Health and the economy: three essays

Halfdan Grangård

Dissertation for the degree of
Doctor of Philosophy in Economics at
The London School of Economics and Political Science

January 2011

Declaration

I certify that the thesis I have presented for examination for the MPhil/PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others.

Halfdan Grangård

Abstract

The main questions of this thesis are how a period of in utero malnutrition can impact the health of young children and their later development, and how job promotions can affect health.

In the first chapter I analyse to what extent the South-East Asian financial crisis affected the height of Indonesian children who had in utero exposure to the crisis. I find that they are significantly shorter than children who were exposed at later ages. There is a large difference in effect for urban and rural children. This finding helps attribute the detrimental health effects to the crisis and not other events which occurred during the period of analysis.

The second chapter exploits the exogenous shock of the crisis to analyse how early childhood height causes later cognitive development. I argue that this question should be analysed using instrumental variables. The results show a large and significant effect of early childhood height on cognitive ability and the use of instrumental variables changes the results significantly compared to OLS with or without fixed effects.

Lastly, I analyse how on the job promotions of British civil servants affect health. In a cross-section, the direction of causality is almost certainly two-way. I argue that the use of individual fixed effects will alleviate this concern. The results show a large, positive effect of a job promotion on health in the subsequent survey phase.

Acknowledgements

I would like to thank some of the people and institutions who have helped me in completing this thesis. First I would like to express my gratitude to my supervisors, Robin Burgess and Oriana Bandiera. Robin Burgess was vital in directing my work early in the process and has helped immensely with great ideas throughout. I thank Oriana Bandiera for all her help with providing feedback and ideas to all of the drafts and results I have presented to her.

I thank the London School of Economics for giving me the opportunity to work on a PhD and for waiving tuition fees.

The University of Oslo provided a home for me in Norway when I had to leave London due to health issues. I will be ever grateful to UiO and ESOP for providing an office and a stimulating environment in Oslo.

I would also like to thank my family for their support. Special thanks to my father, Erik Grangård, for his enthusiasm and the barrage of more and less relevant medical articles he has sent me the last few years.

Lastly, I would like to thank my wife, Solveig Kristinsdottir. Your unwavering support and belief has been of great help during this process.

Contents

List of figures

List of tables

Preface

The question of how individual health affects the person and the economy as a whole is an important one. Bad health impacts productivity negatively as the inflicted person will either miss work completely or be less effective while working. In the following, I will focus on early life causes of bad health. We know that babies in utero are particularly vulnerable to periods of poor nutrition (Barker, 1998). When a fetus is denied proper nutrition, the body will direct more of the blood flow towards particularly important organs like the brain and heart, while the skeletal development might suffer. This is known as fetal programming and has the potential to create path dependence even prior to birth. The literature has shown that there is a clear association between the height of young children and a number of markers of educational success, cognitive ability and productivity.

Early childhood health has been proposed as 'the smoking gun' behind the somewhat puzzling phenomenon termed 'the health gradient' (Case & Paxson, 2008). More affluent people have better health on average than individuals who are otherwise seemingly similar, but who have lower income. This phenomena was the surprising finding in the Whitehall I Study which interviewed a number of civil servants working in London, United Kingdom. It was surprising since this was a population who all received a pay check and were living in a country with well developed health infrastructure subsidised by the state. Since this finding, the health gradient has been found in virtually every country which has been analysed and also within individual states in the United States and other countries (Wilkinson, 2005). A number of explanations of the health gradient have been proposed in the literature. It has been suggested that the relatively poor engage in worse health behaviours (Smith et al., 1990) and that these could explain the differences in health. The results of Smith et al. (1990) show that health behaviours explains some of the gradient, but much is left to explore to fully explain the relationship between health and income; a relationship which runs all the way through society, from bottom to the very top. Case & Paxson (2008) suggests that small children will have worse cognitive ability during school age and that this lack of ability is penalized with lower salaries in the labour market. This mechanism could explain at least parts of the gradient puzzle if early childhood health is correlated with adult health as per the Barker hypothesis (Barker, 1998).

The health gradient is a feature of societies with free health care (Wilkinson, 2005) which makes it unlikely that the gradient is just a reflection of the higher purchasing power of wealthier individuals.

In the first chapter, I analyse the impact of poor intrauterine nutrition on height in early childhood. The Indonesian Family and Life Survey provides data on Indonesian families, mothers and babies. The South-East Asian financial crisis in the late 90's hit the affected countries without much warning. Of the affected countries in the region, the Indonesian experience was the most tragic partly due to structural weaknesses in the banking sector. After having experienced rapid growth in the preceding decades, real GDP fell by over 13% in 1998 and only recovered to pre-crisis levels in 2002. As a result of the fall in production, poverty increased rapidly. I attempt to analyse how this period of increased poverty affected the height of babies born during the crisis. The period leading up to the crisis was characterised by rapid growth and improvements in nutrition. These trends necessitate using estimation strategies that can eliminate trend effects. I employed a non-standard age group difference in differences approach. Children who experienced at least one trimester in utero during the crisis were relatively stunted compared to older children who lived through the crisis, but were too old to have had their heights affected by the crisis. These two age groups were compared to the same pre-crisis age groups. I found that, children who were in utero during the crisis suffered a height z-score loss of 0.28 points. This effect is large enough to be economically important. My results also indicate that urban children suffered the most which is in line with the distributional effects of the crisis which have been described by others (Pritchett et al., 2002; Sumarto et al., 1999). I also found that children born just prior to the crisis seem to have been affected to a much smaller degree than the children we assigned to the treatment group.

The second chapter exploits the Indonesian financial crisis to analyse how early childhood height causes cognitive development in school age. In the literature it is now well established that to analyse this question it is important to control for household- or maternal effects. Cognitive development and health is likely affected by the tastes, preferences and the level of nurturing of parents. Maternal fixed effects controls for such effects as long as these are fixed over time and children. The need for the use of instrumental variables is less established. If parents make nutritional adjustments based on expectations of future cognitive ability, OLS

estimates will be biased. There is also the potential for genetics or other biological processes to determine both cognitive development and height in early life. Both of these issues necessitate the use of instrumental variables. I exploit the crisis as an exogenous shock to early childhood height and the Indonesian Family and Life Survey data to analyse how height prior to the age of five years causes cognitive ability at school age. As far as I am aware, there has been no other attempt to analyse this question by employing a household fixed effects approach coupled with instrumental variables. The high data requirements are likely to be the limiting factor for such analysis. My results show that the causal effect of early childhood height on school age cognitive ability is large.

These results add to the results in the first chapter and highlight the importance of protecting pregnant women and their babies during times of hardship. This group is particularly vulnerable and neglecting them could not only have a negative impact on the individuals, but the on economy as a whole. These children could grow up without fulfilling their true potential leading to productivity losses for the rest of society.

In the last chapter, I employ a new approach to understand the health gradient puzzle. Wilkinson (2005) suggests that given the income distribution, health distribution will follow as a result of the insulin stress responses caused by the relative socioeconomic status in society. People at the bottom of the social ladder will experience stress responses leading to worse health. It is hard to experimentally alter the socioeconomic status of humans. I analyse how a promotion at work in the Whitehall II population affects Self-Assessed Health (SAH). SAH has been shown to both predict health and to be highly correlated with clinical measures of health. Promotions are highly likely to be partly caused by health in a cross-section. Healthier individuals are more likely to be promoted as documented in Boyce & Oswald (2008). We argue that it is less likely that for a given individual, a promotion will occur in a period of especially good health. I use individual fixed effects to analyse this question and find that a promotion between the prior and the current phase of the Whitehall II data is associated with better health. The effects are relatively large and statistically significant. The data shows that it is unlikely that this association is caused by increased income which follows a promotion. It seems that Wilkinson's model of stress responses is a potential explanation of the results we find.

We cannot measure stress responses of the subjects of the Whitehall II Survey. I attribute the results of the last chapter to the stress responses caused by the relative status in the Civil Service. This conclusion is reached by a process of elimination and only direct research on stress levels and cortisol responses could firmly resolve the underlying mechanisms at work.

1 Health effects of the South-East Asian financial crisis on young children in Indonesia

1.1 Introduction

Indonesia was hit hard by the South-East Asian financial crisis in the late 1990's. After having experienced rapid economic growth since the 1960's, real GDP fell by 13.1% in 1998 and did not rebound to pre-crisis levels until 2002[1]. The crisis in Indonesia commenced in the summer of 1997 when the Indonesian Rupiah came under severe pressure and the strains of currency speculation spread into the real economy as outlined in Section 1.2. The fall in GDP quickly pushed people into poverty as reported in Suryahadi et al. (2002).

We will analyse the impact of the crisis on the health of Indonesian children. In particular, we will focus on changes in height of Indonesian children who were exposed to the crisis in utero. Thomas et al. (1999) analysed to what extent the crisis had negative consequences for the health of children. These results were preliminary and they found no significant change in height-for-age of Indonesian children between 1997 and 1998. Cameron (2002) performs a similar analysis with data reaching 1999. The 1997 data points serve as the pre-crisis comparison in both papers. We wish to expand on the results presented in Thomas et al. (1999) and Cameron (2002). Both papers preface their results by acknowledging that their data is unlikely to capture the medium-term anthropometric changes that could occur as a result of the crisis. The two papers do not focus specifically on the youngest children who are likely the most vulnerable to malnutrition. We will employ a difference in differences estimator that focus specifically on children who experienced the crisis as very young and that had at least four months of in utero exposure to the crisis. By using this estimator, we attempt to answer a slightly different question to the two publications mentioned above. They analyse how average height-for-age changed from before the crisis to a period within the crisis. We attempt to provide answers to the question of how the very youngest children

[1] World Bank Development Indicators

were affected by the crisis compared how they might have fared in the absence of a crisis. The difference in differences estimator takes the trend in our outcome variable into account. According to our data, average z-score of Indonesian children showed increased in the period leading up to the crisis. In absence of the crisis, we would expect average height-for-age to improve between 1997 and 1999.

