0.435
SPEI growing season in $t + 1$	0.001 (0.056)	-0.020 (0.059)	0.74
SPEI growing season in $t + 2$	0.061 (0.041)	0.067* (0.037)	0.418
SPEI growing season in $t + 3$	-0.022 (0.039)	-0.023 (0.050)	0.74
SPEI growing season in $t + 4$	0.023 (0.045)	0.050 (0.054)	0.611
SPEI growing season in $t + 5$	0.061 (0.047)	0.057 (0.061)	0.611
SPEI growing season in $t + 6$	0.077* (0.041)	0.053 (0.043)	0.611
SPEI growing season in $t + 7$	0.013 (0.029)	0.027 (0.039)	0.67
SPEI growing season in $t + 8$	0.025 (0.048)	0.022 (0.059)	0.74
SPEI growing season in $t + 9$	-0.008 (0.049)	-0.022 (0.058)	0.74
SPEI growing season in $t + 10$	0.084** (0.043)	0.092** (0.046)	0.418
SPEI growing season in $t + 11$	-0.032 (0.044)	-0.047 (0.048)	0.611
SPEI growing season in $t + 12$	0.060 (0.054)	0.070 (0.066)	0.611
SPEI growing season in $t + 13$	0.062 (0.039)	0.040 (0.051)	0.64
SPEI growing season in $t + 14$	-0.107*** (0.038)	-0.105** (0.052)	0.418
SPEI growing season in $t + 15$	0.039 (0.046)	0.048 (0.054)	0.611
Born during the growing season (=1)	0.071** (0.032)	0.067* (0.035)	-
Birth Year Fixed Effects	Yes	Yes	
Birth District Fixed Effects	Yes	Yes	
Additional Controls	No	Yes	
F statistic	1.750	1.711	
p-value for F-test	0.035	0.041	
Adj. R <sup>2</sup>	0.028	0.026	
N	4664	4407	

Notes: \*\*\* significant at the 1% level; \*\* significant at the 5% level; \* significant at the 10% level. Standard errors are reported in parentheses and are two-way clustered by birth year and birth district (Cameron et al., 2011). Affect/wellbeing factor score is internally standardised, so the sample mean is 0 by construction. Additional controls include dummies for the education level of the most-educated parent, dummies for religion, and a dummy for gender. Both models use the individual-level attrition-corrected weights provided in the IFLS data. The q-value is the FDR-adjusted p-value that is calculated according to the process in Benjamini and Hochberg (1995). F-statistic and F-test are for the joint hypothesis test  $H_0 : \beta_{-2} = \dots = \beta_{15} = 0$ , i.e. all the coefficients on SPEI growing season are 0.

**Table 7:** *Effect of SPEI growing season on “personality” factor score of individual born in t*

	(1)	(2)	q value of (2)
SPEI growing season in t - 2	0.004 (0.039)	0.005 (0.045)	0.959
SPEI growing season in t - 1	0.007 (0.046)	0.006 (0.050)	0.959
SPEI growing season in t	-0.002 (0.053)	0.026 (0.055)	0.959
SPEI growing season in t + 1	0.042 (0.060)	0.046 (0.063)	0.959
SPEI growing season in t + 2	-0.046 (0.050)	-0.043 (0.054)	0.959
SPEI growing season in t + 3	-0.010 (0.044)	-0.002 (0.044)	0.959
SPEI growing season in t + 4	0.033 (0.055)	0.047 (0.060)	0.959
SPEI growing season in t + 5	-0.101* (0.054)	-0.107 (0.067)	0.959
SPEI growing season in t + 6	-0.020 (0.068)	-0.035 (0.074)	0.959
SPEI growing season in t + 7	-0.075* (0.045)	-0.081 (0.051)	0.959
SPEI growing season in t + 8	0.013 (0.049)	0.008 (0.053)	0.959
SPEI growing season in t + 9	-0.031 (0.035)	-0.029 (0.042)	0.959
SPEI growing season in t + 10	-0.033 (0.034)	-0.020 (0.041)	0.959
SPEI growing season in t + 11	-0.025 (0.048)	-0.026 (0.050)	0.959
SPEI growing season in t + 12	0.005 (0.066)	-0.019 (0.069)	0.959
SPEI growing season in t + 13	-0.027 (0.051)	-0.038 (0.059)	0.959
SPEI growing season in t + 14	-0.001 (0.046)	-0.017 (0.053)	0.959
SPEI growing season in t + 15	-0.078 (0.059)	-0.067 (0.062)	0.959
Born during the growing season (=1)	0.048* (0.027)	0.044* (0.026)	-
Birth Year Fixed Effects	Yes	Yes	
Birth District Fixed Effects	Yes	Yes	
Additional Controls	No	Yes	
F statistic	0.915	0.957	
p-value for F-test	0.561	0.511	
Adj. R <sup>2</sup>	0.028	0.033	
N	4664	4407	

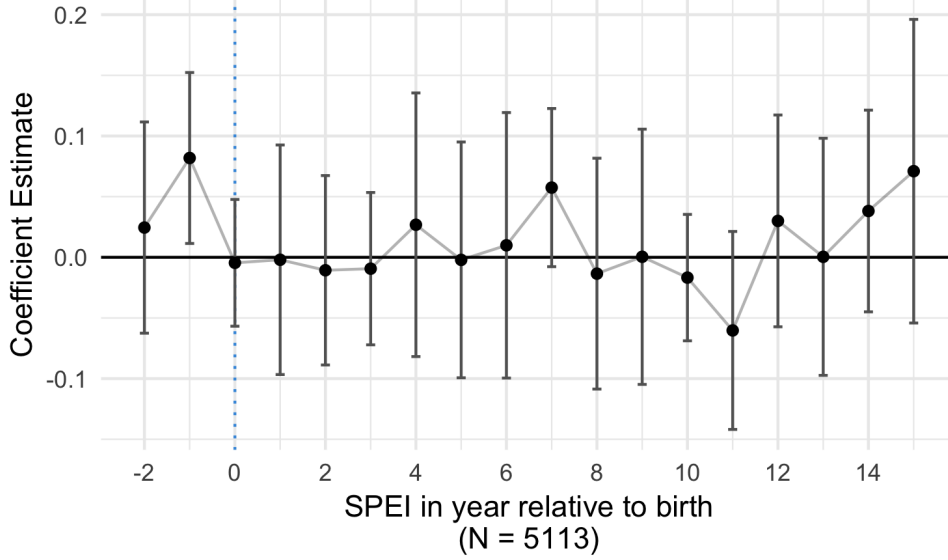
Notes: \*\*\* significant at the 1% level; \*\* significant at the 5% level; \* significant at the 10% level. Standard errors are reported in parentheses and are two-way clustered by birth year and birth district (Cameron et al., 2011). Personality factor score is internally standardised, so the sample mean is 0 by construction. Additional controls include dummies for the education level of the most-educated parent, dummies for religion, and a dummy for gender. Cognitive factor score is internally standardised, so the sample mean is 0 by construction. Both models use the individual-level attrition-corrected weights provided in the IFLS data. The q-value is the FDR-adjusted p-value that is calculated according to the process in [Benjamini and Hochberg \(1995\)](#). F-statistic and F-test are for the joint hypothesis test  $H_0 : \beta_{-2} = \dots = \beta_{15} = 0$ , i.e. all the coefficients on SPEI growing season are 0.

**Table 8:** *Effect of SPEI growing season on “mental health” factor score of individual born in t*

	(1)	(2)	q value of (2)
SPEI growing season in t - 2	0.016 (0.033)	0.025 (0.044)	0.994
SPEI growing season in t - 1	0.076** (0.034)	0.082** (0.036)	0.411
SPEI growing season in t	0.009 (0.024)	-0.005 (0.027)	0.994
SPEI growing season in t + 1	-0.003 (0.038)	-0.002 (0.048)	0.994
SPEI growing season in t + 2	0.012 (0.041)	-0.011 (0.040)	0.994
SPEI growing season in t + 3	-0.027 (0.030)	-0.009 (0.032)	0.994
SPEI growing season in t + 4	-0.006 (0.055)	0.027 (0.055)	0.994
SPEI growing season in t + 5	0.011 (0.043)	-0.002 (0.050)	0.994
SPEI growing season in t + 6	0.017 (0.052)	0.010 (0.056)	0.994
SPEI growing season in t + 7	0.051* (0.030)	0.057* (0.033)	0.757
SPEI growing season in t + 8	-0.001 (0.041)	-0.013 (0.049)	0.994
SPEI growing season in t + 9	-0.018 (0.051)	0.000 (0.054)	0.994
SPEI growing season in t + 10	-0.030 (0.022)	-0.017 (0.027)	0.994
SPEI growing season in t + 11	-0.060* (0.035)	-0.060 (0.042)	0.886
SPEI growing season in t + 12	0.030 (0.041)	0.030 (0.045)	0.994
SPEI growing season in t + 13	0.011 (0.043)	0.000 (0.050)	0.994
SPEI growing season in t + 14	0.024 (0.036)	0.038 (0.042)	0.994
SPEI growing season in t + 15	0.072 (0.053)	0.071 (0.064)	0.994
Born during the growing season (=1)	0.033 (0.027)	0.020 (0.031)	-
Birth Year Fixed Effects	Yes	Yes	
Birth District Fixed Effects	Yes	Yes	
Additional Controls	No	Yes	
F statistic	0.957	0.912	
p-value for F-test	0.511	0.565	
Adj. R <sup>2</sup>	0.038	0.045	
N	5409	5113	

Notes: \*\*\* significant at the 1% level; \*\* significant at the 5% level; \* significant at the 10% level. Standard errors are reported in parentheses and are two-way clustered by birth year and birth district (Cameron et al., 2011). Mental health factor score is internally standardised, so the sample mean is 0 by construction. Additional controls include dummies for the education level of the most-educated parent, dummies for religion, and a dummy for gender. Both models use the individual-level attrition-corrected weights provided in the IFLS data. The q-value is the FDR-adjusted p-value that is calculated according to the process in Benjamini and Hochberg (1995). F-statistic and F-test are for the joint hypothesis test  $H_0 : \beta_{-2} = \dots = \beta_{15} = 0$ , i.e. all the coefficients on SPEI growing season are 0.

**Figure 6:** Effect of SPEI shocks on “mental health” factor



Notes: Each point corresponds to a coefficient from column (2) of Table 8, so represents the effect of an SPEI shock in the specific year relative to birth on the “mental-health” factor z-score. Where measurements of mental health were available in both the 4th and 5th waves of the IFLS survey, the simple arithmetic mean of the factor score was taken for that individual. All coefficients control for birth-year fixed effects, birth-district fixed effects, and individual-level controls (dummies for highest level of parental education, sex, and religion). Error bars are 95% confidence intervals (not adjusted for multiple testing). Points highlighted in red denote coefficients that have a q-value less than 0.05, implying that they are significantly different from 0 after adjusting for multiple testing.

## 6 Mechanisms

### 6.1 Household expenditure

In order to begin to understand the mechanisms through which an SPEI shock at various stages of childhood could impact psychological development, I show below that SPEI shocks have a significant positive impact on contemporaneous household expenditure. To do this, I use a variation of my main specification adjusted to be at the household level:

$$E_{hrt} = \alpha + \sum_{j=-4}^{J=1} \beta_j SPEI_{r(t+j)} + \pi G_{hrt} + \mathbf{X}'_{hrt} \gamma + \delta_r + \sum_{k=2}^5 \mu_k \times \mathbb{1}\{WAVE = k\}_{hrt} + \varepsilon_{hrt} \quad (2)$$

$E_{hrt}$  denotes the log per capita expenditure of household  $h$  in district  $r$  in time  $t$ .  $SPEI_{rt}$  denotes the average SPEI index over the most recent growing season at time  $t$  of the main crop for region  $r$ .  $G_{hrt}$  is an indicator variable denoting whether the household  $h$  was interviewed during a growing season.  $\delta_r$  are regional fixed effects, and  $\mathbb{1}\{WAVE = k\}_{hrt}$  is an indicator variable that takes the value 1 when the household is interviewed in IFLS wave  $k$ .<sup>32</sup>  $\mathbf{X}_{hrt}$  is a vector of additional household-level controls including dummies for the religion of the household head, and whether the household head is female. In order to increase the precision of my estimates, I do not restrict the sample to households which contain an individual from the main sample and instead look at all rural households in all waves of the IFLS data.<sup>33</sup>

The results of this estimation are seen in Table 9 and Figure 7. (Note that Figure 7 is constructed in the “opposite direction” to the previous Figures, since the point at  $t = 1$  on the graph corresponds to the coefficient on  $SPEI_{r(t-1)}$ , and intuitively indicates the persistent effect of a shock in time 0 after 1

<sup>32</sup>Since detailed household expenditure data was not available for the first wave of the IFLS data,  $k$  ranges from 2 to 5 in this specification.

<sup>33</sup>Since some households are observed in multiple waves, the specification therefore acts as a repeated cross-section, with one observation corresponding to one (household  $\times$  wave)



period has passed. The same applies to Figures 8a and 8b.) They show a strongly positive, robust, statistically significant coefficient on SPEI growing season in time  $t$ . After including fixed effects and controls, the coefficient on  $SPEI_{rt}$  in column (2) is 0.112, implying that an increase by one standard deviation of the SPEI index during the growing season of the main crop of the district will lead to an increase in average household expenditure of approximately 11%. Furthermore, using the [Benjamini and Hochberg \(1995\)](#) procedure to adjust for the false-discovery-rate leads to a q-value of 0.015, and the F-test of all the SPEI coefficients yields a p-value of less than 0.001, implying that the significant coefficient in  $t$  is highly unlikely to be a false positive. None of the other lags or leads of  $SPEI$  are statistically significant after controlling for fixed effects (column 1), implying that the effect of weather shocks over the growing season do not have persistent effects that last beyond the next growing season. As expected, there is a strongly negative and statistically significant impact of being surveyed during the growing season on household expenditure. This is likely due to seasonality in household expenditure that leads to drops in consumption just before harvest time (the “lean season”, see e.g. [Mobarak and Ali Akram, 2016](#)). Also notable is that when running the regression with no fixed effects (i.e. including only  $SPEI_{r(t-4)}$  to  $SPEI_{r(t+1)}$  and growing season  $G_{hrt}$ ) the adjusted  $R^2$  is very high at 0.639 (not shown). This implies that weather and seasonality together are the main sources of variation in household consumption. Taken together these results show that rural households in the IFLS sample are highly susceptible to weather shocks, and validates that these shocks themselves have a strong effect on household expenditure.<sup>34</sup> In turn, this starts to build up a picture of the potential mechanisms that could explain the impacts of weather shocks on adult cognitive function seen in Section 5.1. In particular, while we cannot rule out direct effects of weather on early childhood development (for example, higher rainfall leading to worsening disease environment), we now have strong evidence to suggest that weather may be affecting early childhood development *through* changes in household economic welfare.

## 6.2 Anthropometrics

In this section I examine another measure – children’s anthropometrics – in order to provide a further validation of the definition of my weather shock, and further evidence that SPEI during growing season has an impact not only on household economic welfare but on the health status of children. More specifically, I look at the effect of SPEI shocks on height-for-age and weight-for-age z-scores for children below the age of 5. I construct the z-scores by standardising each child’s height and weight measurement relative to the mean and standard deviation of their age group and sex given in World Health Organization’s child growth standards data ([WHO and UNICEF, 2009](#)). The specification used is again a variation on the main specification seen in Section 4:

$$A_{irt} = \alpha + \sum_{j=-4}^{J=1} \beta_j SPEI_{r(t+j)} + \pi G_{irt} + \mathbf{X}'_{irt} \gamma + \delta_r + \sum_{k=1}^5 \mu_k \times \mathbb{1}\{WAVE = k\}_{irt} + \varepsilon_{irt} \quad (3)$$

$A_{irt}$  denotes the z-score of either height or weight (by age and sex-group) for individual  $i$  in region  $r$  measured at time  $t$ .  $SPEI_{rt}$  denotes the average SPEI index over the most recent growing season at time  $t$  of the main crop for region  $r$ .  $G_{irt}$  is an indicator variable denoting whether individual  $i$  was measured during the growing season.  $\delta_r$  are regional fixed effects, and  $\mathbb{1}\{WAVE = k\}_{irt}$  is an indicator variable that takes the value 1 when the child was measured in IFLS wave  $k$ .  $\mathbf{X}_{irt}$  is a vector of additional controls including household-level variables (religion and sex of the household head), individual-level controls (religion, sex, and dummies for the education level of the parent). In addition, all regressions for height include a control for the method used to measure height (which was either lying down or standing up). Just as with household expenditure, I increase power by running this specification for all children aged 0-5 years old from rural household in the IFLS sample, rather than

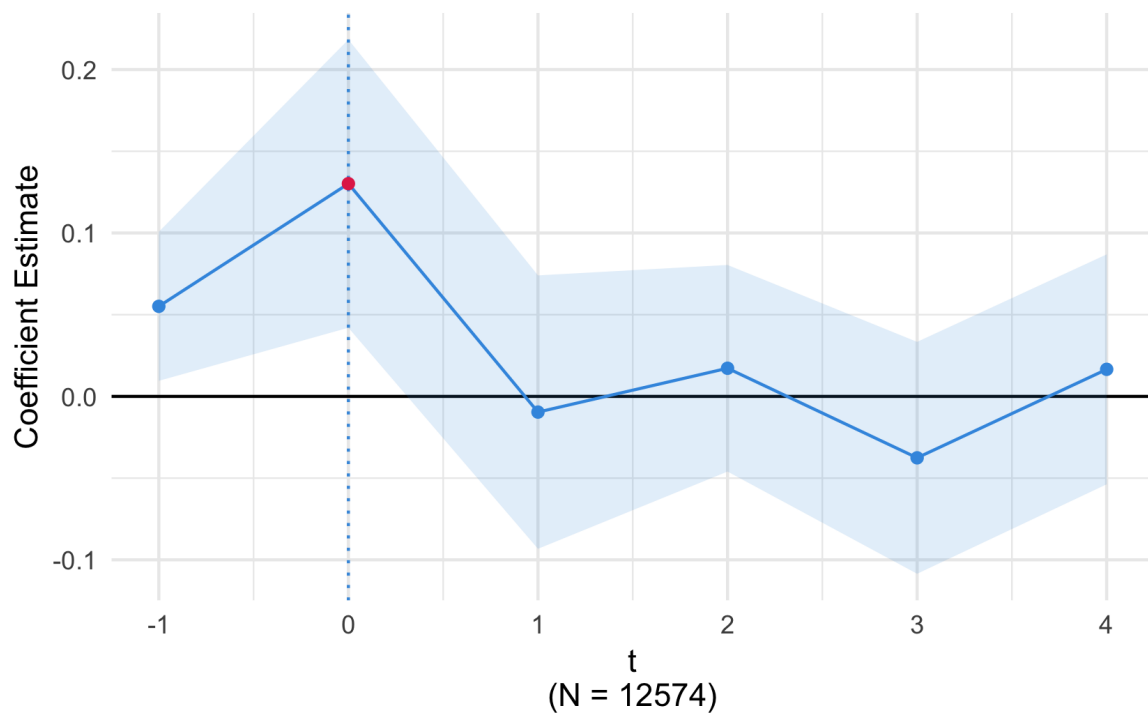
<sup>34</sup>It also demonstrates that the way I have constructed the shock variable is a reasonable proxy for the type of weather shocks that affect household welfare.