Our estimator will compare the changes in height-for-age of young children (aged 0-3 years) between 1993 and 2000 to the changes of adolescents (13-18) over the same period. Research (Martorell, 1995; Hoddinott & Kinsey, 2001) has shown that only very young children experience growth retardation as a result of prolonged periods of hunger. A growing body of literature in economics[2] and medicine[3] has recently demonstrated the extreme vulnerability of infants and very young children to a lack of nutrition especially during gestation. The in utero environment has been shown to have a long lasting effect on anthropometrics. As a result, children who were in utero during the crisis need special attention if we want to examine the health effects of the crisis in Indonesia on children. This paper will focus on children who were in utero during the crisis and the results will show a significant impact of the crisis on their health. We will focus on height as a latent variable of health, an approach that is well established both in economics and medicine. We care about the height of children as not only is childhood height highly correlated with adult height (Barker, 1998), but intra-uterine environment is significantly associated with educational attainment (Maccini et al, 2007; Almond, 2006; Meng & Qian, 2006), adult wages (Almond, 2006), labour supply (Meng & Qian, 2006), socioeconomic status (Almond, 2006) and health of the subsequent generation (Almond & Qian, 2003). These results document a pathway along which the financial crisis in Indonesia potentially could have long-run effects on the economy as a result of a fall in human capital of the children that were born during the crisis.

Our results show that when we focus our attention on children that had in utero exposure to the financial crisis, we find a significant decrease in health as measured by height-for-age. We find no such effect for children born just prior to the onset of the crisis. We also demonstrate that children in urban households were hit hard by the crisis, while rural children were less affected. The crisis in Indonesia was

[2] Maccini & Yang (2007), Almond & Chay (2003), Banerjee et al. (2007), Almond (2006), Meng & Qian (2006), Hoddinott & Kinsey (2001)

[3] For a summary, see Barker (1998)

a result of a collapse in the credit markets which lead to soaring inflation, lower wages and increased unemployment rates. The effects of the crisis affected urban wage earners to a much larger extent than rural households living off the land. As we compare changes over the period 1993-2000, we cannot be positive that the results we find are the result of the crisis only. We do however, find that the distribution of the detrimental effects of the crisis follows an urban/rural divide with urban children taking the brunt of the hardship. This result makes us able to attribute the fall in height-for-age to the crisis with greater confidence.

We will start by giving a brief background of the crisis and the medical research linking in-utero environment and early childhood health. Section 1.3 describes the data and our estimation strategy. Section 1.4 presents our main results and analyse distributional effects and birth timing effects. The last section concludes.

1.2 Background

At the end of 1996, Indonesia had experienced rapid economic growth over the last 30 years. Average real GDP growth of Indonesia over the period 1961-1996 was 6.4%[4]. The country was hailed as an example to other developing countries. In a speech given by the Managing Director of the IMF, Michel Camdessus, in Jakarta November 1996, he described the economic performance of Indonesia and the other ANSEAN countries as 'an inspiration to many other countries around the world'[5]. After decades of astonishing economic growth, the collapse that followed in the next couple of years was difficult to predict. In 1998, the real GDP of Indonesia fell by 13.1%[6]. The real GDP growth in the 1980's and 1990's is shown in Figure 1.1. Pressure on the Thai Baht spread to other countries in the region with currencies supported by the US dollar. In Indonesia the managed float of the Rupiah came under pressure in the second quarter of 1997. As interest rates increased to support the currency, structural problems in the banking and credit sector became apparent and a substantial amount of bad debt resulted in several banks going under. The whole banking sector came under immense pressure and credit became very hard to obtain. As the Rupiah was allowed to float in August

[4] World Bank Development Indicators
[5] Address by Michael Camdessus Managing Director of the International Monetary Fund at the Conference on "Macroeconomic Issues Facing ASEAN Countries"
[6] World Bank Development Indicators

1997 and quickly depreciated, a number of firms with foreign currency denominated debt, became insolvent and bankrupt.

These problems in the credit market coupled with a reduction in government spending, resulted in a substantial fall in production. The fall in government spending was one of the conditions of an IMF rescue package introduced in October 1997. For the purpose of the following analysis, the important feature of the fall in production is not the mechanisms behind the fall, but rather that the crisis was very difficult to predict and as such people had no reason to take precautions before the onset of the crisis. As GDP fell, people were driven into poverty. Not only did production fall, but accelerating inflation made it increasingly expensive to buy food. Inflation in the food sector was especially severe with food prices increasing by 160% between February 1996 and February 1999 compared to an inflation rate of non-food items of 81% (Suryahadi et al., 2002). Suryahadi et al. (2002) employ a number of data sources to estimate the evolution of poverty rates in Indonesia during the crisis. Their headcount measure shows a striking spike between November 1997 and late 1999. Starting in October 1997, the headcount poverty measure increased by more than 150% in the following 12 months and reached its peak in October 1998. Their sample period ends in August 1999 when the poverty count had fallen by 42% from its high point, but was still about 50% higher than at the onset of the crisis.

The extreme contraction of GDP and the ensuing spike in poverty rates, provides the natural experiment needed for the following analysis. The so-called Barker hypothesis has drawn recent attention in economics[7] and provides a possible medical justification for the main hypothesis in this paper. Simplified, the Barker hypothesis states that babies that are undernourished in utero, will experience health problems later in life, particularly an increased risk of diabetes. This initial discovery has lead to a body of work in medicine that shows the importance of in utero environment on the size of the baby at birth. 'Although the growth of a fetus is influenced by its genes, studies in humans and animals suggest that it is usually limited by the nutrients and oxygen it receives' (Barker, 1998). As the supply of nutrients to the baby falls, the body adjusts blood flow to develop particularly important organs like the brain and by doing so can retard skeletal

[7] ex. Almond & Chay (2003), Banerjee et al. (2007), Almond (2006), Meng & Qian (2006)

growth. This is known as fetal programming. Both maternal diet during pregnancy and maternal body composition determine the placental supply of nutrients and possibly the latter more than the former. '...even extreme restrictions in mother's food intake during pregnancy have only modest effects on birth weight. From this it cannot be concluded that fetal growth is not regulated by its nutrient supply. Rather it suggests that maternal nutrient intakes during pregnancy have relatively small effects on birth size, which may depend more on mother's nutritional state before pregnancy, that is the turnover of her protein and fatty acid stores in her muscle and fat' (Barker, 1998). Studies analysing the effect of maternal nutrition on birth weight[8] show an effect, but it is usually rather small. These results suggest that it is not necessarily just the poverty rates in Indonesia at a given point in time that will drive our results, but also poverty rates in the time span leading up to the period of analysis. The former would directly influence the mother's food intake, while the latter would determine the body composition of the mother.

The Barker hypothesis is not the only possible mechanism linking the crisis and early childhood stunting. There is some evidence suggesting that maternal stress during pregnancy can cause health problems for the children. The human body responds to stressors by activating the hypothalamus-pituitary-adrenal (HPA) cortex and the sympathetic nervous system. A number of hormones, including cortisol, are released into the blood. Individuals respond to stress very differently depending on many factors including genes and the support system they are surrounded by. Animal experiments have shown that stress during pregnancy can cause growth retardation and lower birth weight of the offspring (deCatanzaro & Macniven, 1992). Such experiments are usually performed by introducing stressors like capture, noise, introduction of a strange male and other elements that can be experimentally varied. These experiments are obviously not performed on humans and the complicated relationship between stress and the human response makes the research on humans less than conclusive. However, some studies have found a negative relationship between maternal stress and birth length of the baby. Dancause et al. (2011) found that mothers who were exposed to an ice storm during early to mid pregnancy had children who were shorter on average than the reference population who had not been exposed. Torche (forthcoming) exploits

[8] Lechtig & Klein (1981), Smith (1947), Stein et al. (1975), Dean (1951), Antonov (1947)

a major earthquake as a natural experiment to test how stress during pregnancy affects the newborn baby. Using a diffence-in-differences approach she finds a significant decline in birth weight for children with mothers who were exposed during the first trimester.

It seems like methodically sound research in this area is in its infancy. Some of the more recent contributions find negative effects on child growth of maternal stress during pregnancy. Whether the Barker hypothesis or the maternal stress story explains our results, is difficult to glean from our data. We lack good nutritional data for our sample prior and during the crisis and maternal stress levels are unknown. The crisis could have caused both under nutrition and maternal stress as mothers could worry about their own and their baby's future. Research is stronger on the link between malnutrition during gestation and stunting, but we cannot conclude that this is the only, or even the major cause of our results. The policy implications could change depending on which story we believe, but our results do suggest a stronger focus on pregnant women during crisis times.

1.3 Data and estimation strategy

1.3.1 Data

The data used in this paper is exclusively collected from the Indonesian Family Life Survey (IFLS) of households and communities. IFLS[9] is an ongoing longitudinal survey which first round was collected in 1993 (IFLS1). The second round was conducted in 1997 (IFLS2). The third published round of the IFLS was collected in 2000 (IFLS3) and we will use data from IFLS1 and IFLS3. IFLS1 interviewed 7 200 households and 22 000 individuals from 13 of Indonesia's 26 provinces which was representative of 83% of Indonesia's population. The IFLS3 re-contact rate was 95.3% of IFLS1 households. As we are not following individual children over time, but age groups as explained below, selective sample attrition is unlikely to be an issue. The data from 2000, gives us a picture of the state of Indonesian households

[9] Frankenberg, E. and L.Karoly. "The 1993 Indonesian Family Life Survey: Overview and Field Report." November, 1995. RAND. DRU-1195/1-NICHD/AID

Strauss, J., K. Beegle, B. Sikoki, A. Dwiyanto, Y. Herawati and F. Witoelar. "The Third Wave of the Indonesia Family Life Survey (IFLS3): Overview and Field Report." March 2004. WR-144/1-NIA/NICHD

at the tail end of the financial crisis, while the 1993 data provides a pre-crisis comparison. The IFLS questionnaire consists of questions covering socioeconomic-, health status- and community variables. The health variables were collected by a trained nurse and height measurements were taken lying down for the youngest children to allow the spine to be fully extended.