**Table 9:** Effect of SPEI shock on log household expenditure in time  $t$ 

	(1)	(2)	q value of (2)
SPEI growing season in $t - 4$	0.016 (0.036)	0.017 (0.036)	0.774
SPEI growing season in $t - 3$	-0.037 (0.036)	-0.038 (0.036)	0.599
SPEI growing season in $t - 2$	0.019 (0.033)	0.017 (0.032)	0.774
SPEI growing season in $t - 1$	-0.010 (0.043)	-0.010 (0.043)	0.822
SPEI growing season in $t$	0.134*** (0.046)	0.130*** (0.045)	0.023
SPEI growing season in $t + 1$	0.056** (0.023)	0.055** (0.023)	0.054
Surveyed during growing season (=1)	-0.154*** (0.026)	-0.154*** (0.026)	-
Wave Fixed Effects	Yes	Yes	
District Fixed Effects	Yes	Yes	
Additional Controls	No	Yes	
Sample Mean	12.561	12.560	
Adj. R <sup>2</sup>	0.674	0.675	
N	12618	12574	

Notes: \*\*\* significant at the 1% level; \*\* significant at the 5% level; \* significant at the 10% level. All models are generated according to the specification in Equation 2. One observation corresponds to one household observed in wave  $w$ . Standard errors are reported in parentheses and are two-way clustered by birth year and birth district (Cameron et al., 2011). All models use the household-level attrition-corrected weights provided in the IFLS data. Additional controls include dummies for the religion and gender of the household head. The q-value is the FDR-adjusted p-value that is calculated according to the process in Benjamini and Hochberg (1995). F-statistic and F-test are for the joint hypothesis test  $H_0 : \beta_{-4} = \dots = \beta_1 = 0$ , i.e. all the coefficients on SPEI growing season are 0.

**Figure 7:** Effect of SPEI shock on log household per capita expenditure over time



*Notes:* Each point corresponds to a coefficient from column (2) of Table 9, so represents the effect of an SPEI shock on log household expenditure as a function of time. The graph should be read as the evolution of the effect of a shock over time. So the effect in  $t$  of an SPEI shock that occurs in time  $t$  is the coefficient on 0, while the coefficient at  $t = 1$  represents the persistent effect of that same shock in  $t$  after 1 year has passed. The model controls for survey-wave fixed effects, district fixed effects, and household-level controls including religion of the household head and a dummy for a female household-head. Shaded bands are 95% confidence intervals (not adjusted for multiple testing). Points highlighted in red denote coefficients that have a q-value less than 0.05, implying that they are significantly different from 0 after adjusting for multiple testing.

restricting only to the “main sample” for whom I have psychological outcome data<sup>35</sup> (see overview of the samples in Table 2).

The results for height-for-age can be seen in Table 10, while weight-for-age is shown in Table 11. The results from the main specification (column 2) in both tables are summarised in Figures 8a and 8b. Columns (1) and (2) in Table 11 shows that once fixed effects and subsequently controls are added to the specification for weight-for-age, the coefficients on  $SPEI_{r(t-2)}$  and  $SPEI_{r(t-3)}$  are both strongly significant at the 5% level with a coefficient of 0.074. Moreover, they are robust to multiple-testing considerations, both with a q-value of 0.038, suggesting that we can reject the null if we are willing to accept a false discovery rate of 5%. While  $SPEI_{rt}$  has a significant coefficient (at the 10% level) that is negative in column (1), this reduces in magnitude, is no longer significant after controlling, and has a q-value of 0.293 after adjusting for multiple testing, suggesting that we cannot reject the null. I cannot reject that the coefficients on  $SPEI_{r(t+1)}$ ,  $SPEI_{r(t-1)}$ , and  $SPEI_{r(t-4)}$  are different from 0 at the 10% level after controlling for fixed effects. Overall this evidence seems to suggest that an SPEI shock at a certain time period  $t_0$  has a strong effect on children’s weight-for-age but that this effect takes approximately two years to show up and so can only be observed as an expected 0.07 standard deviation increase in weight-for-age at  $t_0 + 2$  and  $t_0 + 3$ . This would be consistent with a nutritional mechanism whereby a shock in  $t_0$  leads to a change in expenditure and nutrition in  $t_0$  but only affects weight after a child has had time to grow (faster or slower). The fact that the coefficient on  $SPEI_{r(t-4)}$  is smaller in magnitude (0.030) and insignificant suggests a ‘fade out’ of the effects over time, possibly due to compensatory investments made by parents or recovery from negative shocks (Lundeen et al., 2014).

Examining the evidence from Table 10 paints a less clear picture. Comparing Tables 10 and 11 shows that in general, the coefficients for height tend to be much less precisely estimated than for weight (with standard errors of 0.06 compared to 0.03). This may be a reflection of higher measurement error when measuring height than weight (for example, due to survey participants being in different positions). It could also be a result of more unobserved factors determining height-for-age, though it is unclear *a priori* why this should affect height z-scores more than weight z-scores. Partly because of this difference, examining column (1) in Table 8a shows that when district and wave fixed effects are added, no coefficients are significantly different from 0 except the surprising strong negative coefficient on  $SPEI_{rt}$ . This coefficient is significant at the 5% level and remains robust when adding controls, moving from -0.148 to -0.145. Taken at face value, this would imply that an improvement in weather conditions during the growing season of 1 standard deviation as measured by the SPEI has an average negative effect of approximately 0.15 standard deviations on height-for-age *in the same period* that the child is measured in. However, making the multiple-testing adjustment described in Section 3 and Appendix Section C shows that the q-value on this coefficient is 0.103, implying that we cannot reject the null if we are not willing to accept a False Discovery Rate of more than 5% or 10%. Given this<sup>36</sup>, and the lack of plausible theoretical explanation for the strong negative coefficient, it seems likely that such a value is actually the result of random noise in the data rather than a true causal effect.

Despite the difficulties in interpreting the results on height-for-age due to the implausible negative coefficient on  $SPEI_{rt}$  and the large standard errors, the overall patterns seen in Figures 8a and 8b are nevertheless compatible. In particular, the coefficients on  $SPEI_{r(t-2)}$  and  $SPEI_{r(t-3)}$  in the main specification are of strikingly similar magnitudes (0.6–0.9) across both height-for-age and weight-for-age, and it is chiefly because of the greater confidence intervals on height-for-age that we cannot deem

<sup>35</sup>As some children are observed across multiple waves, I effectively treat the data as a repeated cross section, with each observation corresponding to one (individual  $\times$  wave).

<sup>36</sup>I also find that the strongly negative coefficient is not robust to the exclusion of one survey wave at a time. In particular, if I exclude the data from IFLS3, the coefficient increases to -0.028 (standard error 0.068, p-value 0.679), suggesting the negative coefficient may be an artefact of random noise. By contrast, the positive coefficients on weight-for-age in  $t + 2$  and  $t + 3$  are robust to the exclusion of any single wave.

them as significant. We must interpret this evidence with due caution given such challenges, but overall it seems indicative that SPEI shocks not only affect household economic welfare as in Section 6.1, but also the health status of the children in those households. Given the wealth of evidence suggesting a link between early life growth outcomes with cognitive development (Sudfeld et al., 2015), this is an important step in explaining the mechanism through which early-life weather shocks can affect adult psychological outcomes.

**Table 10:** *Effect of SPEI shock on height-for-age for 0-5 year olds in time t*

	(1)	(2)	q value of (2)
SPEI growing season in t - 4	0.092 (0.057)	0.084 (0.056)	0.261
SPEI growing season in t - 3	0.054 (0.054)	0.066 (0.053)	0.313
SPEI growing season in t - 2	0.089 (0.057)	0.090 (0.058)	0.261
SPEI growing season in t - 1	0.047 (0.063)	0.048 (0.064)	0.451
SPEI growing season in t	-0.148** (0.061)	-0.145** (0.061)	0.103
SPEI growing season in t + 1	0.048 (0.049)	0.054 (0.048)	0.313
Surveyed during growing season (=1)	0.064 (0.049)	0.063 (0.049)	-
Wave Fixed Effects	Yes	Yes	
District Fixed Effects	Yes	Yes	
Additional Controls	No	Yes	
Sample Mean	-1.659	-1.662	
Adj. R <sup>2</sup>	0.044	0.053	
N	8979	8862	

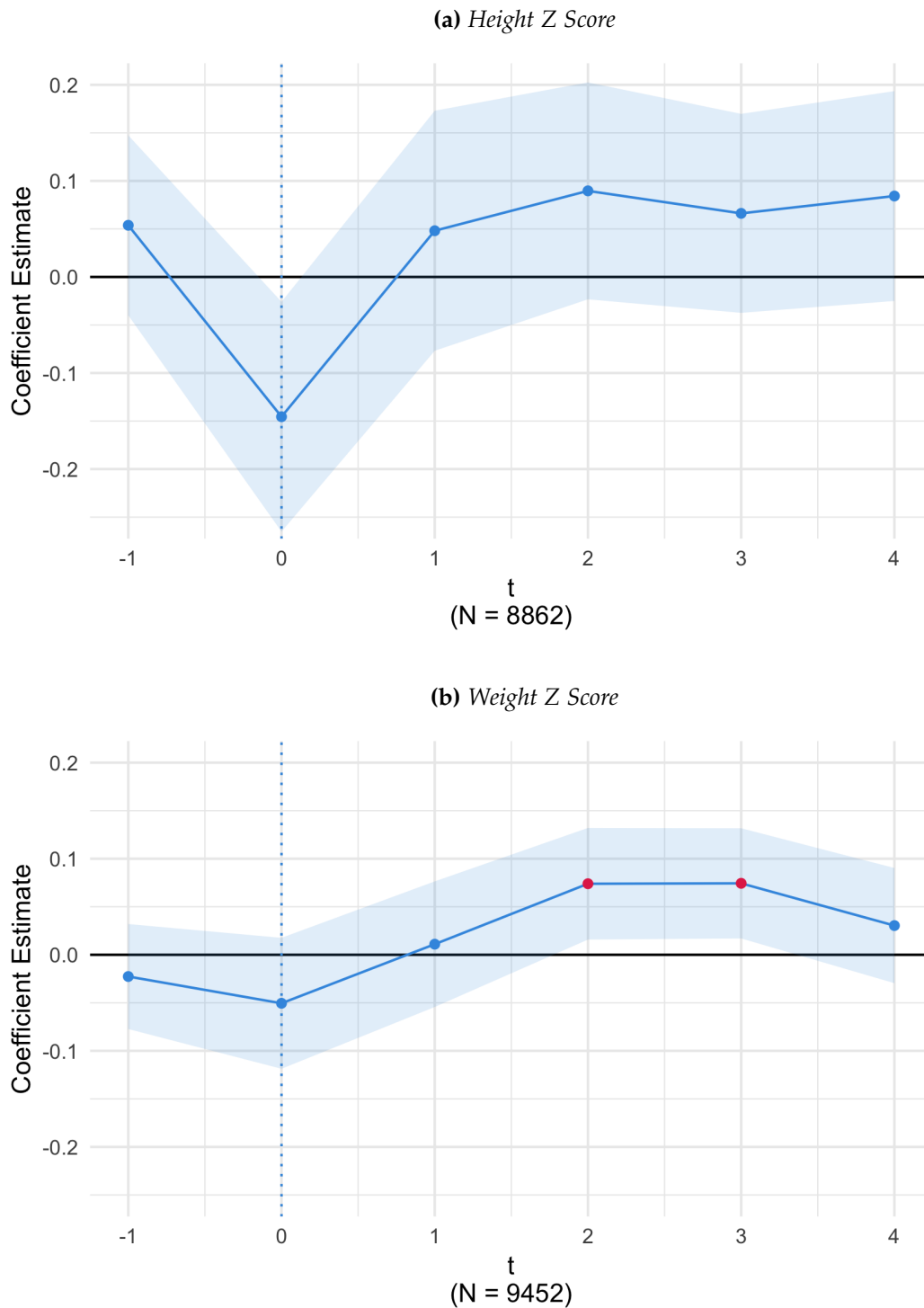
Notes: \*\*\* significant at the 1% level; \*\* significant at the 5% level; \* significant at the 10% level. Standard errors are reported in parentheses and are two-way clustered by birth year and birth district (Cameron et al., 2011). All models are generated according to the specification in Equation 3. Additional controls include individual-level controls (religion, sex, and dummies for the education level of the parent). All regressions include dummies for the method of height measurement as controls. All models use the individual-level attrition-corrected weights provided in the IFLS data. The q-value is the FDR-adjusted p-value that is calculated according to the process in Benjamini and Hochberg (1995). F-statistic and F-test are for the joint hypothesis test  $H_0 : \beta_{-4} = \dots = \beta_1 = 0$ , i.e. all the coefficients on SPEI growing season are 0.

### 6.3 Mediation Analysis

The results presented in Sections 6.1 and 6.2 provide indicative evidence that there are economic and nutritional mechanisms that could explain the impact of early-life weather shocks on adult cognitive function. This evidence is incomplete, however: it only demonstrates the impact of the shock on a mediator, rather than completing the causal chain by showing the subsequent impact of the mediator on cognitive function. An extensive literature shows an empirical correlation between household expenditure, child anthropometrics, and cognitive development, suggesting that such an impact may indeed exist<sup>37</sup> (Grantham-McGregor et al., 2007; Adair et al., 2013; Alderman et al., 2006; Crookston et al., 2011; Victora et al., 2008), but this literature is typically unable to examine the full causal chain from shock to mediator to outcome. This is what I aim to examine in the current section.

<sup>37</sup>On the other hand, there is an open debate about the extent to which weaker anthropometric growth (such as stunting) is truly a *cause* of worse cognitive development, or just an indicator of other conditions that themselves cause worse cognitive development (Leroy and Frongillo, 2019)

**Figure 8:** Effect of SPEI shock over time on anthropometric outcomes



Notes: Coefficients correspond to model (3) in Tables 10 and 11 respectively. Each point corresponds to a coefficient from column (2) of Tables 10 and 11 respectively, and therefore represents the effect of an SPEI shock on children’s anthropometrics as a function of time. The graph should be read as the evolution of the effect of a shock over time. So the effect in  $t$  of an SPEI shock that occurs in time  $t$  is the coefficient on 0, while the coefficient at  $t = 1$  represents the effect of the same shock in  $t$  when the child’s height and weight are observed a year later. The model controls for survey-wave fixed effects, district fixed effects, and individual-level controls including sex, religion, and parental education. Shaded bands are 95% confidence intervals (not adjusted for multiple testing). Points highlighted in red denote coefficients that have a q-value less than 0.05, implying that they are significantly different from 0 after adjusting for multiple testing.

**Table 11:** Effect of SPEI shock on weight-for-age for 0-5 year olds in time  $t$ 

	(1)	(2)	q value of (2)
SPEI growing season in $t - 4$	0.033 (0.030)	0.030 (0.031)	0.481
SPEI growing season in $t - 3$	0.066** (0.029)	0.074** (0.029)	0.038
SPEI growing season in $t - 2$	0.072** (0.030)	0.074** (0.030)	0.038
SPEI growing season in $t - 1$	0.010 (0.033)	0.011 (0.033)	0.741
SPEI growing season in $t$	-0.063* (0.035)	-0.051 (0.035)	0.293
SPEI growing season in $t + 1$	-0.025 (0.028)	-0.023 (0.028)	0.501
Surveyed during growing season (=1)	0.017 (0.028)	0.018 (0.028)	-
Wave Fixed Effects	Yes	Yes	
District Fixed Effects	Yes	Yes	
Additional Controls	No	Yes	
Sample Mean	-1.016	-1.015	
Adj. $R^2$	0.034	0.043	
N	9578	9452	

Notes: \*\*\* significant at the 1% level; \*\* significant at the 5% level; \* significant at the 10% level. Standard errors are reported in parentheses and are two-way clustered by birth year and birth district (Cameron et al., 2011). All models are generated according to the specification in Equation 2. Additional controls include individual-level controls (religion, sex, and dummies for the education level of the parent). All models use the individual-level attrition-corrected weights provided in the IFLS data. The q-value is the FDR-adjusted p-value that is calculated according to the process in Benjamini and Hochberg (1995). F-statistic and F-test are for the joint hypothesis test  $H_0 : \beta_{-4} = \dots = \beta_1 = 0$ , i.e. all the coefficients on SPEI growing season are 0.

### 6.3.1 Methodology

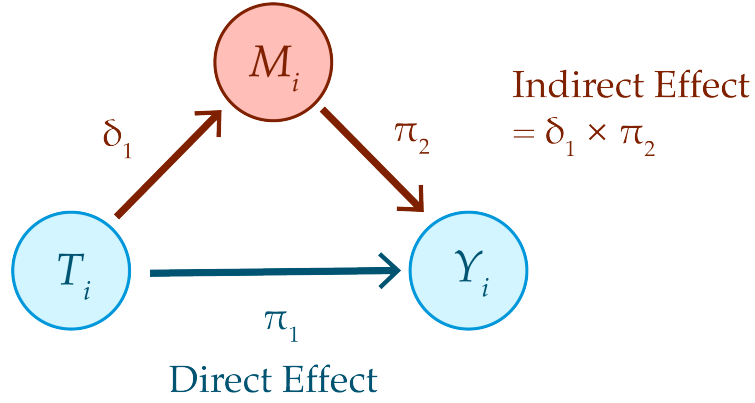
In order to provide the seeds of an answer to this question, I run a mediation analysis that examines the key mechanisms that may act as a channel for the effect of an SPEI shock on cognitive function in adulthood. For this exercise I rely on a basic linear mediation analysis framework that is outlined clearly in VanderWeele (2016).<sup>38</sup> The aim of the linear framework is to be able to distinguish between the *direct effect* (the effect of the shock that is not mediated by a mediator  $M_i$ ) and the *indirect effect* (the effect of the shock that is caused by a change in  $M_i$ ).<sup>39</sup> These ideas are summarised in Figure 9, which shows the theorised (linear) causal relationships between the shock  $T_i$ , the mediator  $M_i$ , and the outcome  $Y_i$ .

Under assumptions outlined below, these can be calculated by estimating the following system of

<sup>38</sup>I initially used a similar framework seen in Huber (2014) that relies on inverse probability weights. However, this framework turned out to be unsuitable in my empirical context. The sample size issue described below meant that I was not able to run the analysis using all mediators in one regression without losing around 80% of my sample. This meant the calculation of the inverse probability weights was erratic and did not lead to reliable estimates of the relevant coefficients. Given these issues, I use a simpler framework that does not use inverse probability weights.

<sup>39</sup>Note that the linear framework does not allow me to tackle the important issue of whether there is complementarity between different inputs in the human capital production. This would require a fully calibrated human capital production function model: this is left for future research.

**Figure 9:** Simple mediation analysis framework



Notes: Schematic diagram depicting the causal relationships between the treatment (shock) variable  $T_i$ , a mediator  $M_i$ , and the outcome variable  $Y_i$ . Each arrow represents a causal relationship between two variables, and the coefficient from equations 6, 4 and 5 that determines that relationship is marked besides.

equations:

$$Y_i = \alpha + \beta T_i + \mathbf{Z}'_i \lambda + \varepsilon_i \quad (4)$$

$$M_i = \delta_0 + \delta_1 T_i + \mathbf{Z}'_i \mu + \eta_i \quad (5)$$

$$Y_i = \pi_0 + \pi_1 T_i + \pi_2 M_i + \mathbf{Z}'_i \gamma + v_i \quad (6)$$

$Y_i$  denotes the outcome variable, which in this case will be the adult cognitive factor score seen in the results in Section 5.1. The proposed mediator is denoted  $M_i$ . The standard mediation analysis framework assumes a binary (rather than continuous) treatment variable denoted  $T_i$ . I define the binary shock variable for an individual  $i$  born in year  $t$  as:

$$T_i \equiv T_{it} = \mathbb{1} \left\{ \sum_{j=-1}^4 SPEI_{i(t+j)} \geq 0 \right\}$$

where  $SPEI_{i(t+j)}$  is defined in the same way as in the main specification. In other words, I take the average SPEI during the growing seasons from the entire early life period (the year before birth to 4 years after birth). I then transform this average into a binary shock variable by setting it to 1 if it is equal to or above 0, and 0 otherwise. In each of the equations I control for birth-year and birth-district fixed effects along with individual level controls (gender and religion). Together these are denoted as the vector  $\mathbf{Z}_i$ .