We have excluded observations with z-scores of an absolute size of 6 or more following Hoddinott & Kinsey (2001) to remove outliers and potentially faulty recordings. The z-score represent the deviation of an individual's height from the median of the reference population expressed in number of standard deviations. Table 1.1 summarizes the data used by IFLS wave.

The IFLS3 survey reached 10 435 households and interviewed 43 649 individuals. The households consisted of the original IFLS1 households plus split-offs between 1993 and 2000. Our increased sample size in 2000 shown in Table 1.1, reflects the general increase in respondents. The unit of analysis is individual children. IFLS3 children are on average a little less healthy which could be a result of being older on average. IFLS3 individuals are older on average since the IFLS3 data in our sample consists of a relatively higher proportion of older control observations. IFLS3 individuals are also slightly less urbanized on average. This is not a general trend between the IFLS1 and IFLS3 surveys, but rather a feature of our specific sample.

As we explain further below, we compare young children (treated) with adolescents (controls). We estimate a difference-in-differences estimator using two IFLS surveys (IFLS1 and IFLS3). Table 1.2 examines the data for potential sample selection bias in our estimations. If the characteristics of the households these children live in, change between the two periods and the two groups (treated and controls), we should interpret our results carefully. We have shown the means and standard deviations for two household characteristics; log of per capita expenditure (logpce) and maternal education. We believe that these are the two most relevant variables and we show later that z-scores are significantly associated with both. Households with an adolescent are better off on average, but this difference is fairly stable between the two periods. In IFLS1 there is no difference in the fractions of mothers with elementary education only between households in the two groups. In IFLS3 this fraction is slightly higher in households with a non-crisis/control child, but the difference is small and not statistically insignificant. There seems to be

little evidence of systematic sample selection in our data along these dimensions. If anything, the development in these variables counteract the effects we document later.

1.3.2 Estimation and identification

We use a difference in differences (DD) estimator to estimate the effect of the financial crisis in the late 90's on children's health. Table 1.3 illustrates our estimation strategy. Children who were in utero during the crisis will form the treatment group. These children are born after 1/1/1998 and have experienced a minimum of four months of in utero exposure to the crisis. Research shows that only very young children experience growth retardation in periods of hunger (Martorell, 1995; Hoddinott & Kinsey, 2001). We use children the age of 10-15 years at the onset of the crisis as our control group. Theses children are born between 1/9/1982 and 1/9/1987 and are assumed to be unaffected by the crisis for the purpose of the following analysis. Children born between the control and treatment periods (1/9/1987-1/1/1998), are excluded to create a clean break between affected and unaffected children. We wish to focus on children who were affected in utero and therefore exclude children who could have been affected in early childhood. To sum up; crisis children (the treatment group) are younger than the control group and might have experienced growth retardation during the crisis while the older control group could not.

Now that we identified the control and treatment group in the IFLS3 sample, we identify the same age groups at height measurement in the IFLS1 survey which will act as a pre-crisis survey comparison. Crisis group children had their heights measured between the ages of 0-35 months in IFLS3. Children measured within the same age range in IFLS1, serve as the pre-crisis treatment group. The pre-crisis control group is found equivalently. We identify the pre-crisis treatment- and control groups based on the age range at height measurement of the crisis groups to minimize the likelihood that our results reflect changes in age at measurement pre- and post crisis. Country specific age trends in z-scores (h_i) could influence our results unless we take measures to account for those. The treatment variable, T, takes the value 0 for controls (i.e. older children) and 1 for treated, crisis children. Our survey dummy, t, takes the value 0 for IFLS1 observations and 1 for IFLS3 observations.

We will be using height measurements as the indicator of health. Height increases with age and to decrease the age-dependency of our dependent variable, h_i is measured using z-score. Z-score is the number of standard deviations the child's measurement deviates from the relevant age- and gender standardized measurement as reported by the US National Center for Health Statistics. We use NCHS instead of the more commonly used WHO height for age tables as the former include children into adolescence while WHO only publish tables up to the age of 60 months. As we wish to include children who are as old as 219 months in our control group, we require expanded height-for-age tables. The two height-for-age tables are very similar for children below the age of 60 months. The correlation is 99.99% and levels deviate by 1.7% at most.

h_{isr} is modelled by the following equation which draws on the work of Case et al. (2002) and Gigante et al. (2006).

$$h_{isr} = \alpha + \beta T_s + \gamma t_r + \delta(T_s * t_r) + \eta X_i + \lambda H_i + \theta M_i + \epsilon_{isr} \quad (1.1)$$

The health variable, h_{isr}, is indexed by treatment status (s), survey round (r) and the individual index (i). δ captures the effect of the crisis on health of the treatment group. The estimate of δ is the difference in health changes between the two periods of the treatment group relative to the control group. The treatment group consists of younger children who were in utero during the crisis, while the control group consists of older, assumed unaffected children. If this estimator is negative, it provides evidence of a detrimental health effect of the crisis on children who were in utero during crisis time. β is a young children specific effect while γ is the survey specific effect. X_i is a matrix of individual controls including monthly age dummies. The age dummies are included to account for the possibility that Indonesian children deviate from the NCHS measures in a way that is related to age. As reported in Case et al. (2002), older children tend to be less healthy as a result of chronic conditions and although this effect should be captured in the NCHS reference population, we cannot exclude the possibility of country specific age-trends in health for Indonesian children. H_i is a matrix of household controls which includes household per capita expenditure (y_i). The health gradient is well documented[10] and provides the motivation for including y_i in H_i. Children's health

[10] see ex. Case et al. (2002), Currie et al. (2004)

is usually found to be positively associated with family income. As Indonesia is an economy dominated by agriculture and self-employment, data on current income is not reliable and we therefore use per capita real consumption expenditure (y_i) as a proxy for family income. M_i contains maternal characteristics. ϵ_{isr} is a random error.

We would like to control for parental health, but the IFLS3 survey does not include a meaningful number of respondents to questions on general health. As far as parental health is correlated with our independent variables, this could bias our estimates. It could potentially be correlated with y_i and mother's education which is included in M_i. We will be careful to compare our estimates on the control variables with other research and in general our estimates are very similar to what others have found.

The identification of our main coefficient, δ, relies on the assumption that the only development between 1993 and 2000 which could differentially affect younger children compared to older, is the financial crisis which started in 1997. Stated differently; the only reason why the difference between crisis and non-crisis children in Table 1.3 is not equal to zero, is the financial crisis. We have argued that our control group should not suffer growth retardation due to the crisis and Table 1.3 seems to indicate this while younger children could have been affected. It is very difficult to find other events between 1993 and 1997 that would negatively affect the height of young children. This is a period in Indonesia history characterized by tremendous economic growth and a general expansion of health care infrastructure (Lanjouw et al., 2001). δ should capture the fall in early life nutritional intake caused by the crisis. Between 1993 and 1997, poverty rates fell by 9% (Surono, 1999). It seems implausible that the relative fall in health of younger children born during the crisis could be caused by anything beside the crisis as the variables found to determine the health of young children has either been controlled for directly or are related to nutritional intake of the child or mother. Until the crisis in the fall of 1997, the average Indonesian child and mother would have experienced improvements in nutritional intake as a result of increased production and lower poverty rates.

1.4 Results

It is not straightforward to put a time frame on the crisis in Indonesia. The financial crisis commenced with the floating of the Rupiah in August 1997 and the quick depreciation which followed. Inflation increased over the following year, especially on food, and combined with a collapse in the banking sector, the financial crisis spilled into the real economy. In terms of the current analysis, we focus on a period when poverty rates soared in Indonesia. Suryahadi et al. (2002) estimated the evolution of poverty rates in Indonesia over the span of the crisis. Over the period October 1997 and March 1998, poverty rates doubled and continued to grow until reaching its peak in October 1998. At that point, poverty rates had increased by 150%. From October 1998 until the end point of Sutyahadi et al.'s analysis in October 1999, the poverty rate fell, but remained 50% higher in October 1999 than at the onset of the crisis. Ideally, we would like to have a measure of the intensity of the crisis and analyse the effect of crisis intensity on fetal development. However, as described in the introduction, fetal development is not only affected by the mother's nutritional intake during pregnancy, but also the body composition of the mother at conception. There is both an intensity and extensivity dimension to the effect of the crisis on fetal development. The longer a mother has been undernourished before conceiving, the more severe the effect on the fetus. This latter effect might even dominate the former (Barker, 1998). If we believe that the lack of proper nutrition during gestation is what drives our results, we could expect that crisis duration has an effect on the estimates.

Unless otherwise stated, the following analysis will treat children born on January 1, 1998 and later as children affected by the crisis. The older children included in our crisis cohort, will have experienced a minimum of 4 months of the financial crisis in utero. Martorell (1999) and Hoddinott & Kinsey (2001) find that very young children (up to the age of 3) are also in danger of suffering growth retardation as an effect of poverty while children above that age in developing countries seem to grow at the same pace as children in USA even though their nutritional environment is much worse (Martorell, 1999). The DD estimation will shed light on the difference in health changes between the 'treated', younger children who could have been affected by the crisis versus older children who could not. The 'untreated', older children are between the age of 10 and 15 years at the onset of the crisis and 13-18 at the time of height measurement. Table 1.4 and 1.5 show

the results of our DD estimation.

The DD estimator is estimating $\{E[h_1^Y] - E[h_0^Y]\} - \{E[h_1^O] - E[h_0^O]\}$ where Y indicates younger children exposed to the crisis in utero if taken from the 2000 survey and the same age group at height measurement if taken from the 1993 survey. O denotes older children who were not exposed during gestation. 1 indicates observations from the 2000 survey and 0 from the 1993 survey. The DD estimator is negative and significant indicating that children exposed in utero have had a negative development of average health between 1993 and 2000 compared to the older children. For the average child in our sample, the coefficient of -0.28 in Column (4) of Table 1.5 translates to a relative decline in height of 1.512 cm for children exposed prior to birth. This is a large effect considering the average height of the treatment group is 73.5 cm.