Intuitively, equation 4 calculates the overall effect of the shock  $T_i$  on the outcome  $Y_i$  (similar to the main specification earlier in the paper), and thus  $\beta$  identifies the *total effect* of the shock on the outcome. Equation 5 identifies the effect of the shock on the mediator ( $\delta_1$ ), and equation 6 identifies the effect of the mediator  $M_i$  on the outcome  $Y_i$  conditional on treatment ( $\pi_2$ ). Using these together it is clear that taking the product of the effect of the shock on the mediator and the effect of the mediator on the outcome will give us the *indirect effect* of the shock. In other words,  $(\delta_1 \times \pi_2)$  identifies the magnitude of the effect of  $T_i$  on  $Y_i$  that is mediated by  $M_i$ .

The identification statements made in the previous paragraph rely on a number of stringent assumptions, described in detail in [Huber \(2014\)](#):

- (i) *Conditional independence of the treatment variable  $T_i$ .*
- (ii) *Conditional independence of the mediator variable  $M_i$ .*
- (iii) *Common support assumption.* The conditional probability to be treated given  $M_i$  and  $\mathbf{Z}_i$  is greater than 0 in both treatment states.



Assumption (i) allows us to calculate the effect of the treatment variable  $T_i$  on both  $Y_i$  and  $M_i$ . This is the same assumption as the main identifying assumption of the paper, namely that conditional on birth-year and birth-district fixed effects, the effect of SPEI growing season shocks on cognitive development is as good as random and thus unconfounded. Assumption (iii) will hold if the covariates used  $Z_i$  do not perfectly predict treatment status. Assumption (ii) is less likely to hold in practice. It states that the mediator variable is exogenous conditional on all covariates. We may doubt this: for example, if unobserved factors such as motivation are correlated with both cognitive factor scores and parental education level, then this would be confounded with the true effect of  $M_i$  on  $Y_i$  and thus lead to a biased estimate of  $\pi_2$ .

The assumptions this framework rely on are thus highly restrictive, but in exchange for making them I am able to estimate the magnitudes of the indirect effects of a number of different mediators and begin to understand the full causal chain of the mechanisms through which early childhood weather shocks could have on cognitive outcomes in adulthood. Given the fact that mediators are likely to be endogenous and thus assumption (ii) is likely to fail, we should interpret the results with extreme caution, providing only suggestive evidence of potential mechanisms rather than robust causal estimates.

In fact, rather than there being one candidate mediator variable  $M_i$ , as in the simplified framework presented above, there are in fact multiple candidate mediators (denote them  $\{M_i^1, M_i^2, \dots, M_i^m\}$ ) that each provide a possible channel for an SPEI shock in early childhood to have an impact on adult cognitive outcomes. Given the wealth of measures available in the IFLS data, I narrow my search using as a guide Heckman's theoretical framework of the human capital production function described in Section 2 (Heckman and Cunha, 2007; Cunha et al., 2010). In this framework, the key determinants of an adult's stock of skills are the child's initial endowment of skills, parental investments made during childhood, and parental human capital. In addition, strong cross-links between cognitive development and health capital may also exist (Heckman, 2007). I therefore use the following variables as candidate mediators: (i) child health, including anthropometrics and child acute health, (ii) measures of parental health, (iii) measures of early investments such as investments in health and nutrition, (iv) measures of parental time investments using time in employment as a proxy for the (inverse) of time spent with children, and (v) parental education level. The variables used in the mediation analysis are the factor scores generated by a latent factor model similar to the one used to generate the cognitive and socioemotional factor scores. This process is described in Appendix Section B. The survey questions used to generate these factors are also described in Appendix Section A.1.

There are at least four major empirical challenges to running the exercise outlined above with the chosen mediators. First, many of the variables that could act as relevant mediators are likely to be subject to significant measurement error. For example, some measures of early investment, such as whether a health care provider was present during the birth of a child, may be subject to recall bias if mothers are asked to recall many years after pregnancy. Second, there are some mediators that are imperfectly measured by a number of proxies. For example, there is no single perfect measure of early investment, although coarse measures such as breastfeeding, expenditure on education, smoking behaviour, and nutrition may capture some important variation. To deal with the first two considerations, I use the same type of latent factor model used to generate the cognitive and socioemotional factor scores. This procedure is based on Cunha et al. (2010) and is described fully in Appendix Section B.

A third challenge I face in using the IFLS data relates to its panel data structure.<sup>40</sup> The survey has 5 waves, different children are observed at different ages in each wave, and children may be observed once or multiple times across waves. The data on different variables for each individual child therefore

---

<sup>40</sup>Many papers look at the impact of an intervention targeting all children of a single age (e.g. Heckman et al., 2013 or Attanasio et al., 2015), or use a single cross-section of data (e.g. Adhvaryu and Fenske, 2016). They therefore either measure all children repeatedly at the same age, or only measure each child once. This obviates the concern outlined above.

varies dramatically. I deal with this problem by age-standardising each candidate mediator measure, taking the z-score based on the mean and standard deviation for that age, and taking the mean z-score over all waves in the cases where children are observed more than once. The process is described in more detail in Appendix Section B. These average z-scores are then used as measures in the latent factor model. The above steps create factor ‘scores’ for each individual that denote the (normalised) level of that factor. For example, a factor score of 1 for “early investment” indicates that a child received 1 standard deviation higher level of early investment than the sample average. These scores are then used as the mediator  $M_i$  variable in the specification outlined above.

Fourth, the fact that different children are observed at different stages in each wave means that there is a high proportion of missing values for the resulting mediator factor scores. Out of the 5903 individuals in the “main sample”, there are only between 3242 and 5528 values available for each mediator (see column “N” in Table 12 for details). This is largely an artefact of the IFLS survey design. For example, measures included in “Early Investment” such as whether the child was breastfed were only available if the mother of the child was interviewed an IFLS household in at least one wave. Moreover, the probability of having a missing value is largely uncorrelated across mediators, so that there are very few observations with non-missing values for *all* mediators. To avoid a drastic reduction in sample size (and thus statistical power), rather than run a mediation analysis with all mediators in one framework, I run a separate mediation analyses for each variable individually. For this to generate unbiased estimates of the mediator contribution, I need to make two additional assumptions alongside (i), (ii) and (iii) above:

- (iv) “Missing completely at random” assumption i.e. whether an individual has a missing value for a mediator is not correlated with any other observed or unobserved factor such as the individual’s potential outcomes, mediator variables, explanatory variables, or covariates  $Z_i$ .
- (v) Zero correlation between each of the mediators  $\{M_i^1, M_i^2, \dots, M_i^m\}$ .

Both of these are exceptionally strong assumptions, but in exchange for making these assumptions I can use the complete set of values for each mediator in estimating the indirect effect on the outcome variable (cognitive factor score). If assumption (v) is not verified, say because  $M_i^1$  and  $M_i^2$  are correlated, this means that the estimated coefficient on  $M_i^1$  will capture some of the effect of  $M_i^2$  (and vice versa) and thus be a biased from the population value. Table 25 shows descriptive evidence that most pairs of mediators have Pearson correlation coefficients below 0.1. For those that don’t, we will need to exercise caution in interpretation. For example, since the correlation between on Early Investment and Parental Education is high (0.398), the result of a strong indirect effect of Early Investment could actually be due to the effects of Parental Education, or vice versa.

### 6.3.2 Results

Table 12 shows the results of carrying out the mediation analysis exercise described in Section 6.3.1. For each mediator variable, I run the system of equations 4, 5 and 6 separately. The coefficient in the column “shock effect on mediator” is the estimate of  $\delta_1$ , i.e. the estimated effect of  $T_i$  on  $M_i$ . The coefficient in the second column, “Mediator Effect on Cognitive Factor”, is the estimate of  $\pi_2$ . Under the assumption that the mediator is conditionally exogenous (assumption (ii)), this will identify a causal effect of  $M_i$  on  $Y_i$ . The indirect effect is calculated by multiplying  $\delta_1$  in column (1) and  $\pi_2$  in column (2), as seen in Figure 9. The implied % contribution is then calculated by dividing the indirect effect by the *total effect*  $\beta$ . The total effect is taken from an estimation of the equation 4 which yields a coefficient of 0.0989 (standard error 0.0363, p-value 0.0064, N = 5770),<sup>41</sup> implying that average SPEI growing season during early life (in utero to 4 years old) being above 0 leads to an average increase in adult cognitive score of approximately 0.1 standard deviations.

<sup>41</sup>Note that this is different to the results of the main specification in Table 4 because of the different (binary) definition of the shock.

Figure 10 summarises the results seen in Table 12 in a readable fashion. Under the five assumptions outlined above (including, importantly, the conditional exogeneity of the mediator), we can interpret the results as follows. First, notice that none of the indirect effects can be rejected to be different from 0 at the 10% level. This could be due to the reduced sample size and demanding data requirements for the mediation analysis exercise, which would lead to a reduction in the statistical power to reject false null hypotheses. This seems likely to at least partly explain the lack of significant results, given the large standard errors on all the estimates. Alternatively, it could be that the majority of the effect of an *SPEI* shock in the early childhood years is direct, or at least is not mediated by the factors included in the analysis here. This is also possible, since there are many unmeasured factors (such as a detailed measure of time spent with children) that are likely to be affected by an *SPEI* shock and also affect a child's cognitive development.

Despite this caveat, some tentative explanations can be drawn from the results seen in Table 12 and Figure 10. The anthropometrics factor (composed of height and weight z-score) appears to contribute positively to cognitive development, with a coefficient of  $\pi_2^{\text{anthro}} = 0.054$  that is significant at the 1% level. If interpreted causally, this would imply that an increase in anthropometrics factor score of 1 standard deviation is associated with an average increase of 0.05 standard deviations in adult cognitive score. This is a somewhat small coefficient compared to those seen in previous literature (e.g. [Sudfeld et al., 2015](#)). Although in this restricted sample, the effect of the shock on the anthropometric factor ( $\delta_1^{\text{anthro}}$ ) is insignificant, it is large and positive with a coefficient of 0.079. Moreover, the results seen in Section 6.2 (which use a more extensive sample of 0-5 year olds) would suggest that *SPEI* shocks have a strong positive impact on weight-for-age. Piecing together these results, we can plausibly make the case that childhood anthropometry is a causal mediator for the impact of early childhood weather shocks on adult cognitive ability, and that the estimated percentage contribution of 4.3% is likely to represent a true positive indirect effect rather than just random noise. This conclusion in turn suggests that childhood nutrition and health will be important channels for such a causal impact. However, the results on the Child Acute Health factor (composed of measures of acute morbidity and frequency of health care use) provides little evidence to support this view, as it appears that weather shocks have no significant impact on this factor.

Table 12 suggests that parental education level is not a mediator when estimating using this framework.<sup>42</sup> Weather shocks during a child's early years do little to affect parental education level (the coefficient is a precisely estimated 0.04). This is unsurprising, given that the vast majority of a parent's education will be undertaken when they themselves are young.

Intriguingly, there seems to be some evidence suggesting a compensatory response to weather shocks in early life. Both the early investment factor (composed using measures of early food security, educational expenditure, and breastfeeding behaviour) and the mother's employment factor have significantly positive effects on child's cognitive factor scores, with coefficients of 0.136 and 0.023 respectively. But mother's employment responds negatively to early-life *SPEI* (-0.103) and so does early investment (-0.064, although this is not estimated with enough precision to be significant at the 10% level). If the estimates of  $\pi_2$  are truly causal, this would imply that mothers respond to better weather by reducing the amount of time worked (presumably thereby spending more time at home). Plausibly this leads to a reduction in household income, which then impairs cognitive development of children through nutritional or material investment channels. Along similar lines, a possible reduction in early investment as a result of the shock, which may reflect weaker diet diversity (less vitamin A and iron), reduction in breastfeeding, or less educational expenditure, also compensates for the overall positive effects of the shock. Parental health also represents a possibly compensating channel for the effect of the shock on cognitive outcomes. Parental health responds negatively to the shock,

<sup>42</sup>To some extent, however, this result may bely one of the weaknesses using a linear mediation analysis. In Heckman's framework, early life shocks do not *change* parental education level, but parental human capital nevertheless plays an important role in mediating cognitive and socio-emotional development through its interactions with other inputs to the human capital production function. Such complexities cannot be captured in my estimation.

and given the positive correlation between parental health and child's cognitive factor score, these leads to an (insignificant) negative indirect effect of the shock through parental health. This result could be rationalised by postulating non-economic mechanisms. For example, a change in the disease environment in the presence of higher rainfall<sup>43</sup> could worsen adult health, or alternatively positive weather shocks could increase the amount of time spent in agriculture, which in turn negatively affects health.

As touched upon above, there are a number of challenges with interpreting the evidence presented in Table 12 and Figure 10 at face value. First, the assumptions required to interpret the coefficients as causal effects (particularly assumption (ii), the conditional exogeneity of the mediator) are unlikely to hold. Nevertheless, the results give an indication as to which mediators may be the most important in our context, and pave the way for mediation analysis that makes use of better identification strategies (such as using plausibly exogenous instruments for investments). Second, the severe restrictions on sample size mean that standard errors are very large, meaning that I have low statistical power to reject the null. This is likely to explain at least partly why no candidate mediators made a statistically significant contribution to the effect of the shock on adult cognitive score. Third, the fact that the framework used is a linear mediation analysis framework may neglect empirically important effects. In Heckman's human capital production function model, the technology of cognitive skill formation is highly non-linear and involves potential complementarities between different inputs such as investments. If this framework is correct, then my mediation analysis will fail to capture such complementarities, potentially biasing estimates.<sup>44</sup>

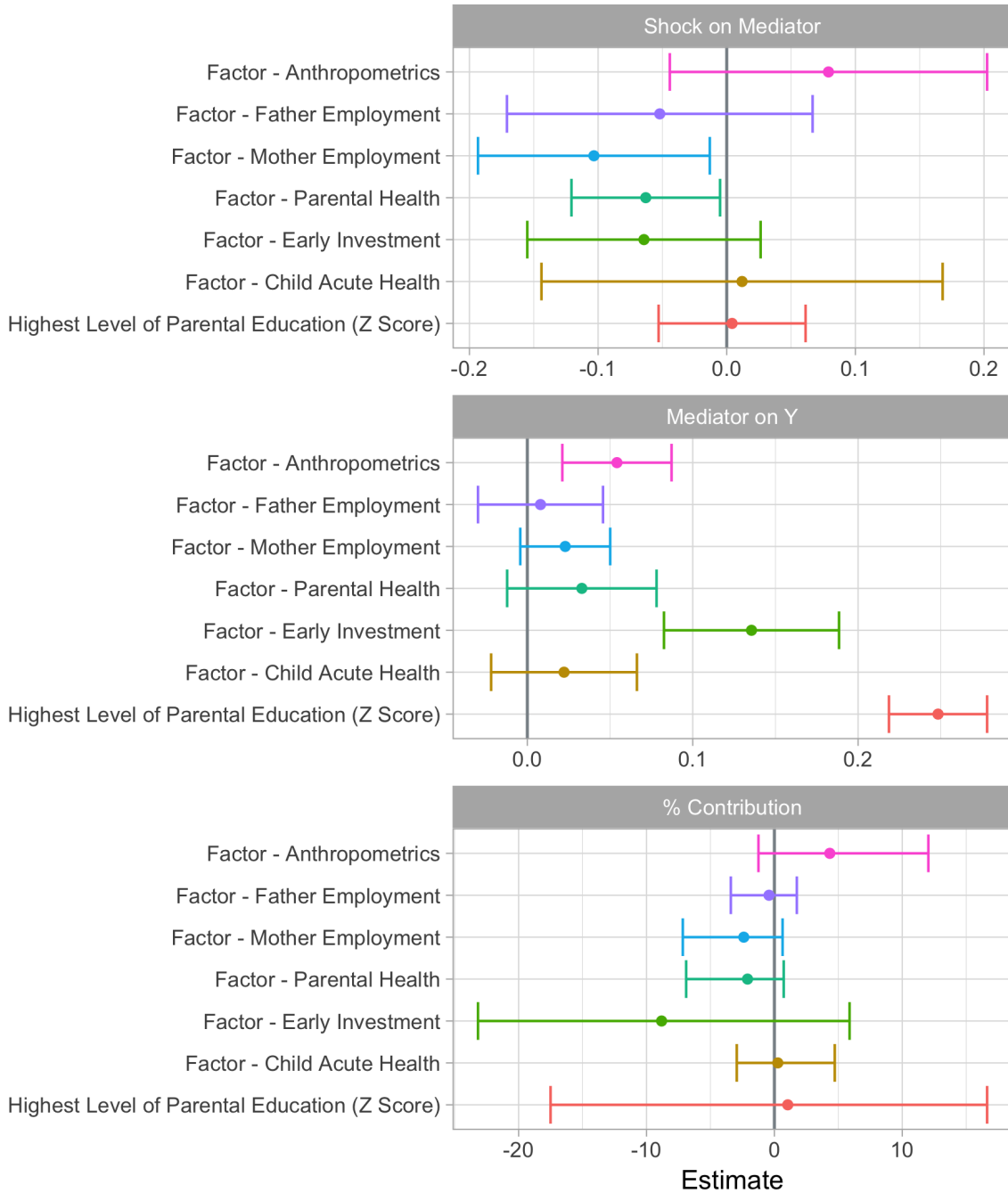
These challenges at least partially explain the mixed, and sometimes inconsistent, picture of the mechanisms through which early life weather shocks can affect adult cognitive outcomes. Despite this, some tentative insights can be gathered from the results. In particular, it seems that there may be important compensatory responses that could dampen the effects of shocks, where early investments and mother's employment act in the opposite direction to shocks, and parental health also responds negatively to shocks. The extent to which these truly reduce the effect of shocks is largely unclear, however. I also find some evidence that positive effects on anthropometric outcomes may be an important mediating channel (or at least an indicator of an important mediating channel) through which weather shocks affect cognitive development.

---

<sup>43</sup>For example, malaria transmission is affected by precipitation patterns because malaria-carrying mosquitoes depend on standing water for breeding (Bomblies, 2012)

<sup>44</sup>Moreover, neglecting complementarities means that I am unable to answer interesting and policy-relevant questions. In particular, I cannot test for dynamic complementarity for investments, neither for cross-productivity or self-productivity in skills (see Cunha et al., 2010 for more details). I leave this for future research in which I will follow Attanasio et al. (2017) in estimating a fully calibrated human capital production function model.

**Figure 10: Mediation Analysis**



Notes: All coefficients are taken from Table 12. See notes on Table 12 for description of methodology.

**Table 12:** Mediation Analysis for the effect of early life SPEI shock on adult cognitive factor score

Mediator	N	Shock Effect on Mediator	Mediator Effect on Cognitive Factor	Indirect Effect	Implied % Contribution
Highest Level of Parental Education (Z Score)	5528	0.004 (0.029)	0.248 (0.015)***	0.001 [0.009]	1.046 [8.753]
Factor - Child Acute Health	3275	0.012 (0.080)	0.022 (0.022)	0.000 [0.002]	0.267 [1.750]
Factor - Anthropometrics	3258	0.079 (0.063)	0.054 (0.017)***	0.004 [0.003]	4.334 [3.445]
Factor - Early Investment	3242	-0.064 (0.046)	0.136 (0.027)***	-0.009 [0.007]	-8.820 [7.117]
Factor - Parental Health	3887	-0.063 (0.029)**	0.033 (0.023)	-0.002 [0.002]	-2.098 [1.930]
Factor - Mother Employment	4416	-0.103 (0.046)**	0.023 (0.014)*	-0.002 [0.002]	-2.393 [1.999]
Factor - Father Employment	4416	-0.052 (0.061)	0.008 (0.019)	0.000 [0.001]	-0.419 [1.251]

Notes: Shock is defined as a binary variable that takes the value 1 when the average SPEI over the growing seasons in the years  $t - 1$  to  $t + 4$  where  $t$  is the year of birth, and 0 otherwise. Each line is the result of a separate estimation of the system of equations 4, 5 and 6. “Shock Effect on Mediator” is the estimate of  $\delta_1$  in equation 5. “Mediator Effect on Cognitive Factor is the estimate of  $\pi_2$  in equation 6. Indirect effect results from multiplying these two quantities together i.e.  $\delta_1 \times \pi_2$ . Implied % contribution gives the indirect effect as a percentage of the total effect  $\beta$ . Total effect is calculated using equation 4, which yields a coefficient of  $\beta = 0.0989$  (standard error 0.0363, p-value 0.0064, N = 5770). Statistics in round parentheses denote standard errors directly gathered from model estimations. Statistics in square brackets denote bootstrapped standard errors that are calculated over 1000 bootstrapped samples. All models use the individual-level attrition-corrected weights provided in the IFLS data.