The survey dummy variable, t, is 1 for observations from the 2000 survey and 0 for 1993. This dummy is generally insignificant except for Column (1) and (2) of Table 1.4. In the regressions run in these two columns, we have included very few control variables. The survey dummy likely picks up improvements between the time of the two surveys in these control variables and is therefore positive and significant. The young children specific effect, T, is a dummy of value 1 if the observed child is in the affected group as detailed above and 0 otherwise. An observation of 1 indicates a young child and should be expected to be positive following Case et al. (2002). We find this effect in all but Column (2) in Table 1.5 where the number of observations is very low. As we include monthly age dummies in our regression, this coefficient loses significance as we would expect. The age dummies are all negative compared to age=0 which is our omitted variable and falling until the age of 22 months where it seems to stabilize. Since T is a linear combination of age dummies, the age dummies and T are clearly collinear. The main issues with multicollinearity are inflated standard errors and unstable coefficients. We do not seem to observe any of those problems with the estimations we run.

Urban children are generally healthier than their rural counterparts which is not surprising as urban centres generally have better health infrastructure and people are more educated. In my sample, 97.8% of urban mothers have ever attended school while the number for rural mothers is 90.4%. The fraction of urban mothers with more than elementary school is 72.7% which is almost 20% higher

then their rural counterparts. Mother's education has been shown to be an important determinant of children's health (Case et al., 2002). As we included mother's education in Column (3) of Table 1.5, the coefficient falls somewhat in size and significance. The gender dummy is insignificant at conventional levels and changes sign depending on our specification. Log household size is negative and significant in our final specification, but generally insignificant. The literature on the relationship between household size and children's health, seem to be inconclusive (Bawah, 2001), but these results indicate a negative, but largely insignificant effect conditional on the other controls, of growing up in larger households. This could be the effect of growing up in cramped and less sanitary quarters or larger households might prioritize the more productive members of the household and pregnant women could suffer as a result.

Per capita expenditure is highly significant and positive in all but one specification as should be expected. The health gradient is the well-known[11] albeit not fully understood, positive relationship between wealth and health. For adults the relationship is not necessarily causal as the direction of causality is almost certainly two-way. However, for children under working age, the effect on health of growing up in wealthier households, is positive and more likely to be causal. Education of the mother has been positively associated with children's health (Case et al., 2002) and this result is mirrored in Column 3 and 4 of Table 1.5. The excluded group is mothers with elementary school only. Compared to mothers with elementary school only, mothers with a higher level of education have healthier children. This effect seems to not be linear in the levels of education as women with senior high school do not have significantly healthier children than women with junior high school.

Gigante et al. (2006) found a negative relationship between maternal smoking during pregnancy and children's health. Unfortunately, the IFLS surveys provide very few responses to the questions on smoking. For respondents who had smoked previously but had since stopped, the information on when they stopped was not detailed enough to determine whether the mother was smoking during the relevant pregnancy. As a result, these observations had to be dropped and we were left with only 241 observations. Due to the increase in standard errors, almost all coefficients become insignificant. The coefficient on the smoking dummy is positive,

[11] ex. Case et al. (2002), Banks et al. (2006)

but insignificant and due to the low number of observations, little can be learned from Column (2) of Table 1.5. In Column (4), we find a positive and statistically significant relationship between the age of the mother at pregnancy and the child's health. Gigante et. al (2006) finds a similar effect. This could be the result of accumulated experience and knowledge not reflected in the educational variables.

Our main coefficient of interest is the DD estimator, δ, which is consistently negative and significant over all but the specification in Column (2) of Table 1.5. This coefficient falls by 70-120% when we include the age dummies. At the same time the coefficient on T, β falls in absolute value, but increases in significance while t remains relatively unchanged. As the coefficient on $T * t$, δ is a function of the size of β, the DD estimate decreases as a result of the reduction in β.

1.4.1 Urban-rural differences

The results above clearly show that between 1993 and 2000, the health of young children took a turn for the worse compared to older adolescents. So far we have attributed this relative deterioration of health to the financial crisis which started in 1997. It is possible that this development is caused by something completely unrelated to the crisis even though it is difficult to find other plausible mechanisms.

We do know that the crisis did not affect everyone equally. As described earlier, wage earners and people without land were especially vulnerable as prices soared on food stuffs and unemployment increased. 'Urban areas on the island of Java-which were among the wealthiest areas of the crisis - were the epicentre of the financial and modern sector crisis. In contrast, traditionally poorer natural resource exporting areas actually benefited from the crisis.' (Pritchett et al., 2002). 'Urban areas have been hit harder by the crisis than rural areas.' (Sumarto et al., 1999). If the impact of children's health follows the same urban-rural pattern, it seems we could attribute the health effects more confidently to the financial crisis. In Table 1.6, we show the result of running the following model:

$$h_{isr} = \alpha + \beta_1 T_s * (urban) + \beta_2 T_s * (1 - urban) + \gamma_1 t_r * (urban) + \gamma_2 t_r * (1 - urban) + \delta_1 (T_s * t_r) * (urban) + \delta_2 (T_s * t_r) * (1 - urban) + \eta X_i + \lambda H_i + \theta M_i + \epsilon_{isr} \quad (1.2)$$

δ_1 and δ_2 are the DD estimates for urban and rural children respectively. By running one regression for both groups, we can statistically test whether the coef-

ficients are different and whether urban children are affected more severely as we expect.

Table 1.6 clearly shows that there is a large difference between urban and rural children. While urban children have been severely affected, rural children have a positive albeit not significant DD estimate[12]. The estimate for urban children is almost exactly twice the absolute size of our original regression. The F-test confirms that the difference between δ_1 and δ_2 is statistically significant at the 5% level for all our specifications. The control variables show similar effects to the ones found in Table 1.5. Above all, the results in Table 1.6, reinforce the idea that the relative loss of health of young children was caused by the financial crisis and not other events occurring during the period of analysis unless these events followed the same urban-rural divide.

An interesting result from Table 1.6, is the difference between $urban * t$ and $(1 - urban) * t$. The coefficients on these two variables are the estimates of the development of health in the control group for urban and rural children respectively. As in Table 1.5, we see the size of the coefficients fall as we include more control variables. The coefficient for urban children is significantly larger than for rural children who experienced a significantly negative development between the two survey rounds. It is hard to find an obvious reason for this result, but it seems to suggest a negative development in health of rural children of all ages between the surveys.

As described, the crisis would affect the purchasing power of wage earners both through prices and wages. Inflation increased greatly and wages fell or were eliminated completely through lay-offs. There are striking differences between the fate of a wage earners and someone living off their land during the crisis. The last group could benefit from increased prices and their crops could serve as insurance against the detrimental effects of the crisis experienced by wage earners. Deaton (1989) examined the effects of increases in the price of rice in Thailand on the income distribution. He finds that rural households are the beneficiaries when the price of rice increases and that rural households in the middle of the income distribution gain the most. Our aim is not to analyse such questions in depth, but the distributional effects we find here are similar to Deaton's. Urban children seem

[12] There is no a priori reason why the controls should differentially affect urban and rural children. We ran regressions interacting all controls with the urban dummy and the controls are not significantly different in urban and rural areas.

to have faired much worse than their rural counterparts and this result helps us attribute our results to the financial crisis with more confidence.

We also ran (1.2) interacting t, T and $t*T$ with indicators for whether the household had above or below mean per capita expenditure. We would expect children born in poorer household to be more severely affected than children from relatively rich households. Affluent household would to a greater degree be insulated from the crisis. Both DD estimates, for below and above mean per capita expenditure households, were negative and significant and the size of the below mean households estimate was about twice the absolute value of the above mean households. The difference is only significant at 15%, but it goes some way to confirm the general observation that poorer households suffered more from the crisis.

1.4.2 In utero effects

So far we have concentrated on children who were exposed to the crisis in utero. Most of our sample also had considerable post natal exposure, but we have not included any children without in utero exposure in our 'crisis group'. Martorell (1995) found that nutritional interventions in Latin-America improved growth rates of children up to the age of three years. Hoddinott & Kinsey (2001) found that children in Zimbabwe who were exposed to severe drought at the age 12-24 months, had a significant slowdown in growth. It is important for the targeting of potential interventions to know which groups are most vulnerable. Again we will use children between the age of 10 and 15 years at the onset of the crisis as the comparison group. This group is assumed to be unaffected by the financial crisis. Instead of comparing two groups, the unaffected against the affected, we want to compare the unaffected with several age groups that might or might not have suffered ill health as a result of being born around the time of crisis. To do that, we estimate the following equation:

$$h_{isr} = \alpha + \gamma t_r + \sum_{k=1}^{5}(\beta_k T_k + \delta_k(T_k * t_r)) + \eta X_i + \lambda H_i + \theta M_i + \epsilon_{isr} \quad (1.3)$$

k denotes the different groups of children. We have six groups; first the group of adolescents between 10 and 15 years at the onset of the crisis. The following five treatment groups consist of children born 1-2 years before the onset, 0-1 years before, 0-1 years after, 1-2 years after and 2-3 years after the onset at 1st of

September 1997. δ_k is the estimate of the DD estimate of group k. Figure 1.2 shows the points estimates of δ_k. The horizontal axis shows the year of birth as detailed above. The point estimate at 1995 shows the coefficient of children born between 1/9/1995 and 1/9/1996 and the dotted line indicates a 5% confidence interval around the point estimates. Only the estimates between 1997 and 1999, are based on children who experienced the crisis in utero.

Figure 1.2 shows that children born 0-24 months before the crisis started, seem to have had no negative health effect. All children with in utero exposure, have a negative coefficient and this coefficient seems to increase in absolute value as we move towards the end of the crisis period. As discussed earlier, there is medical evidence showing that not only is nutritional intake during pregnancy important, but the body composition of the mother at the time of conception might be even more important for the growth of the fetus and subsequent early childhood growth (Barker, 1998). The results shown here, suggest that this mechanism might have played a role for children and mothers during the Indonesian crisis. Even though the poverty levels were higher in 1998, women who were pregnant during 1999 and 2000, did not give birth to babies with significantly better health. These results highlight the need for special support to pregnant women during times of economic crisis. The failure to protect mothers and their children, could have long-lasting effects on the economy as we will discuss further below.