## 7 Discussion

In this paper I present new and compelling evidence for the existence of a critical period for cognitive development in early childhood, and I am able to isolate this period to the third year after birth (age 2). More specifically, I show that weather shocks at this age on individuals born in rural Indonesia between 1988 and 2000 have a long lasting effect on their cognitive ability as an adult, whereas weather shocks at other ages have no such impact. This effect appears to be strongest among boys. When examining similar evidence on adult socioemotional outcomes, I find no such evidence of a critical period for these weather shocks. I find no evidence of long-term impacts of early life weather shocks on adult personality and mental health. On the other hand, early life weather shocks do appear to affect adult affect/wellbeing for girls. Since these effects occur at various stages of childhood, rather than being concentrated impacts with a narrow window of timing, I interpret this as evidence against a critical period for socioemotional development in this context. The lack of critical period may be explained by compensating investments in early childhood or the increased malleability of socioemotional development into later childhood.

I further show evidence that the observed effects on cognitive development are consistent with an economic channel in which weather affects agricultural income and then infant health outcomes (as proxied by anthropometrics), leading to impacts on cognitive development that persist into adult life.

The impacts I describe are “reduced-form” impacts of early-life weather shocks, in the sense that they incorporate the effects of any subsequent investment responses by parents. In order to understand how much of the reduced-form impacts are explained by direct effects as compared to investment responses, I run a linear mediation analysis that gives no clear evidence as to the primary mechanisms

for the effect of the shock, suggesting that the effect of early-life weather shocks may be direct or not mediated by factors measured in the available data. However, the evidence is possibly consistent with compensatory parental investment responses, which would imply that the reduced-form impacts may even underestimate the direct impact of weather shocks.

These results have important potential policy implications. In particular, the existence of a critical period at age 2 for cognitive development, but the absence of such a critical period for socio-emotional development, can inform the optimal timing of early-childhood interventions. To the extent that the timing of effects of weather shocks mirror the timing of effects of such interventions, the results suggest that interventions aiming to improve cognitive development should optimally be targeted at 2-year-olds. Moreover, my results imply that targeting at this age may be even more effective than interventions targeted at 0- or 1-year-olds, contrary to previous literature suggesting that the entire first 1000 day period is “critical”. On the other hand, my results suggest no such critical window for socioemotional development, which is consistent with Heckman and coauthors’ claim that socioemotional programmes aiming to improve factors like personality, motivation, and mental health, need not prioritise the first 1000 days in the same way.

# Appendix

## A Data Appendix

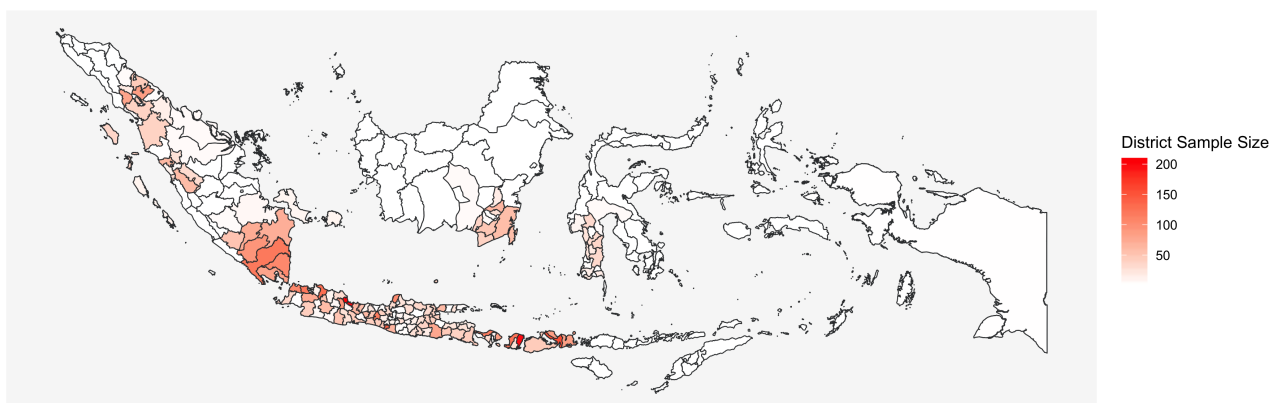
### A.1 IFLS Data

A summary of which outcome variables were available in each of the waves of IFLS (and the years of each survey) is found in Figure 11. Figure 12 shows the location of my sample, which is concentrated in the 13 provinces covered by the IFLS data. The largest concentration of my sample is located in Western, Central, and Eastern Java.

**Figure 11:** Summary of data availability for main outcome variables

IFLS <sub>1</sub> 1993	IFLS <sub>2</sub> 1997	IFLS <sub>3</sub> 2000	IFLS <sub>4</sub> 2007	IFLS <sub>5</sub> 2015
HH Expenditure	HH Expenditure	HH Expenditure	HH Expenditure	HH Expenditure
Anthro	Anthro	Anthro	Anthro	Anthro
		Cognitive (limited)	Cognitive	Cognitive (full)
			Socioemotional (limited)	Socioemotional (full)

**Figure 12:** Birth location of individuals in main sample



**Cognitive Measures.** All IFLS respondents above the age of 15 are asked a number of questions that are designed to test cognitive ability:

1. *Raven's Progressive Matrices* (John and Raven, 2003) - participants have to fill in the gaps of a visual matrix that follows a certain pattern. This is a common metric of "fluid intelligence".
2. *Maths - Written* - participants have to answer 5 simple written maths questions.



3. *Word Recall* - respondents were given a list of 10 words, and then asked to recall as many of the words as possible. They were asked to do this twice, once immediately after hearing the list, and another time 12-15 minutes later. (Score out of 10)
4. *Maths - Oral* (IFLS5 only) - respondents were asked to serially subtract 7s from 100 (Score out of 10)
5. *Number Series* (IFLS5 only) - participants are given a series of 3 numbers and are asked to fill in the fourth number in the series, e.g. "2 4 6 ?". The IFLS data then uses a Rasch scoring model to create a Woodcock-Johnson score based on a participants ability to answer questions of varying difficult.

**Socioemotional Measures** - the full set of socioemotional measures is found in Table 13. Mental health measures are based on the CES-D scale. Personality measures are based on the Big 5 personality test.

**Household expenditure.** IFLS data contains detailed household expenditure data for each household. The categories asked are related to food, self-production, non-food, education, and rent (along with imputed rent). I impute the respondents' estimate of rent for home-owners. Total monthly expenditure is calculated by summing across all available categories. Total per capita monthly expenditure is calculated by dividing this by the reported household size.

**Anthropometrics.** IFLS respondents (including children) had their height and weight measured. To calculate height-for-age and weight-for-age z-scores for children under 5 years old, I combine these measurements with WHO's child growth standards data (WHO and UNICEF, 2009). Following WHO recommendations, I remove all outliers with z-scores outside of the range  $[-5,5]$ .

**Birth location and migration history.** IFLS contains data on all adults' birth district and migration history after the age of 12. I combine the data on district of birth, the district when surveyed, and migration history to construct a year-by-year history of the district each respondent has lived in since birth. This is subsequently used to match individuals to the SPEI values in each year of their life. The main specification in Section 4 also makes use of district location in the two years before birth. For these two years, individuals are matched to the location of their birth.

**Birth date.** The IFLS data contains a "best-guess" birth date for each individual that incorporates information across multiple waves in cases of inconsistencies.

**Control variables.** As controls I make use of data on the gender and religion of each individual. For individuals I construct a parental education variable, either by matching individuals back to their parents if they are also IFLS respondents, or by using the self-reported data on parental education. The variable used in the paper is the highest level of education reached by either of the parents. The possible levels are: no education, elementary school, junior high school, senior high school, or university. I also have data on the identity of the household head which I use to construct the household head education variable in a similar way.

**Mediation variables.** Descriptions of all the measures taken from the IFLS survey to use in the mediation analysis are found in Table 14.

## A.2 District, Weather, and Crop Data

**District data.** District boundary data is taken from IPUMS International (Minnesota Population Center., 2018). This contains the shapefiles for the first-level regional boundaries (*province*) and the second-level district boundaries (*kabupaten*) used in the IFLS. There were multiple district boundary changes and splits (see e.g. Bazzi and Gudgeon, 2016) between the first and last waves of the IFLS. Since the IPUMS data is harmonized so that the geographical unit of analysis stays consistent over time,

**Table 13: Description of all socioemotional variables from IFLS**

Module	Question Name	Pre question	Full Question	Reverse coded?	Included in IFLS4	Answer Scale
Mental Health	Bothered	In the past week...	Bothered by things that usually don't bother me	Yes	Yes	1 to 4 ("Rarely" to "Most of the time")
Mental Health	Concentration	In the past week...	I had trouble concentrating in what I was doing	Yes	Yes	1 to 4 ("Rarely" to "Most of the time")
Mental Health	Depressed	In the past week...	I felt depressed	Yes	Yes	1 to 4 ("Rarely" to "Most of the time")
Mental Health	Everything was an effort	In the past week...	I felt everything I did was an effort	Yes	Yes	1 to 4 ("Rarely" to "Most of the time")
Mental Health	Hopeful	In the past week...	I felt hopeful about the future	-	Yes	1 to 4 ("Rarely" to "Most of the time")
Mental Health	Fearful	In the past week...	I felt fearful	Yes	Yes	1 to 4 ("Rarely" to "Most of the time")
Mental Health	Restless Sleep	In the past week...	My sleep was restless	Yes	Yes	1 to 4 ("Rarely" to "Most of the time")
Mental Health	Happy	In the past week...	I was happy	-	Yes	1 to 4 ("Rarely" to "Most of the time")
Mental Health	Lonely	In the past week...	I felt lonely	Yes	Yes	1 to 4 ("Rarely" to "Most of the time")
Mental Health	Couldn't get going	In the past week...	I could not get going	Yes	Yes	1 to 4 ("Rarely" to "Most of the time")
Personality	Agreeableness - Forgiving	I see myself as someone who ...	Has a forgiving nature	-	-	1 to 5 ("Disagree strongly" to "agree strongly")
Personality	Agreeableness - Considerate and Kind	I see myself as someone who ...	Is considerate and kind to almost everyone	-	-	1 to 5 ("Disagree strongly" to "agree strongly")
Personality	Agreeableness - Rude	I see myself as someone who ...	Is sometimes rude to others	Yes	-	1 to 5 ("Disagree strongly" to "agree strongly")
Personality	Conscientiousness - Thorough	I see myself as someone who ...	Does a thorough job	-	-	1 to 5 ("Disagree strongly" to "agree strongly")
Personality	Conscientiousness - Lazy	I see myself as someone who ...	Tends to be lazy	Yes	-	1 to 5 ("Disagree strongly" to "agree strongly")
Personality	Conscientiousness - Efficient	I see myself as someone who ...	Does things efficiently	-	-	1 to 5 ("Disagree strongly" to "agree strongly")
Personality	Extraversion - Talkative	I see myself as someone who ...	Is talkative	-	-	1 to 5 ("Disagree strongly" to "agree strongly")
Personality	Extraversion - Reserved	I see myself as someone who ...	Is reserved	Yes	-	1 to 5 ("Disagree strongly" to "agree strongly")
Personality	Extraversion - Outgoing	I see myself as someone who ...	Outgoing, sociable	-	-	1 to 5 ("Disagree strongly" to "agree strongly")
Personality	Neuroticism - Relaxed	I see myself as someone who ...	Is relaxed, handles stress well	-	-	1 to 5 ("Disagree strongly" to "agree strongly")
Personality	Neuroticism - Worries	I see myself as someone who ...	Worries a lot	-	-	1 to 5 ("Disagree strongly" to "agree strongly")
Personality	Neuroticism - Gets Nervous	I see myself as someone who ...	Gets nervous easily	Yes	-	1 to 5 ("Disagree strongly" to "agree strongly")
Personality	Openness - Original	I see myself as someone who ...	Is original, comes up with new ideas	Yes	-	1 to 5 ("Disagree strongly" to "agree strongly")
Personality	Openness - Active Imagination	I see myself as someone who ...	Has an active imagination	-	-	1 to 5 ("Disagree strongly" to "agree strongly")
Personality	Openness - Artistic and Aesthetic	I see myself as someone who ...	Values artistic, aesthetic experiences	-	-	1 to 5 ("Disagree strongly" to "agree strongly")
Positive / Negative Affect	Frustrated	Yesterday, did you feel ...	Frustrated	Yes	-	1 to 5 ("Not at all" to "Very")
Positive / Negative Affect	Sad	Yesterday, did you feel ...	Sad	Yes	-	1 to 5 ("Not at all" to "Very")
Positive / Negative Affect	Enthusiastic	Yesterday, did you feel ...	Enthusiastic	-	-	1 to 5 ("Not at all" to "Very")
Positive / Negative Affect	Lonely	Yesterday, did you feel ...	Lonely	Yes	-	1 to 5 ("Not at all" to "Very")
Positive / Negative Affect	Content	Yesterday, did you feel ...	Content	-	-	1 to 5 ("Not at all" to "Very")
Positive / Negative Affect	Worried	Yesterday, did you feel ...	Worried	Yes	-	1 to 5 ("Not at all" to "Very")
Positive / Negative Affect	Bored	Yesterday, did you feel ...	Bored	Yes	-	1 to 5 ("Not at all" to "Very")
Positive / Negative Affect	Happy	Yesterday, did you feel ...	Happy	-	-	1 to 5 ("Not at all" to "Very")
Positive / Negative Affect	Angry	Yesterday, did you feel ...	Angry	Yes	-	1 to 5 ("Not at all" to "Very")
Positive / Negative Affect	Tired	Yesterday, did you feel ...	Tired	Yes	-	1 to 5 ("Not at all" to "Very")
Positive / Negative Affect	Stressed	Yesterday, did you feel ...	Stressed	Yes	-	1 to 5 ("Not at all" to "Very")
Positive / Negative Affect	Pain	Yesterday, did you feel ...	Pain	Yes	-	1 to 5 ("Not at all" to "Very")
Subjective Wellbeing	Life Satisfaction (Cantrill Ladder)	How satisfied are you with your life as a whole		-	-	1 to 5 ("Not at all" to "Completely satisfied")
Subjective Wellbeing	Assessment of current situation	All things together how are things these days		-	Yes	1 to 4 ("Very unhappy" to "Very happy")
Risk/Time Preferences	Gamble Tradeoff 1	Sure gain of 800k/month, or series of 50/50 gambles		Yes	Yes	1 to 5 (Least Risk Averse to Most Risk Averse)
Risk/Time Preferences	Gamble Tradeoff 2	Sure gain of 4m/month, or series of 50/50 gambles		Yes	Yes	1 to 5 (Least Risk Averse to Most Risk Averse)
Risk/Time Preferences	1 year time tradeoff	Receive 1m today or higher amount in 1 year		-	Yes	1 to 5 (Least Patient to Most Patient)
Risk/Time Preferences	5 year time tradeoff	Receive 1m today or higher amount in 1 year		-	Yes	1 to 5 (Least Patient to Most Patient)

**Table 14: Description of candidate mediator variables from IFLS**

<i>Mediator Type</i>	<i>Measure</i>	<i>Description</i>
Child Health	Acute Morbidity	Number of listed symptoms experienced by child in last 4 weeks (including e.g. coughing, diarrhea)
Child Health	Height	Height (cm)
Child Health	Number of visits	Number of visits to hospital, health center, clinic, health worker in last 4 weeks
Child Health	Self-treat	Number of times child treated him/herself in the past 4 weeks
Child Health	Weight	Weight (kg)
Early Life Investment	Birth with Health Care (=1)	Birth took place at a health care facility or with a health care provider present
Early Life Investment	Breastfeed (=1)	Child was breastfed
Early Life Investment	Educational Expenditure	Average monthly expenditure on child's education for the last year
Early Life Investment	Food Frequency	How often does child eat? (from 3+ times a day to 2 or less times a week)
Early Life Investment	Iron Frequency	Number of times per week child eats iron-rich food
Early Life Investment	Father Smokes (=1)	Father currently smokes
Early Life Investment	Mother Smokes (=1)	Mother currently smokes
Early Life Investment	Vitamin A Frequency	Number of times per week child eats vitamin A-rich food
Parental Education	Highest Level of Parental Education	Highest level of education reached by a parent
Parental Health	Days Spent in Bed - Father	Number of days spent in bed due to bad health in last 4 weeks
Parental Health	Days Spent in Bed - Mother	Number of days spent in bed due to bad health in last 4 weeks
Parental Health	Days Missed - Father	Number of days missed of primary activity due to bad health in last 4 weeks
Parental Health	Days Missed - Mother	Number of days missed of primary activity due to bad health in last 4 weeks
Parental Health	Self-reported Health - Father	Self-reported general health (1 = Very healthy to 4 = Unhealthy)
Parental Health	Self-reported Health - Mother	Self-reported general health (1 = Very healthy to 4 = Unhealthy)
Parental Health	Pregnancy Length	Length of pregnancy in months
Parental Time	Employed - Father (=1)	Father earned income or worked for 1 hour in the last week, or was temporarily not working but has a job
Parental Time	Employed - Mother (=1)	Mother earned income or worked for 1 hour in the last week, or was temporarily not working but has a job
Parental Time	Primary Activity is earning - Father (=1)	Father's primary activity during the last week was working / trying to work
Parental Time	Primary Activity is earning - Mother (=1)	Mother's primary activity during the last week was working / trying to work
Parental Time	Primary Activity is housekeeping - Father (=1)	Father's primary activity during the last week was housekeeping
Parental Time	Primary Activity is housekeeping - Mother (=1)	Mother's primary activity during the last week was housekeeping
Parental Time	Weekly Hours worked - Father	Normally, what is the total number of hours worked by week
Parental Time	Weekly Hours worked - Mother	Normally, what is the total number of hours worked by week

some split districts are kept together or combined into a larger unit in order to maintain consistency. The geographical unit of analysis throughout the paper is thus the IPUMS 'harmonised' district.

**Climate data.** The primary weather variable I use is the Standardised Precipitation Evapotranspiration Index (SPEI), a drought index developed recently by [Vicente-Serrano et al. \(2010\)](#). Although most of the economic literature on weather shocks focuses solely on precipitation and temperature, this neglects other features of the climate that can affect the growing cycle of a plant. In particular, the impact of rainfall on crop cycles depends on *potential evapotranspiration* (the ability of soil to retain water), which is in turn affected by multiple features of the environment, including temperature, latitude, windspeed, the number of sunlight hours. SPEI is a rich index that combines all of these features. It has been shown to be a better predictor of crop yields than other climate indexes [Beguería et al. \(2014\)](#). The methodology behind the construction of the SPEI index is described in more detail in [Vicente-Serrano et al. \(2010\)](#) and [Beguería et al. \(2014\)](#).