Figure 1.3 shows the same picture as Figure 1.2, but with all non-control variables interacted with the urban variable. In the same fashion as in Table 1.6, we isolate the DD estimates for the group of children born in urban and rural areas. For the group of urban children, all the estimates in Figure 1.3 are statistically significant except the estimate for 1996. Only 1995 and 1999 estimates are statistically significant for the rural group. We clearly see that urban children were more exposed to the crisis than rural children until 1999-2000. We find a negative coefficient for urban children born in 1995, but it seems improbable that this is a result of the crisis as the coefficient for 1996 is close to zero. It is possible that children with no in utero exposure, but early life exposure to the crisis could be negatively affected, but we would expect to find a similar effect for children born in 1996 if this was the case. From 1997, we see a significant negative effect on urban children. The urban estimates for 1997 and 1998 are significantly different from the equivalent rural estimates. Something interesting happens in 1999. This is the

only year rural children seem to suffer from the crisis and while urban children are also negatively affected, they are less so than the previous two years. In 1999, the volume of rice imported by Indonesia fell by almost 50% compared to 1998 (Surono, 1999). At the same time, production of rice, the staple food of Indonesians[13], fell a little. In 1999 macro economic indicators like inflation and GDP improved. It is difficult to say confidently how these effects interacted to change the distributional effects of the over the crisis period. Figure 1.3 clearly shows a large difference in the effect of the crisis on rural and urban children until 1999. Figure 1.2 and 1.3 are based on Column (2) of Table 1.7 and 1.8, respectively.

1.4.3 Potential future consequences

The next wave of the IFLS surveys was released early spring in 2009. With this data, we are able to follow children born during the crisis into the early stages of the education system. However, even with that data, we are not able to directly analyse the full effect of the crisis on educational- and labour market outcomes as the crisis children are still too young. Inference to the potential future consequences of the crisis might be drawn using research from other settings. The medical journal Lancet recently published an overview of research on long-term consequences of child stunting (Victoria et al., 2008). The paper summarizes the results of five cohort studies from Brazil, Guatemala, India, the Philippines and South Africa and provides new estimates pooling the data from these separate studies. Few papers from a developing country setting follow children from birth to adulthood. Alderman et al. (2004) investigates the relationship between preschool stunting and human capital formation in Zimbabwe. The aim is to link the loss in stature at early childhood with adult or adolescent outcomes, particularly education. This demands a lot of the data. We need early childhood height measurements and to follow children over a long period of time. There are not many papers using data that fulfil these criteria, but the two mentioned above do and will serve as the basis for what follows. The evidence used here is not going to be perfectly compatible to an Indonesian setting, but will still give us some idea of how the financial crisis could have long-term effects working through human capital.

[13] Rice contributes more than 60% of calories consumed by Indonesians (Surono,1999)

The five studies in Victoria et al. (2008) all find a significant relationship between z-score at the age of two and subsequent human capital formation. The estimates range from 0.21 years of schooling per z-score point, to 0.59 years. Their pooled estimation of all studies combined, results in an estimate of 0.51. Alderman et al. (2004) employ a somewhat more stringent estimation strategy. They acknowledge that ordinary OLS estimates may be biased since household and child specific components may be present both in the health variable and in the error of the ultimate regression of the health variable on adult outcomes. To account for this, they run a combined maternal fixed effects and instrumental variables regression. They effectively compare children with the same mother where one of the siblings has been born during a period of drought or civil war, while the other was born far enough removed from this event to have been affected. They find a rather large difference in estimates between OLS and their preferred estimation strategy. Running number of grades attained on preschoolers' z-score, they find a coefficient of 0.678. The OLS estimate is only 0.222.

Our results show a reduction in z-score of the affected cohort of 0.28 z-score points. Depending on the estimate we choose to use, this cohort will on average stand to lose between 0.06 and 0.19 years of schooling. The lower end of this estimate is likely to suffer from the problems highlighted in Alderman et al. (2004) and our preferred estimate is the latter. It is further interesting to speculate about how the loss in human capital will affect future wages for children in utero during the Indonesian crisis. Duflo (2001) exploits the school building program which was initiated in the mid 70's in Indonesia (the INPRES program) to estimate returns to education. She employs a 2SLS approach where the difference in program intensity is exploited and finds a return to education of about 8%. Combined with our earlier estimates, this implies a decrease in wages of about 2% on average for the Indonesian crisis children. These calculations should be interpreted with caution as we use non-Indonesian estimates throughout. In the 1993 IFLS wave, only 45% of individuals worked for a wage and Duflo (2001) finds a somewhat lower return to education when using imputed wages for self-employed people. Whatever the true wage effects of this crisis will eventually be, it seems likely that these crisis children will face non-trivial negative outcomes in the future as a result of the timing of their birth. When we isolate urban children who suffered the most, we found a coefficient of -0.493. That in turn translates into 0.33 years of lost schooling and

a future loss in wages of 2.7%.

1.5 Conclusion

Indonesia suffered more that most other countries hit by the South-East Asian crisis in the late 1990's. As the Rupiah was allowed to float, the credit market collapsed and production fell. Poverty rates increased by 150% over the first year of the crisis. Thomas et al. (1999) and Cameron (2000) examined the effect of this increase in poverty on the health of Indonesian children and found no negative effect. As we have argued throughout this article, we need to pay special attention to children who experienced this episode of poverty during gestation.

The results presented here clearly show a negative effect of crisis exposure in utero. We ran a DD estimation which shows a loss of 0.28 z-score points for the youngest children in our sample who were exposed in utero, compared to older children who were not. We find no such effect for children who were born 0-24 months prior to the onset of the crisis. It is difficult to have a clear prediction a priori as to which part of the crisis period would have the largest in utero effect. As reported by Suryahadi et al. (2002), poverty rates were higher in 1998 than 1999, but our results show that children born the 12 months period beginning September 1st 1999 were no less severely affected than children born the prior year. We know from medical literature that the body composition of the mother at conception is possibly more important than nutritional intake during gestation (Barker, 1998) and our results show that young Indonesian children suffered more towards the end of the crisis. One possible explanation of our results could be that as we move further into the crisis, the body composition of mothers increasingly deteriorate. It is also possible that as the crisis progresses, mothers increasingly understand it's severity and therefore stress levels increase with time, We also found that urban children were much more affected than rural children who were seemingly unaffected until the tail end of the crisis. This result makes the attribution of the negative health effects to the financial crisis more credible as the crisis hit urban areas the hardest.

During times of increased poverty, NGOs and governments often initiate nutritional interventions. The results presented here, show that we need to pay special attention to pregnant women during such times. Their children are likely to suffer

especially harsh consequences of increased poverty. We have speculated on some of the potential future consequences of early life stunting which include lower educational attainment and lower wages. Stunted female babies are also likely to give birth to children with lower stocks of health (Almond & Chay, 2003). Episodes of poverty during gestation could very well influence the people affected and the economy as a whole not only for the current generation, but also for future generations.