The Global SPEI database ([Beguería and Vicente Serrano, 2017](#)) uses precipitation and potential evaporation data to calculate monthly SPEI estimates in grids of 0.5 latitude by 0.5 longitude (corresponding to grids of approximately 50km by 50km at the equator). Monthly SPEI data is available for all years from 1900 to 2015. The index is standardised to have a mean of 0 and a standard deviation of 1 within each grid cell over the entire historical period (1900 to 2015), so that positive SPEI indexes generally represent good climactic conditions for crop-growing relative to the historical average, while negative SPEI indexes represent conditions more likely to lead to drought relative to the historical average. The data is fitted to a log-logistic distribution and can be normalised to a number of different time scales (e.g. 1, 2, 4, 6, 12 months, etc.). Since I am interested in the effect of weather conditions primarily through its effects on agriculture, I follow [Harari and La Ferrara \(2018\)](#) in choosing a relatively short 4-month time scale for the SPEI calculation. This means the SPEI will reflect short- and medium-run changes in moisture conditions that will be relevant for the seasonal changes that are important for crop yields.

In order to calculate the monthly district-level SPEI, I match each district to all the grid-squares with which it overlaps. I then use a weighted average of SPEI for all the matched grid-squares, where the weights are proportional to the area of the square that overlaps with the district. For example, if exactly half of district *A* sits inside grid-square  $s_1$ , and the other half sits inside grid-square  $s_2$ , then the estimated rainfall in district *A* would be the (unweighted) arithmetic mean of the rainfall in squares  $s_1$  and  $s_2$ .

My approach for calculating the district-level weather patterns is likely to be preferable to that seen in previous estimates from Indonesia ([Maccini and Yang, 2009](#); [Bharati et al., 2017](#)). [Maccini and Yang \(2009\)](#) use data on the latitude and longitude of the *centroid* of each district to match each district to the nearest weather station, and then use monthly rainfall data from this station directly.<sup>45</sup> This measurement is likely to induce non-negligible measurement error; while the median distance between the district centroid and weather station in their sample period is only 14km, the 95th percentile is 70km and the maximum is approximately 230km. The main advantage of the approach that I take is that I make use of district shapefiles rather than just data on the district centroid. This means that I accurately account for the different shapes and sizes of each district when calculating district-level weather.

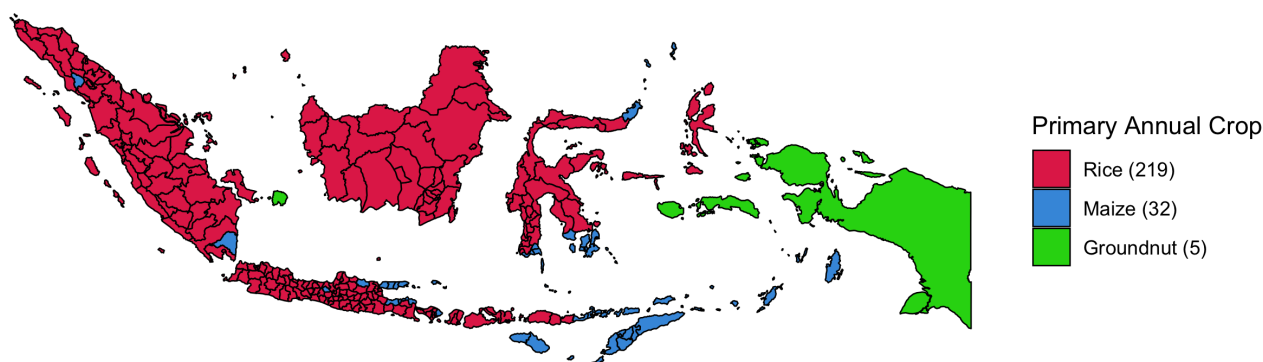
**Crop harvest data.** Crop harvested-area data comes from [Monfreda et al. \(2008\)](#), who combine national, state, and county level census data to create a global data set that describes the harvested area for 175 crops on a 5 minute by 5 minute (approximately 10km by 10km) grid. Using this data, I calculate the primary annual crop in each district according to the total harvested area.<sup>46</sup> Figure

<sup>45</sup>Their station-level rainfall data comes from Global Historical Climatology Network (GHCN) Precipitation and Temperature Data (Version 2) and the Badan Meterologi Dan Geofisika (BMG) agency in Indonesia.

<sup>46</sup>I exclude perennial crops such as oil palm from this calculation, since seasonal weather shocks are unlikely to affect production or yields.

13 shows the primary annual crop in each district in the data. Rice is the primary annual crop for 219 out of 256 district, with production particularly concentrated in many areas of Java. The primary annual crop for 32 districts is maize, and for 5 it is groundnut. The districts in which groundnut is the most harvested annual crop tend to have a lower proportion of harvested area overall, and since the IFLS sample is taken from the 13 provinces indicated **above** in the western areas of Indonesia, these areas are unlikely to be driving the results I show in the main body of the paper. For example, only 9 individuals in the IFLS sample were born in the large eastern island of Papua highlighted in green on the map.

**Figure 13:** Primary Annual Crops (by harvested area) in each Indonesian district



**Crop calendar data.** In order to identify the primary growing season in each district, I make use of crop calendar data from the University of Wisconsin-Madison’s Nelson Institute (Sacks et al., 2010). This provides grid-cell level ( $0.5 \times 0.5$  degree) data for planting and harvesting dates of 19 different crops, including rice, maize, and groundnuts. I aggregate this to the district level by choosing the most common planting and harvesting dates over all the cells in the district.

**Shock definition.** In order to match individuals to shocks, I take the following steps, loosely following Maccini and Yang (2009). For individual  $i$  in year  $y$ :

1. Identify the district  $r$  that  $i$  was in during year  $y$  using the full migration history described above in Appendix Section A.1.
2. Identify the plant and harvest dates of the main crop in district  $r$ .
3. If  $i$  was born before the harvest date of  $y$ , assign  $i$  to the growing season in  $y - 1$ . If  $i$  was born after the harvest date of  $y$ , assign  $i$  to the growing season in  $y$ .
4. Generate  $SPEI_{iy}$  as the average SPEI over the assigned growing season in region  $r$ .

This shock definition is based on the assumption that weather patterns during the growing season affect agricultural income at harvest time, meaning that the effects of the shock will be felt only after harvest. Individuals are therefore matched to the growing season of the most recent harvest.

## B Exploratory Factor Analysis

As discussed in Section 3.3, I use a latent factor model in the style of Cunha et al. (2010) to combine data from multiple measures of skills to a lower-dimensional set of psychological “traits”. The latent factor model explicitly allows for measurement error and combines data from multiple measures in order to mitigate the problem of measurement error on psychological variables.

I carried out an exploratory factor analysis (EFA) exercise on three types of variables: cognitive measures, socioemotional measures, and the candidate mediator variables used in the mediation analysis in Section 6.3.

The EFA process used was similar across all of these categories of measures. The process is described in detail throughout the rest of this section.

For the psychological outcome and mediation measures of interest, I took the following five steps to carry out the exploratory factor analysis:

1. Correct survey responses for acquiescence bias (for socio-emotional measures only).
2. Standardise measures by age and survey wave (where required).
3. Determine the number of latent factors underlying the set of measured variables.
4. Estimate the factor loadings for each measure.
5. Use the factor loadings to generate factor scores for each individual in the survey

I describe each of these steps below. Note that I for psychological outcome measures (only available in IFLS4 and IFLS5), I carried out separate factor analyses for each wave of the survey, so that there are two different factor structures, one for the questions contained in IFLS4 and one for the questions contained in IFLS5. For the mediation measures, I pooled data available from all waves in the analysis.

### B.1 Correcting for acquiescence bias (socioemotional only)

Acquiescence bias (also known as “yay-saying”) is the tendency of a respondent to agree (or disagree) with a statement from the survey enumerator, even if doing so results in contradictory responses that are intended to capture the same trait. For example, a biased respondent may be likely to agree with both the statement “I see myself as someone who is talkative” and with the statement “I see myself as someone who is reserved”, even though the two responses contradict each other. The acquiescence bias may be especially problematic given the lower-middle income context in Indonesia; cross-country evidence using the Big 5 personality tests tend to show higher acquiescence bias in lower income settings (Rammstedt and Farmer, 2013), and Laajaj and Macours (2017) find evidence of systematic measurement error related to acquiescence bias among farmers in western Kenya, which is more severe among the less educated. Using exploratory factor analysis without correcting for such bias can lead to misleading results, including the emergence of a factor that is an artefact of the response bias rather than underlying socio-emotional variation. Laajaj and Macours (2017) further show that correcting for acquiescence bias substantially improves the reliability of socio-emotional constructs, so I follow them and Laajaj et al. (2018) in carrying out a correction that is common in the psychometric literature (Soto et al., 2008; Rammstedt and Farmer, 2013). The correction makes use of the fact that for some socio-emotional traits, there are some questions that are positively-coded, and some that are reverse-coded, and so we can detect the acquiescence pattern for a given individual. I take the

following steps to correct for acquiescence on all socio-emotional outcomes that use a Likert scale:<sup>47</sup>

1. Reverse the reverse-coded items. For example, for a scale that goes from 1 to 5, 1 will be recoded as 5, 2 as 4, etc.
2. For every personality trait that has at least one reverse-coded item and one positively-coded item:
  - (a) Take the average of the positively-coded items for each individual  $i$ .
  - (b) Subtract from this the average of the reverse-coded items for the same individual  $i$ .
  - (c) Divide this by two.
3. Calculate the overall acquiescence score  $AS_i$  for individual  $i$  by averaging over the differences for all personality traits calculated in step 2.
4. Correct individual  $i$ 's raw scores for acquiescence bias by adding  $AS_i$  to every reverse-coded item, and subtracting  $AS_i$  from every positively-coded item.

These corrected socioemotional scores are then used in all the steps below.

## B.2 Age Standardisation

**Cognitive and socioemotional.** IFLS respondents are measured at different ages in adulthood, and both cognitive and socioemotional measures may vary by age. Therefore in order to get comparable measures of cognitive and socioemotional skills across different age groups I need to carry out some form of age standardisation. For each measure, I calculate the mean and standard deviation of the measure for each age group (by years). I then make an internal standardisation (Schady, 2011) by taking the age-specific z-score based on these means and standard deviation.

**Mediator variables.** Mediator variables also vary by age of the child. For example, the sample probability of a mother being in employment (a proxy for early investment) increases after the child reaches one year old (not shown). In order to generate measures of early investment, child health, parental health, and parental employment that are comparable across ages, I make the same internal standardisation using the age-specific z-score. The main difference in my treatment of the mediator variables is that I amalgamate z-scores across waves at this stage. For example, if I have a measure of food frequency z-score in both IFLS2 and IFLS3 for individual  $i$ , then I create an overall z-score by taking the unweighted average of these two measures. These are the measures then used in the subsequent steps.

## B.3 Determining the number of factors

In order to determine the number of factors that should be extracted from the data, I make use of 4 commonly-used methods that suggest how many factors to extract.

The methods are as follows:

1. *Kaiser's criterion* (Kaiser, 1958). This rule-of-thumb suggests retaining all factors with an eigenvalue of above 1. The intuition that underlies this is that factors should only be kept if they explain as much variance as the equivalent of one original variable. This tends to suggest

---

<sup>47</sup>The 'Personality', 'Positive / Negative Affect', and 'Mental Health' modules used a Likert scale and had both positive and reverse-coded items. Because the phrasing of the Likert scale on each of these modules was different (see Table 13), it appears to have led to a very different pattern of response bias (see Figure 15), and so I generated separate acquiescence scores for each module and made one acquiescence correction for each module. In other words, I carried out steps 1-4 separately for each module, treating each Big 5 personality trait as a different trait in step 2, positive/negative affect as a single personality trait, and mental health as a single personality trait.

retaining a low number of factors; newer literature often suggests using a threshold value of 0.7 instead. I show results for both.

2. *Scree plot* (Cattell, 1966). This test is based on visual inspection of a plot of eigenvalues associated with the data. The proposed number of factors is equal to the number of factors before the 'elbow' of the scree plot; that is, the number before the which the smooth decrease of eigenvalues appears to level off towards the right of the plot.
3. *Minimum Average Partial (MAP)* (Velicer, 1976). This rule chooses the number of factors that minimises the unexplained partial correlation. The procedure iterates as follows. For  $r = 1$ , we run a full factor analysis with one factor, partial out this factor from the correlation matrix of the variables of interest, and calculate the average squared coefficient in the off-diagonals of this partial correlation matrix. Then, for  $r = 2$ , we run a similar process in which we partial out two factors, and calculate the average squared off-diagonal from the resulting matrix. This process is repeated up until  $r = K - 1$  where  $K$  is the number of measurements. The recommended number of factors  $r^*$  is the value of  $r$  that minimises the average squared partial correlation. The intuition behind this process is that components are retained as long as the variance in the correlation matrix represents systematic variance. They are dropped when there is proportionately more unsystematic variance than systematic variance (see Attanasio et al., 2017, Appendix p. 15).
4. *Parallel Analysis* (Horn, 1965). Horn's test involves a Monte-Carlo procedure in which we simulate uncorrelated random variables of the same dimension as the actual data of interest. The eigenvalues derived from the actual data are then compared to the eigenvalues from the randomly generated data, and factors are kept if the eigenvalue is greater than the 95th percentile from the simulated data.

**Cognitive and socioemotional.** I carry out each of these methods first for all cognitive measures, and then for all socioemotional measures. All methods recommended one factor for cognitive measures in both IFLS4 and IFLS5. Table 15 shows the recommendation of the number of factors using each method for socioemotional measures. Each different method tends to give different recommendations, so some judgement is required in deciding how many factors to retain. Ledesma and Valecro-Mora (2007) suggest that Velicer's MAP criterion and Horn's Parallel Analysis tend to be more reliable, so the number of factors I retain for analysis is the number most commonly recommended from all the methods, with more weight placed on the parallel analysis and MAP criteria. Since there were fewer socio-emotional questions asked in IFLS4 than in IFLS5 (see Table 13 for details), I expect there to be fewer factors for IFLS4; the fact that all methods suggest a smaller number pf factors in IFLS4 reflects this. For IFLS4, the scree plot (see Fig. 14a), Kaiser criterion (1.0) and MAP all suggest that there is 1 factor. For IFLS5, the same set of methods suggest a value of 3 (see also Fig. 14b). In both cases, parallel analysis appears to overstate the number of latent factors. Based on this process, I retain 1 socioemotional factor for analysis in IFLS4, and 3 socioemotional factors in IFLS5.

I also carry out all of the above methods for each group of mediator variables that has three or more available measures (this therefore excludes parental education, which only has one measure). Based on the recommendations from each of the methods, I retain 2 factors for child health, 1 factor for early investment, 1 factor for parental health, and 2 factors for parental time use.

#### B.4 Estimating the factor analysis model

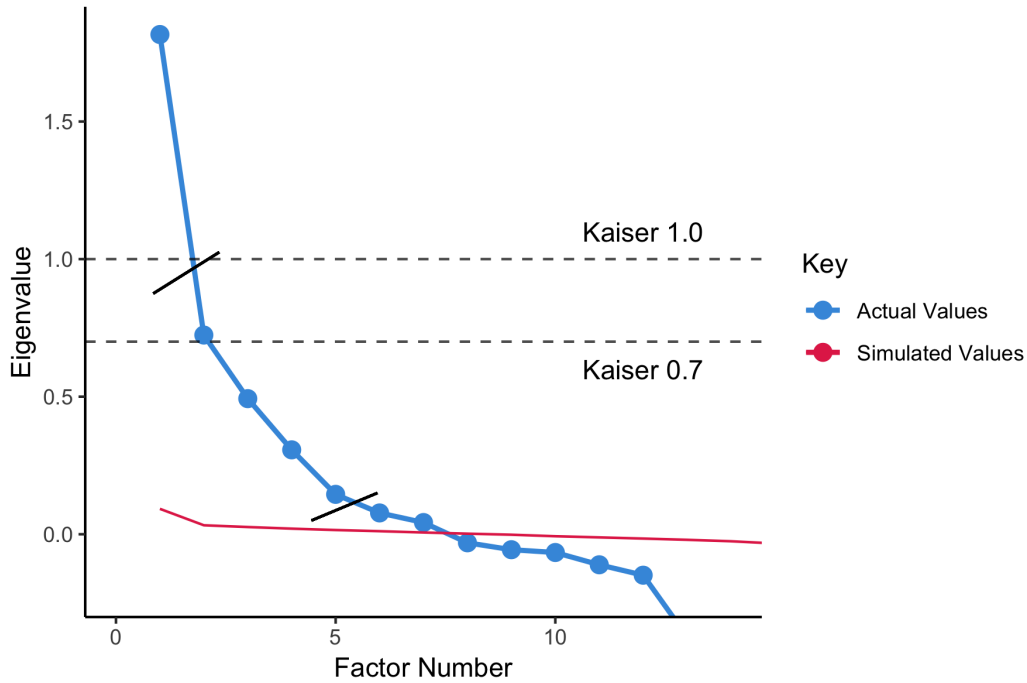
After identifying how many factors should be used in the factor analysis, I run an exploratory factor analysis (EFA) on measures (which have been corrected for acquiescence bias where relevant and age-standardised). The EFA estimates factor loadings by minimising the squared residuals<sup>48</sup>, and then carries out a factor rotation that ensures that each measure primarily loads onto a single factor. I use

<sup>48</sup>This has been shown to give unbiased estimates of factor loadings in the context of ordinal variables such as those used here, see Lee et al., 2012.

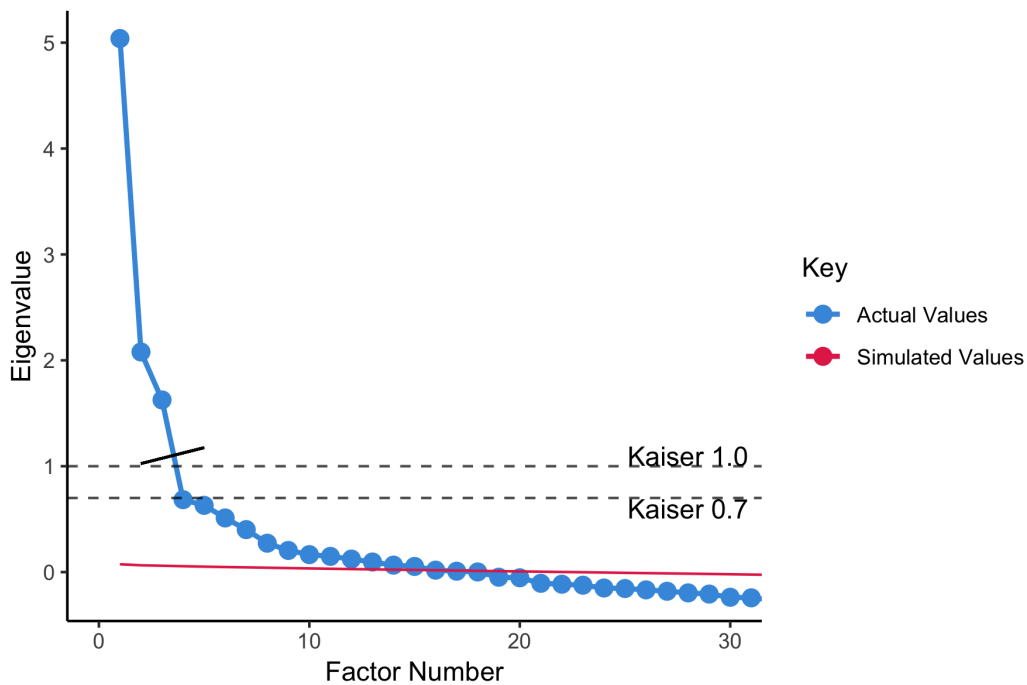


**Figure 14:** Scree Plots for Exploratory Factor Analysis

(a) Socio-emotional variables in IFLS4



(b) Socio-emotional variables in IFLS5



Notes: 'Actual values' denotes the eigenvalues from the true data. 'Simulated values' (in red) denotes the values from the 95th percentile of the randomly generated uncorrelated normal variables. The black horizontal dashed lines indicate the threshold points for the Kaiser criteria (1.0 and 0.7), and the solid black marks indicate possible positions for the 'elbow' of the scree plot in Cattell (1966).