Figure 1.1: Real GDP growth of Indonesia 1980-2000

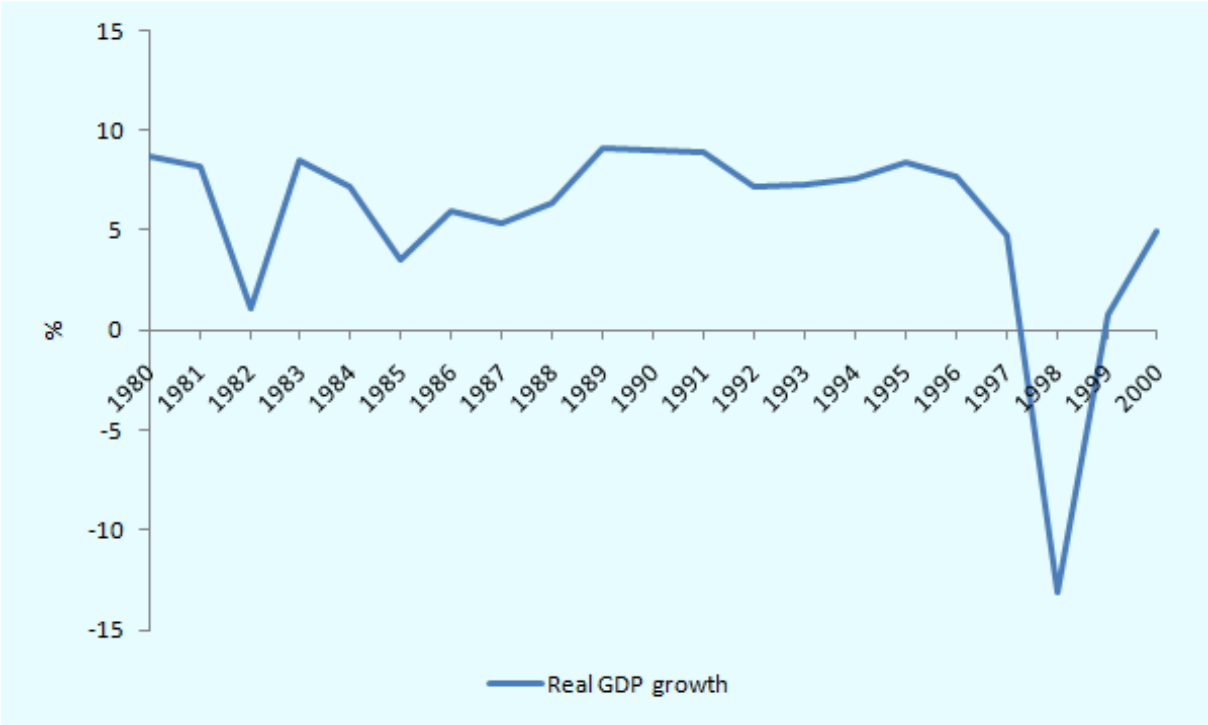
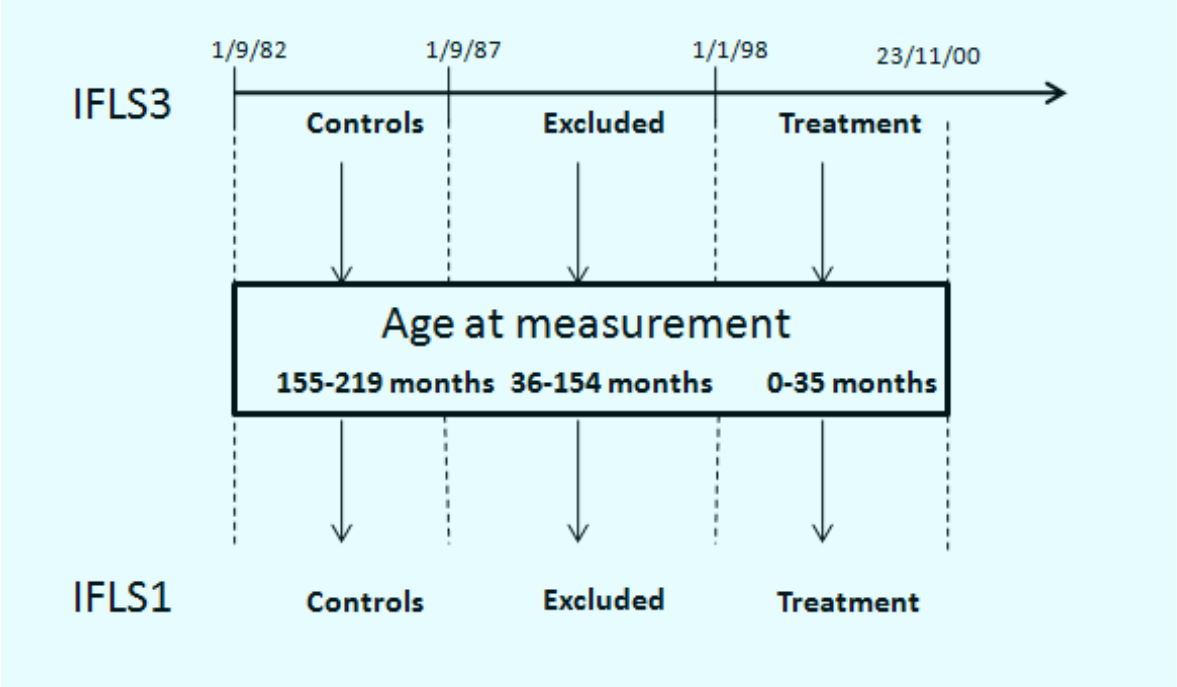
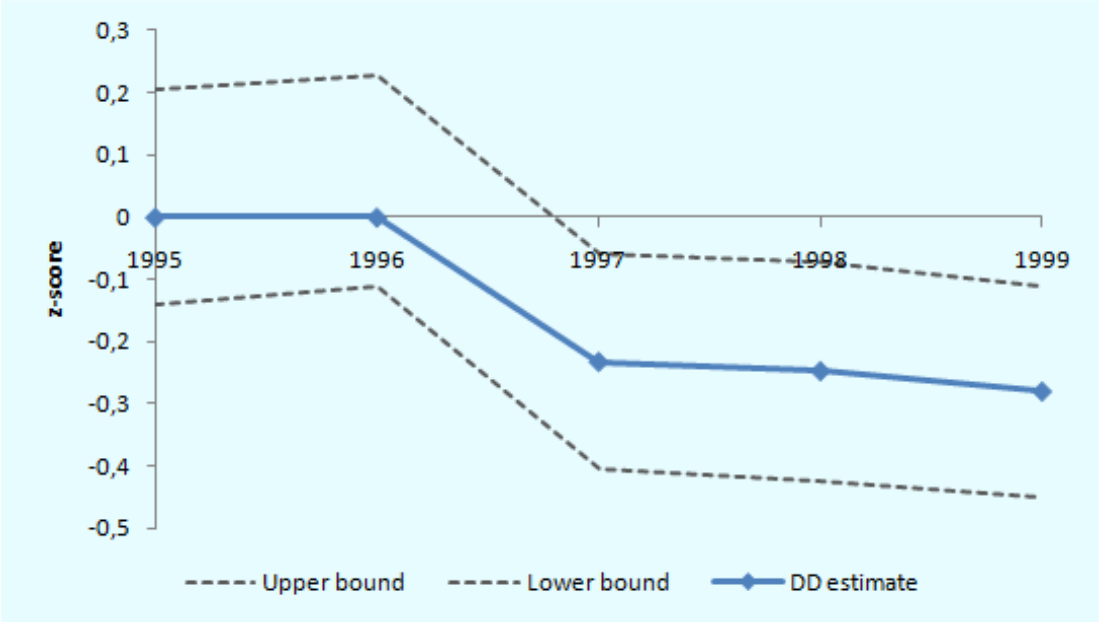


Figure 1.2: DD estimate separated by year of birth



Notes: The point estimate labeled 1995 refers to children born between 1/9/1995 and 1/9/1996 . The other point estimates are labeled similarily.

Figure 1.3: DD estimate separated by year of birth further stratified by household urban/rural status



Notes: The point estimate labeled 1995 refers to children born between 1/9/1995 and 1/9/1996 . The other point estimates are labeled similarily.

Table 1.1: Descriptive statistics of the sample

Survey	1993/IFLS1	2000/IFLS3
Number of observations	2833	6299
Individual variables		
% male	50.40 %	49.96 %
Mean age (months)	83.7	127.1
Mean z-score	-1.43	-1.55
Household variables		
% urban	52.45 %	49.42 %
Mean realpce	181727.2	241533.4
Mean household size	5.35	6.34

Notes: *z-score* is the difference between actual height and the height of a same age and gender reference population, measured in standard deviations. *realpce* is the monthly real per capita household expenditures.

Table 1.2: Sample selection bias

	logpce		**Maternal education**	
	1993	**2000**	**1993**	**2000**
Treated (young children)	10.87	12.10	0.54	0.46
	(1.14)	(0.68)	(0.50)	(0.48)
Controls (adolescents)	11.21	12.26	0.55	0.51
	(1.14)	(0.68)	(0.50)	(0.50)

Notes: *logpce* is the log of monthly real per capita household expenditures. *Maternal education* is measured as the fraction of mothers with elementary school only. *1993* refers to data from the IFLS1 survey, *2000* from the IFLS3 survey. *Treated* children are 0-35 months at height measurement while *Controls* are 13-18 years. Standard deviations in paretheses.

Table 1.3: Difference in differences: z-score

	Crisis	Non-crisis	Difference
Treated (young children)	-1.068	-1.060	-0.008 (0.056)
Controls (adolescents)	-1.810	-1.946	0.136 (0.033)
Difference	0.742 (0.032)	0.886 (0.056)	-0.144 (0.060)

Notes: *z-score* is the difference between actual height and the height of a same age and gender reference population, measured in standard deviations. *Crisis* refers to observations from IFLS3 (2000), *Non-crisis* to observations from IFLS1 (1993). *Treated* children are 0-35 months at height measurement while *Controls* are 13-18 years. Standard errors in parentheses.

Table 1.4: The effect of in utero exposure to the crisis on early childhood z-score

	(1)	**(2)**
LHS variable	z-score	z-score
t (survey dummy)	0.137**	0.122**
	(3.21)	(2.79)
T (treatment dummy)	0.887**	2.188**
	(18.06)	(4.53)
t*T (DD estimator)	-0.144**	-0.247**
	(-2.41)	(-4.12)
Age dummies	No	Yes
Constant	-1.947**	-2.303**
	(-52.05)	(-5.26)
Observations	9132	9132
R-squared	0.082	0.163

Notes: *z-score* is the difference between actual height and the height of a same age and gender reference population, measured in standard deviations. t=1 if observation is from the 2000 survey, 0 if from the 1993 survey. T=1 if the observed child is born after 1/1/98 or from the equivalent age group at height measurement if t=0. **p<0.05, *p<0.10. t-stats in parentheses.

Table 1.5: The effect of in utero exposure to the crisis on early childhood z-score

	(1)	**(2)**	**(3)**	**(4)**
LHS variable	z-score	z-score	z-score	z-score
t (survey dummy)	-0.046	-0.340	-0.050	-0.026
	(-1.00)	(-1.37)	(-0.97)	(-0.50)
T (treatment dummy)	2.113**	0.678	2.107**	1.647**
	(5.41)	(0.53)	(4.13)	(2.82)
t*T (DD estimator)	-0.312**	-0.256	-0.298**	-0.282**
	(-5.32)	(-0.47)	(-4.73)	(-4.26)
Age dummies	Yes	Yes	Yes	Yes
Male	0.033	-0.276*	0.021	0.027
	(1.28)	(-1.89)	(0.78)	(0.93)
Urban	0.347**	0.171	0.343**	0.345**
	(13.25)	(1.11)	(12.03)	(11.40)
Logpce	0.196**	0.434**	0.195**	0.190**
	(12.86)	(4.80)	(11.67)	(10.84)
Loghhsize	-0.018	-0.061	-0.034	-0.072**
	(-0.66)	(-0.31)	(-1.01)	(-1.97)
Smoking		0.188		
		(0.98)		
Mother's education				
Junior high school			0.118**	0.146**
			(2.07)	(2.56)
Senior high school			0.090	0.118*
			(1.42)	(1.84)
College/University			0.574**	0.595**
			(4.49)	(4.71)
Mother's age at pregnancy				0.007**
				(2.84)
Constant	-4.449**	-5.892**	-4.432**	-4.103**
	(-11.63)	(-4.08)	(-8.71)	(-7.12)
Observations	9106	241	8138	7025
R-squared	0.202	0.573	0.209	0.208

Notes: *z-score* is the difference between actual height and the height of a same age and gender reference population, measured in standard deviations. t=1 if observation is from the 2000 survey, 0 if from the 1993 survey. T=1 if the observed child is born after 1/1/98 or from the equivalent age group at height measurement if t=0. *Urban* is a dummy of value 1 if the household resides in an urban area. *Logpce* is the log of monthly real per capita household expenditures. *Loghhsize* is the log transformed number of household members. All education variables are dummies with elementary school as the excluded variable. **p<0.05, *p<0.10. t-stats in parentheses.