**Table 15:** Recommendation of how many factors to retain from each method

Method	IFLS4	IFLS5
Kaiser Criterion (1.0)	1	3
Kaiser Criterion (0.7)	2	4
Scree Plot	1 or 5	3
Parallel Analysis	8	17
Minimum Average Partial (MAP)	1	3
<i>Retained for Analysis</i>	1	3

Notes: Scree plots for each IFLS wave can be seen in Figures 14a and 14b. Each method recommends fewer factors for IFLS4 than IFLS5 because IFLS4 contains fewer socio-emotional questions than IFLS 5 (see Table 13 for details).

an *oblimin* rotation to rotate factors; this has the advantage of being an *oblique* rotation rather than *orthogonal*, meaning that I do not have to make the strong assumption that factors are uncorrelated<sup>49</sup>. The aim of the rotation is to identify measures that primarily load onto a single factor, which can then be used as a ‘dedicated’ measure for that factor. Measures that load onto more than one factor, or onto no factors, will not be used in the measurement system.

Table 17 shows the factor loadings for the socio-emotional measures available in IFLS5, and Table 18 shows the factor loadings for the socio-emotional measures available in IFLS4. Following standard practice in the psychometric literature (Rammstedt and Farmer, 2013; Attanasio et al., 2017), I deem a measure to ‘load’ on a factor if the (absolute value of the) factor loading is greater than 0.3, and have highlighted cells in Tables 17 and 18 where the value meets this condition.

The factor loadings seen in Table 17 shows that the underlying latent variables coincide relatively closely with the survey modules in which they are used, and thus permit a clear interpretation for each factor (see Table 16) Factor 1 is loaded on strongly by all the questions in the Positive/Negative Affect (PNA) module, along with the two Subjective Well-being questions and the question from the Mental Health module relating to whether the individual felt happy in the last week. This factor can therefore be broadly interpreted as a measure of affect and well-being. Factor 3 has a similarly straightforward interpretation, being composed of all mental health measurements apart from ‘Everything was an effort’, which does not load strongly (less than 0.3) on any factor. Promisingly, the factor loadings for the single factor in IFLS4 (Table 18) closely accord with the factor loadings of Factor 3 in IFLS5, with all mental health measures loading except ‘Everything was an effort’. This implies that the factor structure here is congruent across different time periods and populations.

**Table 16:** Dominant interpretation of each factor

Factor	Wave	Interpretation
Factor 1	IFLS4	“Mental Health”
Factor 1	IFLS5	“Affect / Wellbeing”
Factor 2	IFLS5	“Personality”
Factor 3	IFLS5	“Mental Health”

Notes: Primary interpretations for each factor, based on the factor loadings seen in Tables 17 and 18

The factor loadings on Factor 2 in the personality module are less uniform and do not conform with the ‘Big 5’ personality traits on which they are based. This is in keeping with Laajaj and Macours (2017) and Laajaj et al. (2018) who show that factor structures tend to deviate from the Big 5 in low income contexts. Nevertheless, the personality measures do load onto a single factor, with the

<sup>49</sup>This follows recent literature that also use oblique rotations, see Laajaj and Macours, 2017, Attanasio et al., 2017. Multiple methods for carrying out an oblique factor rotation are available, therefore as a robustness check I carried out the same exercise with both *geomin* and *quartimin* rotations and obtained similar results.

most consistent loadings coming from the *agreeableness* and *conscientiousness* traits, all of which have loadings greater than 0.3. Measures of *extroversion*, *neuroticism*, and *openness* do not consistently load on Factor 2, despite a particularly strong factor loading for 'Extraversion - Outgoing' (0.587).

Notably, measures of risk and time preferences do not load onto any factor in either survey wave. This implies that such measures are subject to severe measurement error would lead to misleading results if used as an outcome variable. This is in keeping with evidence that suggests self-reported responses for such measures that involve no real money at stake (unlike experimental measures of risk aversion or patience) tend to be highly unreliable ([Binswanger, 1980](#); [Chuang and Schechter, 2015](#)).

There is little evidence of acquiescence bias in the socio-emotional factor loadings: there is no discernible pattern where only reverse-coded or positively-coded items load onto a factor. Given that such patterns were observable when I ran the factor analysis without acquiescence correction (not shown), this implies that the acquiescence correction described above is effective.

Notably when looking at the interpretations of Factors 1 and 2, very similar questions appear to load on different factors depending on the module they are in. For example, both the Positive-Negative Affect (PNA) and the Mental Health modules ask if the individual was lonely, but the measurements do not load on the same factor. There may be two main reasons for this. First, we may be worried that the responses to similar questions in different modules are biased in different ways due to changes in response pattern, possibly due to the different scales used, or the fact that the Mental Health module comes significantly later in the survey than the PNA module. This worry is mitigated somewhat by the acquiescence correction, which accounts for the proportion of bias that is due to "yay-saying". Alternatively, the difference could be due to the change in temporal scope in each module. In other words, since the PNA module asks about feelings yesterday, while the Mental Health module asks about feelings over the past week (see [Table 13](#)), respondents could give justifiably different answers. This consideration is confused somewhat by the fact that the subjective well-being questions (that target longer-term cognitive evaluation of happiness) tend to load alongside the short-term affect questions rather than the medium-term mental health questions. One possible explanation may be that the long-term subjective well-being questions are mistakenly interpreted by respondents as an evaluation of short-term happiness; this type of problem is well known in the literature on measuring well-being (see e.g. [White, 2018](#)).

Whichever explanation is true, the considerations above point towards the importance of running a factor analysis exercise on socioemotional measures, rather than naively interpreting each question to be measuring the trait it is nominally intended to measure. If the answer to a question about loneliness is so affected by seemingly extraneous factors such as the phrasing of the question or the position within the model, then the answer to such a question should not be taken 'at face value'. Instead, it is better to assume that it is a proxy for some underlying latent factor, and that this factor can in turn be recovered from multiple proxy measures. Nevertheless, we must be equally wary in interpreting the latent factors themselves: even though they are designed to pick up true variation (instead of measurement error), this true variation may still be at least partly due to seemingly extraneous factors like survey design and response biases, and so should be interpreted with caution.

[Table 19](#) shows the cognitive factor loadings for IFLS4, and [Table 20](#) shows the cognitive factor loadings for IFLS5. In both waves, all the cognitive measures load onto the same single factor, implying that this factor can be clearly interpreted as measuring cognitive ability. The written tests (Raven, Written Maths, and Number Series), have higher loadings in both waves than the oral tests (Word Recall and Oral Maths), suggesting that these capture more of the variation in the latent cognitive factor. Overall, the loadings for cognitive a clear interpretation of a single cognitive factor that is in keeping with previous literature [Laajaj and Macours \(2017\)](#).

[Tables 21, 22, 23, and 24](#) show the estimated factor loadings for each of the mediator variable groups with more than 3 measures. The child health measures ([Table 21](#)) split neatly into two factors, the

first of which can be interpreted as representing acute morbidity and health care usage, the second as a summary of anthropometric scores. The factor for early investment (Table 22) is loaded on by the measure of whether birth took place with a health care provider present, the level of educational expenditure, and two measures of nutritional investment (the frequency of iron-rich and vitamin-A-rich foods). I interpret this factor as a broad measure of early investment. The factor for parent health (Table 23) is loaded on by all acute measures of father and mother health, but not by the length of pregnancy. It can therefore again be interpreted broadly as “parental health”. Finally, the parent time measures split into two factors which are interpreted as mother employment and father employment respectively (Table 24).

### B.5 Generating factor scores

I use the factor loadings from Tables 17–24 to generate factor scores for each individual  $i$ . These factor scores represent the ‘value’ of that factor for each individual  $i$ . In order to generate the factor scores for factor  $k$  I carry out the following steps:

1. Assign each question to ‘load’ on factor  $k$  if and only if the (absolute value of) the factor loading is greater than or equal to 0.3 (following standard practice in the psychometric literature, see [Rammstedt and Farmer, 2013](#); [Attanasio et al., 2017](#)). All other questions have their factor loadings set to 0.
2. For each individual  $i$ , take a weighted mean of the questions that load onto factor  $k$ , where the weights correspond to the loadings obtained in Tables 17, 18, 19 and 20.
3. Standardise the factor scores so that each factor has a mean of 0 and a standard deviation of 1.

Using this method to obtain factor scores ensures that questions are only kept if they clearly load onto a particular factor, while questions with particularly weak loadings (such as the questions on risk and time preferences) are not included in the factor scores. Step 3 ensures that the factor scores are clearly interpretable. For example, if individual  $i$  has a factor score of 1 on Factor 2 (“Personality”) in IFLS5, this means that her non-cognitive personality skills are 1 standard deviation above the sample mean.

The process described here generates a separate factor scores in each wave for the two psychological factors that are measured in both IFLS4 and IFLS5 (mental health and cognitive).<sup>50</sup> The factor structures for these factors is strikingly similar across waves. I therefore construct the final outcome variables used in the main part of the paper by combining measurements across waves, in particular by taking a simple arithmetic mean of the factor score in IFLS4 and the factor score in IFLS5 for any individuals who were measured in both waves. The distribution of the finalised factor scores among my main sample are shown in Figure 16. By construction, all the distributions are centred around 0, and they all appear to be approximately symmetrical.

---

<sup>50</sup>Mediator variables, on the other hand, are already combined across waves as described in Step 2.

**Table 17: Socio emotional factor Loadings for IFLS5**

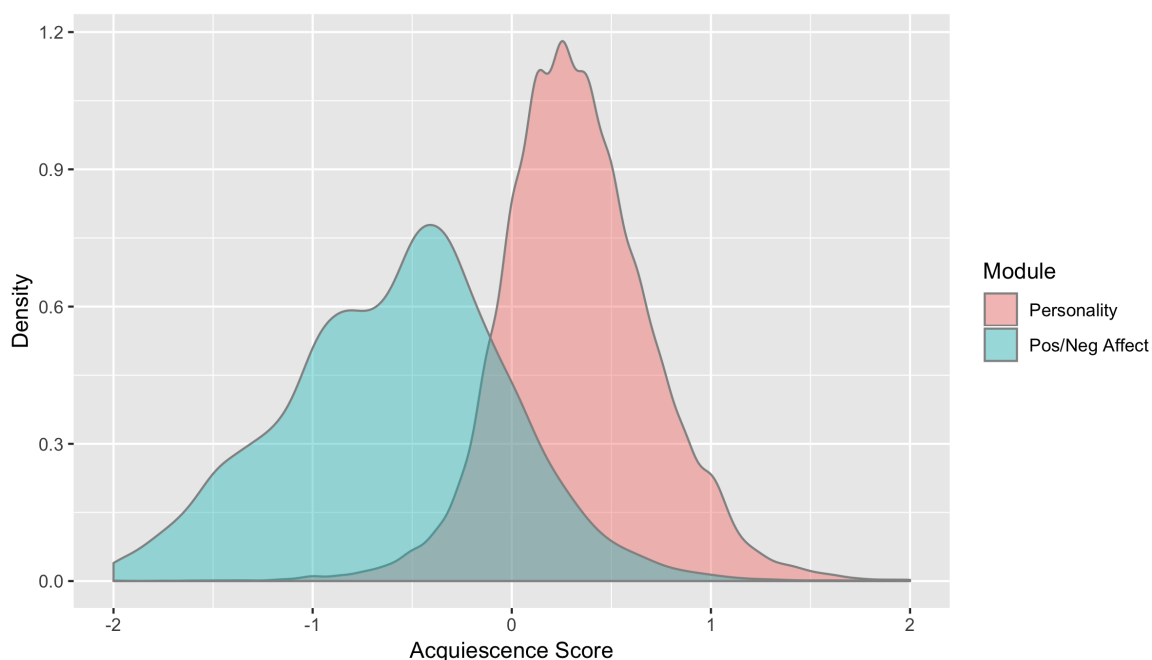
Module	Question	Factor 1	Factor 2	Factor 3
Positive / Negative Affect	Frustrated (-)	0.606	-0.004	0.006
Positive / Negative Affect	Sad (-)	0.693	-0.020	-0.011
Positive / Negative Affect	Enthusiastic	0.330	0.044	0.007
Positive / Negative Affect	Lonely (-)	0.519	-0.017	0.046
Positive / Negative Affect	Content	0.632	-0.029	-0.034
Positive / Negative Affect	Worried (-)	0.466	0.005	0.039
Positive / Negative Affect	Bored (-)	0.594	0.088	-0.052
Positive / Negative Affect	Happy	0.557	0.019	0.026
Positive / Negative Affect	Angry (-)	0.580	0.030	0.007
Positive / Negative Affect	Tired (-)	0.363	0.028	-0.041
Positive / Negative Affect	Stressed (-)	0.672	-0.006	-0.028
Positive / Negative Affect	Pain (-)	0.299	0.001	0.026
Subjective Wellbeing	Assessment of current situation	0.364	0.008	0.030
Subjective Wellbeing	Life Satisfaction (Cantrill Ladder)	0.346	0.020	-0.019
Mental Health	Bothered (-)	0.090	-0.023	0.421
Mental Health	Concentration (-)	0.093	0.042	0.455
Mental Health	Depressed (-)	0.207	-0.039	0.512
Mental Health	Everything was an effort (-)	0.184	-0.015	0.198
Mental Health	Hopeful	-0.191	0.011	0.726
Mental Health	Fearful (-)	0.107	0.070	0.447
Mental Health	Restless Sleep (-)	0.157	0.037	0.363
Mental Health	Happy	0.445	0.012	0.320
Mental Health	Lonely (-)	0.070	-0.024	0.495
Mental Health	Couldn't get going (-)	0.064	0.060	0.440
Personality	Agreeableness - Forgiving	0.000	0.530	-0.016
Personality	Agreeableness - Considerate and Kind	-0.031	0.606	-0.031
Personality	Agreeableness - Rude (-)	0.020	0.385	-0.025
Personality	Conscientiousness - Thorough	0.026	0.479	0.037
Personality	Conscientiousness - Lazy (-)	-0.004	0.427	-0.038
Personality	Conscientiousness - Efficient	0.048	0.292	0.073
Personality	Extroversion - Talkative	-0.009	0.099	0.001
Personality	Extroversion - Reserved (-)	-0.030	0.164	0.052
Personality	Extroversion - Outgoing	0.015	0.543	0.004
Personality	Neuroticism - Relaxed	0.072	0.295	0.096
Personality	Neuroticism - Worries (-)	0.130	0.104	0.170
Personality	Neuroticism - Gets Nervous (-)	0.064	0.289	0.049
Personality	Openness - Original	0.006	0.407	0.043
Personality	Openness - Active Imagination	-0.042	0.283	0.043
Personality	Openness - Artistic and Aesthetic	-0.032	0.393	0.017
Risk/Time Preferences	Gamble Tradeoff 1 (-)	-0.026	0.031	0.011
Risk/Time Preferences	Gamble Tradeoff 2 (-)	0.018	-0.004	-0.052
Risk/Time Preferences	1 year time tradeoff	-0.025	0.009	0.031
Risk/Time Preferences	5 year time tradeoff	-0.002	0.010	0.006
N				4678

Notes: Cells highlighted in blue have an absolute value of greater than 0.3. '(-)' denotes a reverse-coded item.

**Table 18: Socio emotional factor Loadings for IFLS4**

Module	Question	Factor 1
Mental Health	Bothered (-)	0.634
Mental Health	Concentration (-)	0.532
Mental Health	Depressed (-)	0.612
Mental Health	Everything was an effort (-)	0.283
Mental Health	Hopeful	0.463
Mental Health	Fearful (-)	0.493
Mental Health	Restless Sleep (-)	0.455
Mental Health	Happy	0.542
Mental Health	Lonely (-)	0.617
Mental Health	Couldn't get going (-)	0.634
Subjective Wellbeing	Assessment of current situation	0.240
Risk/Time Preferences	Gamble Tradeoff 1 (-)	-0.043
Risk/Time Preferences	Gamble Tradeoff 2 (-)	-0.025
Risk/Time Preferences	1 year time tradeoff	-0.035
Risk/Time Preferences	5 year time tradeoff	-0.043
N		2049

Notes: Cells highlighted in blue have an absolute value of greater than 0.3. ‘(-)’ denotes a reverse-coded item.



**Figure 15: Different Pattern of Acquiescence scores for the Personality module compared to the Positive/Negative Affect module**

**Table 19: Cognitive Factor Loadings (IFLS4)**

Module	Question	Factor 1
Cognitive	Raven (15-24)	0.568
Cognitive	Maths - Written (15-24)	0.488
Cognitive	Word Recall	0.384
N		1922

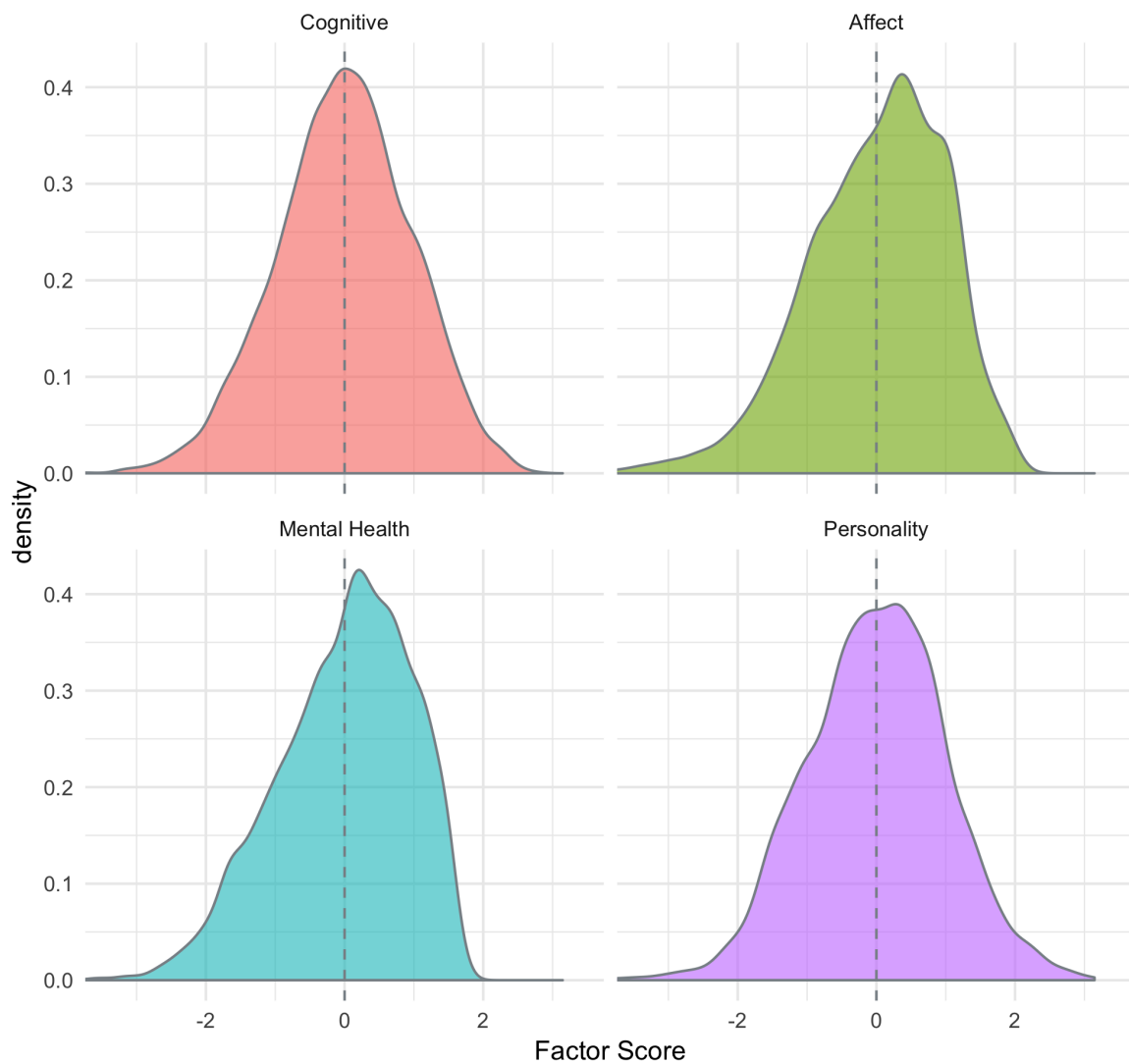
Notes: Cells highlighted in blue have an absolute value of greater than 0.3.