Table 1.6: Analysis of urban/rural differences

	(1)	**(2)**	**(3)**	**(4)**
LHS variable	z-score	z-score	z-score	z-score
Urban*t	0.320**	0.313**	0.075	0.100*
	(6.89)	(6.57)	(1.47)	(1.734)
(1-Urban)*t	-0.055	-0.055	-0.231**	-0.214**
	(-1.18)	(1.18)	(-4.75)	(-3.84)
Urban*T	1.136**	2.212**	2.407**	1.965**
	-19.55	(5.57)	(5.08)	(3.35)
(1-Urban)*T	0.642**	1.688**	1.983**	1.569**
	(-11.12)	(4.26)	(4.19)	(2.68)
Urban*t*T	-0.390**	-0.516**	-0.511**	-0.493**
	(-5.14)	(-6.86)	(-6.86)	(-5.85)
(1-Urban)*t*T	0.132*	0.026	-0.076	-0.045
	(1.78)	(0.35)	(-1.05)	(-0.54)
Age dummies	No	Yes	Yes	Yes
Male			0.033	0.027
			(1.28)	(0.94)
Logpce			0.204**	0.199**
			(13.47)	(11.40)
Loghhsize			-0.011	-0.061*
			(-0.39)	(-1.67)
Mother's education				
Junior high school				0.161**
				(2.82)
Senior high school				0.135**
				(2.10)
College/University				0.608**
				(4.79)
Mother's age at pregnancy				0.007**
				(2.91)
Constant	-1.947**	-1.974**	-4.465**	-4.178**
	(-52.58)	(-5.80)	(-9.64)	(-7.24)
Observations	9132	9132	9106	7025
R-squared	0.101	0.182	0.199	0.205

Notes: t=1 if observation is from the 2000 survey, 0 if from the 1993 survey. T=1 if the observed child is born after 1/1/98 or from the equivalent age group at height measurement if t=0. *Urban* is a dummy of value 1 if the household resides in an urban area. *Logpce* is the log of monthly real per capita household expenditures. *Loghhsize* is the log transformed number of household members. All education variables are dummies with elementary school as the excluded variable. **p<0.05, *p<0.10.

Table 1.7: Time effects

	(1)	**(2)**	**(3)**	**(4)**
LHS variable	z-score	z-score	z-score	z-score
t (survey dummy)	0.137**	0.136**	-0.056	-0.005
	(-3.28)	(-3.27)	(-1.28)	(-0.10)
Birth year dummies				
1995	0.314**	0.315**	0.393**	0.328**
	(4.76)	(4.79)	(6.12)	(4.79)
1996	0.445**	0.448**	0.502**	0.423**
	(7.01)	(7.07)	(8.09)	(6.45)
1997	0.461**	0.463**	0.539**	0.435**
	(7.24)	(7.27)	(8.65)	(6.61)
1998	0.518**	0.514**	0.604**	0.550**
	(8.04)	(7.98)	(9.59)	(8.12)
1999	1.573**	1.580**	1.698**	1.570**
	(24.62)	(24.73)	(27.13)	(23.13)
t*1995	0.091	0.092	0.052	0.032
	(1.11)	(1.12)	(0.66)	(0.37)
t*1996	0.053	0.05	0.041	0.057
	(0.66)	(0.62)	(0.52)	(0.66)
t*1997	-0.198**	-0.200**	-0.209**	-0.233**
	(-2.44)	(-2.45)	(-2.63)	(-2.64)
t*1998	-0.179**	-0.176**	-0.222**	-0.248**
	(-2.18)	(-2.15)	(-2.78)	(-2.76)
t*1999	-0.306**	-0.311**	-0.382**	-0.281**
	(-3.90)	(-3.96)	(-4.97)	(-3.27)
Birth month dummies	No	Yes	Yes	Yes
Male			0.000	-0.018
			(0.02)	(-0.72)
Urban			0.371**	0.365**
			(16.17)	(13.91)
Logpce			0.203**	0.182**
			(15.36)	(12.12)
Loghhsize			-0.017	-0.074**
			(-0.69)	(-2.29)

Table 1.7 cont: Time effects

	(1)	**(2)**	**(3)**	**(4)**
LHS variable	z-score	z-score	z-score	z-score
Mother's education				
Junior high				0.164***
				(3.42)
Senior high				0.192**
				(3.63)
College/University				0.624**
				(6.07)
Mother's age at pregnancy				0.005**
				(2.43)
Constant	-1.947**	-2.017**	-4.495**	-4.296**
	(-53.11)	(-36.86)	(-26.78)	(-22.27)
Observations	12150	12150	12118	9404
R-squared	0.105	0.107	0.151	0.158

Notes: *z-score* is the difference between actual height and the height of a same age and gender reference population, measured in standard deviations. t=1 if observation is from the 2000 survey, 0 if from the 1993 survey. *Urban* is a dummy of value 1 if the household resides in an urban area. *Logpce* is the log of monthly real per capita household expenditures. *Loghhsize* is the log transformed number of household members. All education variables are dummies with elementary school as the excluded variable. **p<0.05, *p<0.10. t-stats in parentheses.

Table 1.8: Time effects and urban/rural differences interacted

	(1)	**(2)**	**(3)**	**(4)**
LHS variable	z-score	z-score	z-score	z-score
t*urban	0.320***	0.321***	0.071	0.129**
	(7.06)	(7.10)	(1.48)	(2.40)
t*(1-urban)	-0.055	-0.057	-0.242**	-0.197**
	(-1.21)	(-1.25)	(-5.17)	(-3.74)
Birth year dummies interacted				
1995*urban	0.665**	0.672**	0.699**	0.621**
	(8.02)	(8.10)	(8.51)	(7.09)
1995*(1-urban)	-0.069	-0.072	0.033	-0.018
	(-0.81)	(-0.84)	(0.39)	(-0.20)
1996*urban	0.616**	0.619**	0.618**	0.534**
	(7.76)	(7.80)	(7.86)	(6.46)
1996*(1-urban)	0.258**	0.264**	0.344**	0.267**
	(3.15)	(3.22)	(4.22)	(3.13)
1997*urban	0.903**	0.905**	0.916**	0.814**
	(11.24)	(11.27)	(11.51)	(9.84)
1997*(1-urban)	-0.000	0.003	0.110	-0.014
	(-0.00)	(0.03)	(1.35)	(-0.17)
1998*urban	0.829**	0.828**	0.838**	0.762**
	(9.96)	(9.96)	(10.18)	(8.58)
1998*(1-urban)	0.226**	0.218**	0.331**	0.298**
	(2.78)	(2.69)	(4.10)	(3.45)
1999*urban	1.600**	1.608**	1.653**	1.467**
	(19.53)	(19.63)	(20.37)	(16.75)
1999*(1-urban)	1.547**	1.554**	1.695**	1.625**
	(19.15)	(19.24)	(21.07)	(18.59)
t*1995*urban	-0.220**	-0.227**	-0.230**	-0.269**
	(-2.03)	(-2.09)	(-2.13)	(-2.28)
t*1995*(1-urban)	0.471**	0.478**	0.390**	0.381**
	(4.34)	(4.40)	(3.62)	(3.24)
t*1996*urban	-0.040	-0.039	-0.012	-0.011
	(-0.38)	(-0.37)	(-0.12)	(-0.09)
t*1996*(1-urban)	0.217**	0.209**	0.141	0.172
	(2.06)	(1.98)	(1.35)	(1.48)
t*1997*urban	-0.480**	-0.483**	-0.454**	-0.483**
	(-4.38)	(-4.41)	(-4.18)	(-4.00)
t*1997*(1-urban)	0.184*	0.184*	0.109	0.112
	(1.74)	(1.74)	(1.03)	(0.96)

Table 1.8 cont.: Time effects and urban/rural differences interacted

	(1)	**(2)**	**(3)**	**(4)**
LHS variable	z-score	z-score	z-score	z-score
t*1998*urban	-0.529**	-0.530**	-0.508**	-0.566**
	(-4.78)	(-4.80)	(-4.64)	(-4.59)
t*1998*(1-urban)	0.181*	0.188*	0.092	0.091
	(1.70)	(1.77)	(0.88)	(0.77)
t*1999*urban	-0.373**	-0.381**	-0.394**	-0.242**
	(-3.55)	(-3.63)	(-3.79)	(-2.07)
t*1999*(1-urban)	-0.213**	-0.216**	-0.333**	-0.293**
	(-2.09)	(-2.11)	(-3.29)	(-2.58)
Month dummies	No	Yes	Yes	Yes
Male			-0.001	-0.020
			(-0.02)	(-0.77)
Logpce			0.210**	0.189**
			(15.91)	(12.63)
Loghhsize			-0.012	-0.066**
			(-0.48)	(-2.05)
Mother's education				
Junior high				0.179**
				(3.74)
Senior high				0.205**
				(3.87)
College/University				0.627**
				(6.09)
Mother's age at pregnancy				0.005**
				(2.49)
Constant	-1.947**	-2.031**	-4.368**	-4.186**
	(-53.90)	(-37.66)	(-25.80)	(-21.55)
Observations	12150	12150	12118	9404
R-squared	0.132	0.134	0.152	0.160

Notes: *z-score* is the difference between actual height and the height of a same age and gender reference population, measured in standard deviations. t=1 if observation is from the 2000 survey, 0 if from the 1993 survey. *Urban* is a dummy of value 1 if the household resides in an urban area. *Logpce* is the log of monthly real per capita household expenditures. *Loghhsize* is the log transformed number of household members. All education variables are dummies with elementary school as the excluded variable. **p<0.05, *p<0.10. t-stats in parentheses.