**Table 20: Cognitive Factor Loadings (IFLS5)**

Module	Question	Factor 1
Cognitive	Raven (15-24)	0.519
Cognitive	Maths - Written (15-24)	0.530
Cognitive	Word Recall	0.367
Cognitive	Maths - Oral	0.380
Cognitive	Number Series	0.585
N		4991

Notes: Cells highlighted in blue have an absolute value of greater than 0.3.

**Figure 16: Kernel density of each of the factor scores**



**Table 21: Child health**

Module	Question	Factor 1	Factor 2
Child Health	Acute Morbidity (-)	1.002	-0.001
Child Health	Height	-0.004	0.994
Child Health	Number of visits (-)	0.332	-0.022
Child Health	Self-treat (-)	0.506	0.011
Child Health	Weight	0.017	0.606
N			3258

Notes: Cells highlighted in blue have an absolute value of greater than 0.3. ‘(-)’ denotes a reverse-coded item.

**Table 22: Early investment**

Module	Question	Factor 1
Early Life Investment	Birth with Health Care (=1)	0.396
Early Life Investment	Breastfeed (=1)	-0.047
Early Life Investment	Educational Expenditure	0.415
Early Life Investment	Food Frequency	0.046
Early Life Investment	Iron Frequency	0.433
Early Life Investment	Father Smokes (=1)	-0.144
Early Life Investment	Mother Smokes (=1)	-0.005
Early Life Investment	Vitamin A Frequency	0.272
N		3377

Notes: Cells highlighted in blue have an absolute value of greater than 0.3. ‘(-)’ denotes a reverse-coded item.

**Table 23: Parent health**

Module	Question	Factor 1
Parental Health	Days Spent in Bed - Father (-)	0.565
Parental Health	Days Spent in Bed - Mother (-)	0.367
Parental Health	Days Missed - Father (-)	0.675
Parental Health	Days Missed - Mother (-)	0.437
Parental Health	Self-reported Health - Father	0.512
Parental Health	Self-reported Health - Mother	0.410
Parental Health	Pregnancy Length	-0.046
N		3928

Notes: Cells highlighted in blue have an absolute value of greater than 0.3. ‘(-)’ denotes a reverse-coded item.

**Table 24: Parent time**

Module	Question	Factor 1	Factor 2
Parental Time	Employed - Father (=1)	0.021	0.817
Parental Time	Employed - Mother (=1)	0.833	0.051
Parental Time	Primary Activity is earning - Father (=1)	-0.015	0.946
Parental Time	Primary Activity is earning - Mother (=1)	0.975	0.008
Parental Time	Primary Activity is housekeeping - Father (=1)	-0.021	-0.550
Parental Time	Primary Activity is housekeeping - Mother (=1)	-0.946	0.060
Parental Time	Weekly Hours worked - Father	-0.011	0.484
Parental Time	Weekly Hours worked - Mother	0.739	0.038
N			4448

Notes: Cells highlighted in blue have an absolute value of greater than 0.3. ‘(-)’ denotes a reverse-coded item.



**Table 25: Pearson Correlation Matrix for Mediation Factors**

	Parental Education	Child Acute Health	Anthropometrics	Early Investment	Parental Health	Mother Employment	Father Employment
Parental Education	1.000	0.014	0.095	0.398	0.117	0.025	0.032
Child Acute Health	0.014	1.000	-0.034	-0.045	0.145	0.076	0.033
Anthropometrics	0.095	-0.034	1.000	0.145	-0.008	-0.002	-0.054
Early Investment	0.398	-0.045	0.145	1.000	0.105	0.021	0.028
Parental Health	0.117	0.145	-0.008	0.105	1.000	0.016	0.172
Mother Employment	0.025	0.076	-0.002	0.021	0.016	1.000	0.114
Father Employment	0.032	0.033	-0.054	0.028	0.172	0.114	1.000

Notes: Uses all values available for each pairwise comparison between two factors.

### C Multiple Testing Adjustment

The False Discovery Rate (FDR) is the expected proportion of false rejections from testing a set of null hypotheses  $H_1, \dots, H_n$  i.e.  $E(H_i \text{ is true} | \text{Test rejects } H_i)$ .

For each of the main results tables in the paper, I calculate the [Benjamini and Hochberg \(1995\)](#) FDR-adjusted p-values using the procedure below based on the one seen in [Anderson \(2008\)](#).

1. For each of the coefficients  $\beta_j, j \in \{-2, -1, \dots, 15\}$ , get the corresponding p-values from the primary specification (denoted  $p_{-2}, p_{-1}, \dots, p_{15}$ )
2. Sort the p-values in ascending order so we have  $p_{(1)} \leq p_{(2)} \leq \dots \leq p_{(17)}$
3. To test whether  $\beta_j$  is rejected at the significance level  $\alpha$ :
  - (a) Find the largest  $k$  such that  $p_{(k)} \leq k \times \alpha / 17$
  - (b) Reject all hypotheses with a lower p-value than  $p_{(k)}$  i.e. the coefficients associated with the p-values  $p_{(1)}, \dots, p_{(k)}$
4. Repeat step 3 for all values  $\alpha$  (starting from 1.000, then 0.999, then 0.998 etc.). The q-value  $q_j$  for  $\beta_j$  is the smallest value of  $\alpha$  at which step 3 rejects the hypothesis for that coefficient.

The degree to which the [Benjamini and Hochberg \(1995\)](#) procedure is conservative varies based on the dependency structure of the tested p-values  $p_{-2}, p_{-1}, \dots, p_{15}$ . Suppose we set a significance threshold  $\alpha_0 = 0.05$  such that we reject the null for  $\beta_j$  if  $\beta_j \leq \alpha_0$ . If the p-values are independent or positively dependent, then the procedure above will ensure that  $FDR \leq \alpha_0$ , implying that we are controlling for the false discovery rate at the desired level of 5%. If, on the other hand, any of the p-values are negatively dependent, the procedure is overly conservative as it ensures that  $FDR \lesssim \alpha_0 / \ln(n) = \alpha_0 / \ln(17)$ . Since the right-hand side of this equation is smaller than  $\alpha_0$ , this means we are controlling for a smaller FDR than intended, and being overly conservative in inferring significant effects. Given my empirical setting, the p-values are likely to be positively dependent (for example, the effect of weather shocks at age 1 is likely to be positively correlated with the effect of weather shocks at age 2). But if they are negatively dependent, I am still successfully controlling for Type I errors.

## D Sample Selection

A relevant worry in my empirical setting is whether the main sample I am using is subject to a selection effect that could bias causal estimates of the impact of early-life weather shocks on adult psychological outcomes. Ideally, my main sample would be a random sample of all individuals born between 1988 and 2000 in rural Indonesia (in the provinces surveyed by the IFLS). However, we may be worried about two primary sources of selection that could lead to a non-random sample. First, there may be selective mortality. For example, it may be the case that extreme weather shocks such as drought lead to infant death, and that deaths are concentrated on infants with poor health. If poor health in infancy is correlated with adult cognitive ability, this would lead to a biased estimate in my main specification. Second, children who were born in IFLS households are only tracked when they move households if they were part of one of the original households from the very first wave, IFLS1. If children who were only incorporated into the IFLS as a result of later waves non-randomly migrate out of their households, the sample of individuals in my main sample will be similarly non-random.

In order to test for such selection effects, I make use of the detailed pregnancy history taken of each woman above aged 15 in the IFLS survey. Using this data, I have a record of every pregnancy and birth for IFLS respondents, which amounts to 9184 records.<sup>51</sup> Approximately 55% of these records (5006) can be matched to individuals in the main sample used in the analysis. But the rest are missing from the main sample: for example, 16% (1515) die before the age of 5 or result in still birth, and so do not have their adult psychological outcomes measured. If this attrition is non-random and correlated with early life weather shocks, this could lead to biased estimates in my main analysis. The test proposed will not detect all sources of attrition bias: in particular, it does not account for any individuals who are not recorded in the pregnancy history data, which could result from a mothers' recall bias or refusal to respond. On the other hand, the data available give me a unique opportunity to test for selective mortality directly. Table 26 shows the results of regressing measures of early life SPEI growing season on indicators for attrition, controlling as for the main analysis for birth-year and birth-district fixed effects.

In columns (1) and (2), I test whether early life SPEI affects the probability of dying before the age of 5 (or resulting in still birth). It appears that in utero SPEI may have a small positive effect on the probability of dying before age 5.<sup>52</sup> The coefficient of 0.024 on  $SPEI_{t-1}$  suggests that an increase in SPEI of 1 standard deviation leads to an increase in the probability of dying before aged 5 by 2.4 percentage points, and the coefficient is significant at the 10% level. The direction of this effect is surprising given the main results in the paper: since SPEI is positively associated with household economic wellbeing and child anthropometrics, we would expect higher SPEI to have a negative impact on infant mortality. On the other hand, there may be non-economic explanations for such a result. For example, the impact shown here could possibly be explained by a change in the disease environment in utero. For example, wetter environments (higher SPEI) may lead to higher incidence of malaria and an increased risk of infant death (Kudamatsu et al., 2012). If in utero SPEI does indeed have a positive effect on the probability of infant death, this would likely lead to a positive bias on the results found in my main analysis: I would be missing weaker children (who are likely to have lower cognitive ability) that are subject to high SPEI from my analysis. On the other hand, the results from column (1) and the multiple testing adjustment suggest that the true effect may not be significantly different from 0, and the coefficient on  $SPEI_{t-1}$  is itself relatively small. While I cannot rule out selection bias due to mortality, it is unlikely to explain the large effects seen on adult cognitive skills in Section 5.1.

In columns (3) and (4), I run a similar test to see whether early life SPEI affects the overall probability to be included in my main sample. No coefficients on early life SPEI are significant at the 10% level, suggesting no evidence for any selective inclusion in the final sample. Despite the potential selective

<sup>51</sup>I match pregnancies to districts using data on the mother's migration history.

<sup>52</sup>This appears to be driven mostly by an increase in the probability of a still birth.

mortality seen in columns (1) and (2), I cannot reject that there is no selection effect of  $SPEI_{t-1}$  overall, implying that selection bias in the sample is unlikely to drive the main results in the paper.

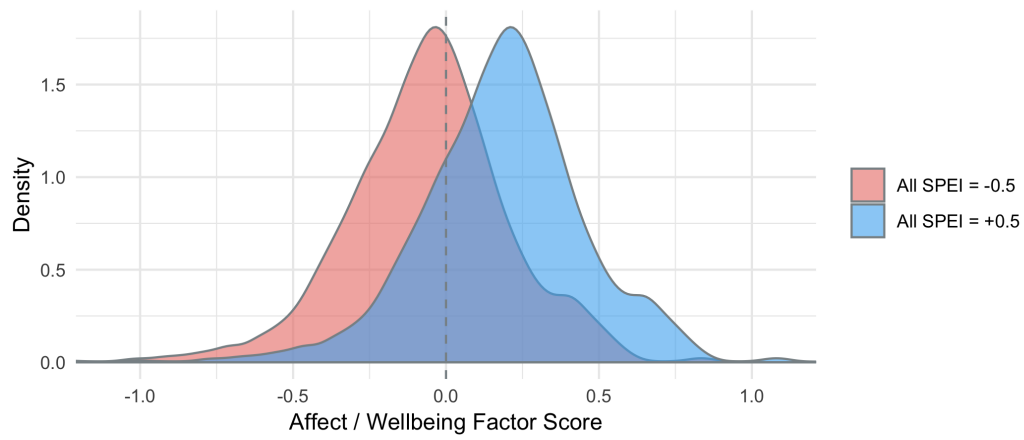
**Table 26: Testing for sample selection**

	<i>Dep Var: Died before age 5</i>			<i>Dep Var: Included in Main Sample</i>		
	(1)	(2)	q value (2)	(3)	(4)	q value (4)
Average SPEI growing season t - 1 to t + 4	0.022 (0.044)			0.024 (0.060)		
SPEI growing season in t - 1		0.024* (0.013)	0.365		-0.015 (0.031)	0.898
SPEI growing season in t		0.002 (0.014)	0.916		0.017 (0.018)	0.898
SPEI growing season in t + 1		0.006 (0.017)	0.916		0.004 (0.028)	0.898
SPEI growing season in t + 2		-0.001 (0.010)	0.916		0.020 (0.019)	0.898
SPEI growing season in t + 3		-0.011 (0.014)	0.916		0.004 (0.022)	0.898
SPEI growing season in t + 4		0.008 (0.014)	0.916		-0.011 (0.020)	0.898
Birth Year Fixed Effects	Yes	Yes		Yes	Yes	
Birth District Fixed Effects	Yes	Yes		Yes	Yes	
Additional Controls	No	No		No	No	
F statistic		1.405			1.067	
p-value for F-test		0.215			0.384	
Adj. R <sup>2</sup>	0.026	0.026		0.140	0.140	
N	9184	9184		9184	9184	

*Notes:* Sample includes all children born between 1988 and 2000 whose mother was an IFLS respondent and gave a pregnancy history in at least one IFLS wave. “Died before age 5” is an dummy variable that takes the value 1 if pregnancy ended in still birth or miscarriage, or if child was born and died before 5th birthday. “Included in Main Sample” is an indicator variable for whether child is in the main sample used in the paper (see Table 2 and surrounding discussion). \*\*\* significant at the 1% level; \*\* significant at the 5% level; \* significant at the 10% level. Standard errors are reported in parentheses and are two-way clustered by birth year and birth district (Cameron et al., 2011). The q-value is the FDR-adjusted p-value that is calculated according to the process in Benjamini and Hochberg (1995). The F statistic and p-value is for the hypothesis  $\beta_{-1} = \dots = \beta_4 = 0$  i.e. that all coefficients on SPEI growing season are 0. No survey weights are used for the models seen in this table.

## E Other Tables and Figures

**Figure 17:** Predicted change in “affect / wellbeing” factor as a result of a 1 standard deviation increase in SPEI at all ages



Notes: Distribution of affect / wellbeing factor when predicting using the model seen in Table 6 Column (2). The red distribution shows the results if all values of SPEI growing season are set to -0.5; the blue distribution shows the results if all values of SPEI growing season are set to +0.5.

**Table 27:** Effect of SPEI growing season on affect/wellbeing factor score of individual born in  $t$  by gender

	(1) Female	q value of (1)	(2) Male	q value of (2)
SPEI growing season in $t - 2$	0.122** (0.059)	0.168	-0.078 (0.069)	0.92
SPEI growing season in $t - 1$	0.144*** (0.056)	0.058	-0.018 (0.062)	0.92
SPEI growing season in $t$	0.086 (0.057)	0.337	0.025 (0.083)	0.92
SPEI growing season in $t + 1$	-0.069 (0.050)	0.374	0.032 (0.107)	0.92
SPEI growing season in $t + 2$	0.163*** (0.049)	0.017	-0.007 (0.063)	0.949
SPEI growing season in $t + 3$	0.043 (0.072)	0.809	-0.083 (0.075)	0.92
SPEI growing season in $t + 4$	0.127* (0.071)	0.261	-0.021 (0.053)	0.92
SPEI growing season in $t + 5$	-0.044 (0.079)	0.809	0.177** (0.072)	0.24
SPEI growing season in $t + 6$	0.006 (0.071)	0.983	0.085* (0.044)	0.482
SPEI growing season in $t + 7$	0.043 (0.056)	0.731	-0.004 (0.060)	0.949
SPEI growing season in $t + 8$	-0.002 (0.099)	0.983	0.034 (0.075)	0.92
SPEI growing season in $t + 9$	-0.008 (0.072)	0.983	-0.038 (0.082)	0.92
SPEI growing season in $t + 10$	0.119 (0.076)	0.337	0.086 (0.100)	0.92
SPEI growing season in $t + 11$	-0.094 (0.074)	0.411	0.006 (0.086)	0.949
SPEI growing season in $t + 12$	0.023 (0.076)	0.913	0.102 (0.104)	0.92
SPEI growing season in $t + 13$	-0.028 (0.077)	0.913	0.105 (0.088)	0.92
SPEI growing season in $t + 14$	-0.154*** (0.053)	0.033	-0.072 (0.090)	0.92
SPEI growing season in $t + 15$	0.065 (0.068)	0.611	0.021 (0.072)	0.92
Born during the growing season (=1)	0.109** (0.049)	-	0.033 (0.043)	-
Birth Year Fixed Effects	Yes		Yes	
Birth District Fixed Effects	Yes		Yes	
Additional Controls	Yes		Yes	
F statistic	2.103		1.202	
p-value for F-test	0.008		0.266	
Adj. R <sup>2</sup>	0.032		0.021	
N	2456		1951	

Notes: \*\*\* significant at the 1% level; \*\* significant at the 5% level; \* significant at the 10% level. Standard errors are reported in parentheses and are two-way clustered by birth year and birth district (Cameron et al., 2011). Affect/wellbeing factor score is internally standardised, so the overall sample mean (with both female and male) is 0 by construction. Additional controls are included in both regressions, and include dummies for the education level of the most-educated parent, dummies for religion, and a dummy for gender. Both models use the individual-level attrition-corrected weights provided in the IFLS data. The q-value is the FDR-adjusted p-value that is calculated according to the process in Benjamini and Hochberg (1995). F-statistic and F-test are for the joint hypothesis test  $H_0 : \beta_{-2} = \dots = \beta_{15} = 0$ , i.e. all the coefficients on SPEI growing season are 0.

**Table 28:** Effect of SPEI growing season on personality factor score of individual born in  $t$  by gender

	(1) Female	q value of (1)	(2) Male	q value of (2)
SPEI growing season in $t - 2$	-0.080 (0.070)	0.872	0.054 (0.079)	0.925
SPEI growing season in $t - 1$	-0.023 (0.068)	0.941	0.014 (0.067)	0.925
SPEI growing season in $t$	-0.058 (0.069)	0.872	0.066 (0.073)	0.925
SPEI growing season in $t + 1$	0.006 (0.086)	0.941	0.098 (0.073)	0.925
SPEI growing season in $t + 2$	-0.045 (0.057)	0.872	-0.026 (0.108)	0.925
SPEI growing season in $t + 3$	-0.010 (0.059)	0.941	-0.020 (0.078)	0.925
SPEI growing season in $t + 4$	0.049 (0.091)	0.941	0.045 (0.086)	0.925
SPEI growing season in $t + 5$	-0.133* (0.072)	0.591	-0.056 (0.077)	0.925
SPEI growing season in $t + 6$	-0.099 (0.083)	0.872	0.053 (0.120)	0.925
SPEI growing season in $t + 7$	0.006 (0.070)	0.941	-0.176*** (0.065)	0.12
SPEI growing season in $t + 8$	0.020 (0.067)	0.941	0.014 (0.084)	0.925
SPEI growing season in $t + 9$	-0.041 (0.065)	0.941	0.006 (0.065)	0.928
SPEI growing season in $t + 10$	-0.019 (0.052)	0.941	0.019 (0.092)	0.925
SPEI growing season in $t + 11$	-0.058 (0.055)	0.872	-0.020 (0.078)	0.925
SPEI growing season in $t + 12$	-0.116 (0.077)	0.801	0.090 (0.111)	0.925
SPEI growing season in $t + 13$	0.008 (0.065)	0.941	-0.032 (0.104)	0.925
SPEI growing season in $t + 14$	-0.114** (0.057)	0.591	0.033 (0.082)	0.925
SPEI growing season in $t + 15$	0.066 (0.074)	0.872	-0.189** (0.079)	0.152
Born during the growing season (=1)	0.068 (0.048)	-	0.018 (0.049)	-
Birth Year Fixed Effects	Yes		Yes	
Birth District Fixed Effects	Yes		Yes	
Additional Controls	Yes		Yes	
F statistic	1.368		1.378	
p-value for F-test	0.153		0.150	
Adj. R <sup>2</sup>	0.023		0.042	
N	2456		1951	

Notes: \*\*\* significant at the 1% level; \*\* significant at the 5% level; \* significant at the 10% level. Standard errors are reported in parentheses and are two-way clustered by birth year and birth district (Cameron et al., 2011). Personality factor score is internally standardised, so the overall sample mean (with both female and male) is 0 by construction. Additional controls are included in both regressions, and include dummies for the education level of the most-educated parent, dummies for religion, and a dummy for gender. Both models use the individual-level attrition-corrected weights provided in the IFLS data. The q-value is the FDR-adjusted p-value that is calculated according to the process in Benjamini and Hochberg (1995). F-statistic and F-test are for the joint hypothesis test  $H_0 : \beta_{-2} = \dots = \beta_{15} = 0$ , i.e. all the coefficients on SPEI growing season are 0.

**Table 29:** Effect of SPEI growing season on mental health factor score of individual born in  $t$  by gender

	(1) Female	q value of (1)	(2) Male	q value of (2)
SPEI growing season in $t - 2$	0.012 (0.058)	0.887	0.036 (0.081)	0.931
SPEI growing season in $t - 1$	0.078* (0.041)	0.887	0.070 (0.056)	0.931
SPEI growing season in $t$	0.016 (0.046)	0.887	-0.022 (0.045)	0.931
SPEI growing season in $t + 1$	-0.011 (0.050)	0.887	0.005 (0.086)	0.962
SPEI growing season in $t + 2$	-0.027 (0.043)	0.887	0.012 (0.070)	0.962
SPEI growing season in $t + 3$	-0.031 (0.054)	0.887	0.019 (0.067)	0.931
SPEI growing season in $t + 4$	0.015 (0.075)	0.887	0.041 (0.069)	0.931
SPEI growing season in $t + 5$	-0.075 (0.077)	0.887	0.076 (0.076)	0.931
SPEI growing season in $t + 6$	-0.037 (0.073)	0.887	0.037 (0.081)	0.931
SPEI growing season in $t + 7$	0.067 (0.055)	0.887	0.048 (0.080)	0.931
SPEI growing season in $t + 8$	-0.056 (0.062)	0.887	0.027 (0.068)	0.931
SPEI growing season in $t + 9$	0.017 (0.066)	0.887	0.004 (0.090)	0.962
SPEI growing season in $t + 10$	0.034 (0.059)	0.887	-0.028 (0.045)	0.931
SPEI growing season in $t + 11$	-0.069 (0.058)	0.887	-0.039 (0.068)	0.931
SPEI growing season in $t + 12$	-0.009 (0.066)	0.887	0.049 (0.058)	0.931
SPEI growing season in $t + 13$	0.043 (0.067)	0.887	-0.023 (0.077)	0.931
SPEI growing season in $t + 14$	0.033 (0.067)	0.887	0.052 (0.063)	0.931
SPEI growing season in $t + 15$	0.082 (0.067)	0.887	0.079 (0.085)	0.931
Born during the growing season (=1)	-0.037 (0.038)	-	0.083* (0.049)	-
Birth Year Fixed Effects	Yes		Yes	
Birth District Fixed Effects	Yes		Yes	
Additional Controls	Yes		Yes	
F statistic	0.891		0.585	
p-value for F-test	0.590		0.907	
Adj. R <sup>2</sup>	0.064		0.028	
N	2763		2350	

Notes: \*\*\* significant at the 1% level; \*\* significant at the 5% level; \* significant at the 10% level. Standard errors are reported in parentheses and are two-way clustered by birth year and birth district (Cameron et al., 2011). Mental health factor score is internally standardised, so the overall sample mean (with both female and male) is 0 by construction. Additional controls are included in both regressions, and include dummies for the education level of the most-educated parent, dummies for religion, and a dummy for gender. Both models use the individual-level attrition-corrected weights provided in the IFLS data. The q-value is the FDR-adjusted p-value that is calculated according to the process in Benjamini and Hochberg (1995). F-statistic and F-test are for the joint hypothesis test  $H_0 : \beta_{-2} = \dots = \beta_{15} = 0$ , i.e. all the coefficients on SPEI growing season are 0.



## References

- Adair, L. S., Fall, C. H., Osmond, C., Stein, A. D., Martorell, R., Ramirez-Zea, M., Sachdev, H. S., Dahly, D. L., Bas, I., Norris, S. A., Micklesfield, L., Hallal, P., and Victora, C. G. (2013). Associations of linear growth and relative weight gain during early life with adult health and human capital in countries of low and middle income: findings from five birth cohort studies. *The Lancet*, 382(9891):525–534.
- Adhvaryu, A. and Fenske, J. (2016). Early Life Circumstance and Adult Mental Health. *Journal of Political Economy*, page Forthcoming.
- Alderman, H., Hoddinott, J., and Kinsey, B. (2006). Long term consequences of early childhood malnutrition. *Oxford Economic Papers*, 58(3):450–474.
- Almlund, M., Duckworth, A. L., Heckman, J., and Kautz, T. (2011). Personality Psychology and Economics. In *Econometrica*, volume 78, pages 1–181. Elsevier B.V.
- Almond, D., Currie, J., and Duque, V. (2018). Childhood Circumstances and Adult Outcomes: Act II. *Journal of Economic Literature*, 56(4):1360–1446.
- Anderson, M. 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):1481–1495.
- Attanasio, O., Cattan, S., Fitzsimons, E., Meghir, C., and Rubio-Codina, M. (2015). Estimating the Production Function for Human Capital: Results from a Randomized Control Trial in Colombia. *SSRN Electronic Journal*, (1987).
- Attanasio, O., Cattan, S., Fitzsimons, E., Meghir, C., and Rubio-Codina, M. (2017). APPENDIX: Estimating the Production Function for Human Capital: Results from a Randomized Controlled Trial in Colombia. *Ssrn*, 03(560450).
- Attanasio, O. P., Fernandez, C., Fitzsimons, E. O. A., Grantham-McGregor, S. M., Meghir, C., and Rubio-Codina, M. (2014). Using the infrastructure of a conditional cash transfer program to deliver a scalable integrated early child development program in Colombia: cluster randomized controlled trial. *BMJ*, 349(sep29 5):g5785–g5785.
- Baez, J., Fuente, A. D., and Santos, I. (2010). Do Natural Disasters Affect Human Capital? An Assessment Based on Existing Empirical Evidence. *IZA Discussion Paper*, 5164(5164).
- Barham, T., Macours, K., and Maluccio, J. A. (2013). Boys' Cognitive Skill Formation and Physical Growth: Long-Term Experimental Evidence on Critical Ages for Early Childhood Interventions. *American Economic Review*, 103(3):467–471.
- Barham, T., Macours, K., and Maluccio, J. A. (2018). Experimental Evidence of Exposure to a Conditional Cash Transfer During Early Teenage Years: Young Women's Fertility and Labor Market Outcomes. (August).
- Barker, D. J. P. (1995). Fetal origins of coronary heart disease. *BMJ*, 311(6998):171–174.
- Bazzi, S. and Gudgeon, M. (2016). Local Government Proliferation, Diversity, and Conflict.
- Beguera, S. and Vicente Serrano, S. M. (2017). SPEIbase v.2.5 Global 04-month 1901-2015 SPEI, 2017.
- Beguera, S., Vicente-Serrano, S. M., Reig, F., and Latorre, B. (2014). Standardized precipitation evapotranspiration index (SPEI) revisited: Parameter fitting, evapotranspiration models, tools, datasets and drought monitoring. *International Journal of Climatology*, 34(10):3001–3023.
- Benjamini, Y. and Hochberg, Y. (1995). Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society, B(57)*:289–300.
- BenYishay, A. (2013). Childhood Shocks and Trust in Adulthood: Impacts of Early-Life Rainfall in Sub-Saharan Africa.
- Bharati, T., Chin, S., and Jung, D. (2017). Recovery from an early life shock through improved access to schools: Evidence from Indonesia. pages 1–65.
- Binswanger, H. P. (1980). Attitudes toward Risk: Experimental Measurement in Rural India. *American Journal of Agricultural Economics*, 62(3):395.
- Black, S. E., Bütikofer, A., Devereux, P. J., and Salvanes, K. G. (2018). This Is Only a Test? Long-Run and Intergenerational Impacts of Prenatal Exposure to Radioactive Fallout. *The Review of Economics and Statistics*, pages 1–16.
- Bomblies, A. (2012). Modeling the role of rainfall patterns in seasonal malaria transmission. *Climatic Change*, 112(3-4):673–685.
- Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2011). Robust Inference With Multiway Clustering. *Journal of Business & Economic Statistics*, 29(2):238–249.
- Cattell, R. B. (1966). The Scree Test For The Number Of Factors. *Multivariate Behavioral Research*, 1(2):245–276.
- Chuang, Y. and Schechter, L. (2015). Stability of experimental and survey measures of risk, time, and social preferences: A review and some new results. *Journal of Development Economics*, 117:151–170.
- Crookston, B. T., Dearden, K. A., Alder, S. C., Porucznik, C. A., Stanford, J. B., Merrill, R. M., Dickerson, T. T., and Penny, M. E. (2011). Impact of early and concurrent stunting on cognition. *Maternal & Child Nutrition*, 7(4):397–409.
- Cunha, F., Heckman, J., and Schennach, S. (2010). Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Econometrica*, 78(3):883–931.
- Cunha, F. and Heckman, J. J. (2009). The economics and psychology of inequality and human development. *Journal of the European Economic Association*, 7(2-3):320–364.
- Currie, J. and Almond, D. (2011). Human capital development before age five. In *Handbook of Labor Economics*, volume 4, pages 1315–1486. Elsevier B.V.
- Eriksson, J. G., Kajantie, E., Osmond, C., Thornburg, K., and Barker, D. J. (2010). Boys live dangerously in the womb. *American Journal of Human Biology*, 22(3):330–335.

- Evans, G. W. and English, K. (2002). The Environment of Poverty: Multiple Stressor Exposure, Psychophysiological Stress, and Socioemotional Adjustment. *Child Development*, 73(4):1238–1248.
- Frederick, W. H., Worden, R. L., and of Congress. Federal Research Division, L. (2011). *Indonesia: A Country Study*. Area handbook series. Federal Research Division, Library of Congress.
- Fryer, R., Levitt, S., and List, J. (2015). Parental Incentives and Early Childhood Achievement: A Field Experiment in Chicago Heights. Technical report, National Bureau of Economic Research, Cambridge, MA.
- Glewwe, P. and King, E. M. (2001). The Impact of Early Childhood Nutritional Status on Cognitive Development: Does the Timing of Malnutrition Matter? *The World Bank Economic Review*, 15(1):81–113.
- Grantham-McGregor, S., Cheung, Y. B., Cueto, S., Glewwe, P., Richter, L., and Strupp, B. (2007). Developmental potential in the first 5 years for children in developing countries. *Lancet*, 369(9555):60–70.
- Harari, M. and La Ferrara, E. (2018). Conflict, Climate, and Cells: A Disaggregated Analysis. *The Review of Economics and Statistics*, 100(4):594–608.
- Heckman, J. and Cunha, F. (2007). The Technology of Skill Formation. *American Economic Review*, 97(2):31–47.
- Heckman, J., Pinto, R., and Savelyev, P. (2013). Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes. *American Economic Review*, 103(6):2052–2086.
- Heckman, J. J. (2007). The economics, technology, and neuroscience of human capability formation. *Proceedings of the National Academy of Sciences*, 104(33):13250–13255.
- Heckman, J. J. and Mosso, S. (2014). The Economics of Human Development and Social Mobility. *Ssrn*.
- Heckman, J. J., Stixrud, J., and Urzua, S. (2006). The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics*, 24(3):411–482.
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2):179–185.
- Huber, M. (2014). IDENTIFYING CAUSAL MECHANISMS (PRIMARILY) BASED ON INVERSE PROBABILITY WEIGHTING. *Journal of Applied Econometrics*, 29(6):920–943.
- John and Raven, J. (2003). Raven Progressive Matrices. In *Handbook of Nonverbal Assessment*, pages 223–237. Springer US, Boston, MA.
- Kahneman, D. and Deaton, A. (2010). High income improves evaluation of life but not emotional well-being. *Proceedings of the National Academy of Sciences*, 107(38):16489–16493.
- Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23(3):187–200.
- Knudsen, E. I. (2004). Sensitive Periods in the Development of the Brain and Behavior. *Journal of Cognitive Neuroscience*, 16(8):1412–1425.
- Krutikova, S. and Lilleør, H. B. (2015). Fetal Origins of Personality : Effects of early life circumstances on adult personality traits.
- Kudamatsu, M., Persson, T., and Strömberg, D. (2012). Weather and Infant Mortality in Africa. *C.E.P.R. Discussion Papers*, 9222.
- Kuklina, E. V., Ramakrishnan, U., Stein, A. D., Barnhart, H. H., and Martorell, R. (2006). Early childhood growth and development in rural Guatemala. *Early Human Development*, 82(7):425–433.
- Laajaj, R. and Macours, K. (2017). Measuring Skills in Developing Countries. *World Bank Policy Research Working Papers*, WBS 8000(January):82.
- Laajaj, R., Macours, K., Alejandro, D., Hernandez, P., Arias, O., Gosling, S., Rubio-codina, M., and Vakis, R. (2018). Challenges to Capture the Big Five Personality Traits in non-WEIRD populations.
- Ledesma, R. D. and Valecro-Mora (2007). Determining the number of factors to retain in EFA: an easy-to-use computer program for carrying out Parallel Analysis. *Practical Assessment, Research & Evaluation*, 12(2):2–11.
- Lee, C.-T., Zhang, G., and Edwards, M. C. (2012). Ordinary Least Squares Estimation of Parameters in Exploratory Factor Analysis With Ordinal Data. *Multivariate Behavioral Research*, 47(2):314–339.
- Leroy, J. L. and Frongillo, E. A. (2019). Perspective: What Does Stunting Really Mean? A Critical Review of the Evidence. *Advances in Nutrition*, 10(2):196–204.
- Levine, D. and Yang, D. (2014). The Impact of Rainfall on Rice Output in Indonesia. Technical report, National Bureau of Economic Research, Cambridge, MA.
- Lundeen, E. A., Behrman, J. R., Crookston, B. T., Dearden, K. A., Engle, P., Georgiadis, A., Penny, M. E., and Stein, A. D. (2014). Growth faltering and recovery in children aged 1–8 years in four low- and middle-income countries: Young Lives. *Public Health Nutrition*, 17(09):2131–2137.
- Maccini, S. and Yang, D. (2009). Under the Weather: Health, Schooling, and Economic Consequences of Early-Life Rainfall. *American Economic Review*, 99(3):1006–1026.
- Macours, K., Schady, N., and Vakis, R. (2012). Cash Transfers, Behavioral Changes, and Cognitive Development in Early Childhood: Evidence from a Randomized Experiment. *American Economic Journal: Applied Economics*, 4(2):247–273.
- Minnesota Population Center. (2018). Integrated Public Use Microdata Series, International: Version 7.1. Indonesia Spatially Harmonized Second-Level Geography (1970-2011).
- Mobarak, A. M. and Ali Akram, A. (2016). Seasonal migration to increase incomes of poor households in Bangladesh. *Bangladesh Priorities, Copenhagen Consensus Center*.
- Monfreda, C., Ramankutty, N., and Foley, J. A. (2008). Farming the planet: 2. Geographic distribution of crop areas,

- yields, physiological types, and net primary production in the year 2000. *Global Biogeochemical Cycles*, 22(1):1–19.
- Nilsson, J. P. (2017). Alcohol Availability, Prenatal Conditions, and Long-Term Economic Outcomes. *Journal of Political Economy*, 125(4):1149–1207.
- Pasha, M., Rockmore, M., and Tan, C. M. (2018). Positive Early Life Rainfall Shocks and Adult Mental Health. *Ssrn*, pages 1–31.
- Persson, P. and Rossin-Slater, M. (2018). Family Ruptures, Stress, and the Mental Health of the Next Generation. *American Economic Review*, 108(4-5):1214–1252.
- Pollitt, E., Gorman, K. S., Engle, P. L., Martorell, R., Rivera, J., Wachs, T. D., and Scrimshaw, N. S. (2006). Early Supplementary Feeding and Cognition: Effects over Two Decades. *Monographs of the Society for Research in Child Development*, 58(7):i.
- Radloff, L. S. (1977). The CES-D Scale. *Applied Psychological Measurement*, 1(3):385–401.
- Rammstedt, B. and Farmer, R. F. (2013). The impact of acquiescence on the evaluation of personality structure. *Psychological Assessment*, 25(4):1137–1145.
- Rammstedt, B. and John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, 41(1):203–212.
- RAND (1994). Indonesia Family Life Survey, Wave 1.
- RAND (1997). Indonesia Family Life Survey, Wave 2.
- RAND (2000). Indonesia Family Life Survey, Wave 3.
- RAND (2008). Indonesia Family Life Survey, Wave 4.
- RAND (2015). Indonesia Family Life Survey, Wave 5.
- Redelmeier, D. A., Katz, J., and Kahneman, D. (2003). Memories of colonoscopy: a randomized trial. *Pain*, 104(1):187–194.
- Rose, E. (1999). Consumption Smoothing and Excess Female Mortality in Rural India. *Review of Economics and Statistics*, 81(1):41–49.
- Sacks, W. J., Deryng, D., Foley, J. A., and Ramankutty, N. (2010). Crop planting dates: An analysis of global patterns. *Global Ecology and Biogeography*, 19(5):607–620.
- Sapolsky, R. M. (2017). *Behave: The Biology of Humans at Our Best and Worst*. Bodley Head Limited.
- Schady, N. (2011). Parents' education, mothers' vocabulary, and cognitive development in early childhood: Longitudinal evidence from Ecuador. *American Journal of Public Health*, 101(12):2299–2307.
- Singhal, S. (2018). Early life shocks and mental health: The long-term effect of war in Vietnam. *Journal of Development Economics*.
- Soto, C. J., John, O. P., Gosling, S. D., and Potter, J. (2008). The developmental psychometrics of big five self-reports: Acquiescence, factor structure, coherence, and differentiation from ages 10 to 20. *Journal of Personality and Social Psychology*, 94(4):718–737.
- Sudfeld, C. R., Charles McCoy, D., Danaei, G., Fink, G., Ezzati, M., Andrews, K. G., and Fawzi, W. W. (2015). Linear Growth and Child Development in Low- and Middle-Income Countries: A Meta-Analysis. *Pediatrics*, 135(5):e1266–e1275.
- Thompson, R. A. and Nelson, C. A. (2001). Developmental science and the media. Early brain development. *The American psychologist*, 56(1):5–15.
- VanderWeele, T. J. (2016). Mediation Analysis: A Practitioner's Guide. *Annual Review of Public Health*, 37(1):17–32.
- Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, 41(3):321–327.
- Venkataramani, A. S. (2012). Early life exposure to malaria and cognition in adulthood: Evidence from Mexico. *Journal of Health Economics*, 31(5):767–780.
- Vicente-Serrano, S. M., Beguería, S., and López-Moreno, J. I. (2010). A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index. *Journal of Climate*, 23(7):1696–1718.
- Victora, C. G., Adair, L., Fall, C., Hallal, P. C., Martorell, R., Richter, L., and Sachdev, H. S. (2008). Maternal and child undernutrition: consequences for adult health and human capital. *The Lancet*, 371(9609):340–357.
- White, M. D. (2018). The Problems with Measuring and Using Happiness for Policy Purposes. *SSRN Electronic Journal*.
- WHO and UNICEF (2009). WHO child growth standards and the identification of severe acute malnutrition in infants and children. pages 1–12.
- World Bank (2018). Rural Population in Indonesia.