2 Early childhood height and medium-term cognitive development

2.1 Introduction

The association between height and adult outcomes like socioeconomic status and wages is well established. Taller people earn more on average. The causal channels that link height and wages or status are still controversial. The proposed explanations range from systematic discrimination in the workplace to differences in adolescent social experiences as a result of height and the resulting human capital formation (Persico et al., 2004). In a recent paper, Case & Paxson (2008) document the association between early childhood height and cognitive abilities. Their main conclusion is that early childhood height determines both cognitive abilities and adult height and therefore the association of adult height and wages.

We will attempt to analyse the causal effect of early childhood height on school age cognitive abilities. The estimation strategy in this paper has been heavily influenced by a few papers investigating the causal effect of early childhood nutrition on primary school enrolment age. Glewwe & Jacoby (1995) has been influential in their emphasis on employing instrumental variables and maternal- or household fixed effects. While much of the work in the nutrition-cognitive development area seems to ignore the necessity of employing instrumental variables, this approach has become the norm for work following Glewwe & Jacoby (1995), like Glewwe & King (2001) and Alderman et al. (2006). It is clear that finding good instruments is not easy in many settings. We believe that we have an instrument which satisfy the exclusion restriction and has a clear effect on nutritional status in early childhood. Chapter 1 showed that Indonesian children who experienced the South-East Asian crisis in utero, had relatively worse health as measured by z-scores, compared to older children who did not. Our instrumental variables (IV) routine will exploit the crisis, by instrumenting z-score with a dummy which indicates whether the child was born during the defined crisis period (1998-2000). Without such instruments, we should not expect to find the causal effect, but rather view correlations as statistical associations. The main idea of the paper is that the South-East Asian crisis caused increased chance of stunting for babies in utero and that reduced early

childhood health has hampered their cognitive development into school age. Our sample is somewhat small for running household fixed effects. We include these results as a sensitivity exercise, and find that the results are similar to our IV results. The results presented here are qualitatively similar to the results in Alderman et al. (2006). They analyse other outcome variables, but also find that the use of IV changes the results substantially and that there is little difference between OLS with extended family and individual controls and maternal fixed effects.

The analysis presented here demands a lot from the data. We need to follow individual children over time and we require access to a valid instrument or instruments. Our data satisfy these conditions. We do however face another issue with our data. Our early childhood height measurement has not been performed at a specific age for all children. Some were measured shortly after birth, while others were measured close to the age of five. We use standardized z-scores as a proxy for health status and computing z-scores requires a healthy reference population. We found that Indonesian children do not exhibit exactly the same growth pattern as the reference population. Specifically, they are much closer to the reference in the first 0-20 months. We discuss below how this might affect our results and present our best solution to this issue. With the frequent use of z-score in a variety of contexts, this shows that country specific age patterns might be important to determine in order for our results to be interpreted correctly. To account for the relationship between age and z-score in our data, we restrict our non-crisis part of the sample to children that had height measurements done within the same age range as the 'crisis children'. Crisis children are those born in the period 1998-2000.

Lynn (1990) summarises the state of similar research at the end of the 1980's. He points to the increase in average height and intelligence over the last half century. Both height and IQ has increased by about one standard deviation over the period 1940-1990. He points to better nutrition as the cause of both. Better nutrition tends to increase the size of the brain and potentially improve the brain's neurological development and functioning. Height is a commonly used measure of the nutritional status of young children and is the one we will adopt.

Most of the literature on the association between stature and cognition that Lynn (1990) refers to is somewhat primitive in their econometric approach. They are able to establish an association, but not causality. There are two issues facing us in this context. The first issue relates to the unobserved effect of home environ-

ment on both early childhood height and school age cognitive abilities. Second, parental decisions, genes or other biological mechanisms may very well determine both height and cognitive abilities. We will include a number of household control variables to account for parental- and household characteristics. Household fixed effects will also be employed to more fully capture unobserved household characteristics. Since genes are generally hard to control for without a data set of monozygotic twins, we attempt to deal with these issues through instrumental variables. Most modern research in this area, has acknowledged the importance of the prior issue, but the effects of the second are largely ignored or dismissed as unimportant.

In a developed country setting, there has been relatively little research dealing specifically with the causal relationship between early childhood height and cognitive development. As mentioned above, Case & Paxson (2008) is the most recent contribution. The authors examine the relationship between height and wages in a cohort of British men and women followed over a 30-40 year period. Their conclusion is that the height premium is a result of early childhood circumstances. Central to their argument, is the association of z-score at the age of five and school age cognitive ability. They find using maternal fixed effects, that z-score at the age of 5-6 years is significantly associated with a battery of cognitive test scores both at the time of height measurement and later in school age. Their results indicate that the effect of a one standard deviation increase in z-score at age 7 is comparable to a two standard deviation increase in log household income of the children in the study. In contrast to the results presented here, Case & Paxson (2008) find a significant decrease in the size of this coefficient when controlling for maternal fixed effects compared to estimates using extended family and individual controls. By using maternal fixed effects, they effectively control for the genetic material that is shared between siblings. Sundet et al. (2005) however, found in their analysis of monozygotic and dizygotic twins that the shared environment of dizygotic twins only account for 65% of the observed height-intelligence correlation.

Almond (2006) exploits the 1918 influenza pandemic to estimate how children who experience suboptimal gestational environment suffer consequences into adult life. He finds that people who had in utero exposure to the pandemic displayed lower educational attainment, increased probability of attracting a physical disability, lower incomes and socioeconomic status, and higher transfer payments.

These results are found comparing the crisis cohorts to the trend prior and after the pandemic. Almond suggests that the pandemic is a good natural experiment to test the fetal origins hypothesis (Barker, 1992). Barker introduced the idea that the lack of proper nutrition in utero could explain certain chronic health conditions later in life. We analyse a different outcome variable than Almond (2006) and exploit a period of crisis as an instrumental variable rather than comparing crisis cohorts to cohorts who did not experience the crisis in utero using OLS. Chapter 1 demonstrated that children who were exposed to the financial crisis in utero were relatively stunted compared to Indonesian children who did not. This result provides the first stage in testing the fetal origins hypothesis and one which is difficult to document for the influenza pandemic due to the lack of data.

More effort has been put into research focusing on developing countries. Most of this research however is very specific and has very small sample sizes. Several studies has dealt with experimental nutritional interventions which often focus on giving young children specific nutrients and testing the effect on the intervention on cognitive ability shortly after nutrients have been provided. The Institute of Nutrition of Central America and Panama (INCAP) conducted a study of 4 Guatemalan villages between 1969 and 1977. Martorell (2008) is one of the most frequently quoted analyses of this study. Randomly selected children between the ages of 0-24 months were given a protein enhanced nutrition supplement and their progress tracked. These individuals were interviewed and tested in adulthood at an average age of 30 years. They find that subjects given the supplement scored 3.46 points (mean score approximately 68) higher on a reading comprehension test, but this effect was not statistically significant. These subjects also scored 2.16 points (mean score approximately 18) higher on a cognitive functioning test (Raven Progressive Matrices Test) similar to the IFLS test and this effect was statistically significant.

Rose (1994) examines the association between visual recognition memory and cross-modal transfer and measures of infant growth. In the study they tested 183 infants aged 5-11 months living in Bombay, India. She compared the average test performance of under weight children and those who had adequate weight in a weight-for-age and weight-for-length sense. The results show that cognitive performance was significantly associated with physical growth after controlling for family characteristics, history of illness and birth weight. These controls are

also significantly associated with the test outcome. The author acknowledges the difficulties of directly controlling for all variables which constitute the level of nurturing in the home environment. The study does not address the issue of simultaneity other than controlling for parental height and weight.

There are a number of nutritional intervention studies[14] and most show some relationship between better nutrition in childhood and later cognitive functioning. Such studies require a great deal of resources and tend to be very specific in nature. It is not clear that the results from each of these studies are possible to generalize to other settings and other nutritional supplements. Our analysis is more general in scope as we focus on how variation in z-score will affect later cognitive functioning. These results could prove very useful when analysing the cost effectiveness of nutritional interventions to pregnant women and infants in poor countries.

Our results show that IV estimation increases the effect of early childhood z-score on school age cognition compared to OLS. We find that a one standard deviation increase in early childhood z-score improves cognitive test scores by 8% on average. Combining IV and household fixed effects does not change the coefficient significantly.

We start by presenting the estimation problem we face and the possible solutions to the econometric issues. Section 2.3 presents the data and discusses limitations to our analysis. The following section presents results and their potential weaknesses. The final section concludes.

2.2 Estimation problem

The exposition below draws heavily on the work of Glewwe & Jacoby (1995) and Alderman et al. (2006). As these papers and others have provided the behavioural framework for the type of problem we analyse in this paper, we will only briefly present the main econometric issues we face. The reduced form equation for early childhood health can be written as follows:

$$H_{ih1} = \alpha X_{ih1} + \upsilon_{ih1} \quad (2.1)$$

H_{ih1} denotes the health of child i in household h in Period 1. The period of the

[14] see ex. Wanatabe et al. (2005), Freeman et al. (1980), Grantham-McGregor et al. (1997) and Waber et al. (1981)

child's life which precedes school age is labelled Period 1. H_{ih1} is determined by a vector of individual-, household- and community characteristics, X_{ih1}, which have been found in existing literature to play an important role in determining early childhood (Period 1) health. Height is a frequently used proxy for child health and reflects history of disease, genetics and nutritional status. In the following we will use health and height interchangeably. υ_{ih1} is the composite error which contains an individual term, a household term and a white noise error term:

$$\upsilon_{ih1} = \epsilon_i + \epsilon_h + \epsilon_1$$

The unobserved individual characteristics could be genetic endowments or other time invariant attributes. ϵ_h captures time invariant unobservable household characteristics such as parent's tastes and discount rates.

The relationship of primary interest here is the association between Period 2 cognitive test score (A_{ih2}) and Period 1 health. This relationship can be illustrated by the following:

$$A_{ih2} = \beta X_{ih2} + \alpha_2 H_{ih1} + \eta_{ih2} \text{ where } \eta_{ih2} = \varpi_i + \varpi_h + \varpi_2 \quad (2.2)$$

This structure creates two immediate problems for OLS estimation. $E(H_{ih1}\varpi_i) \neq 0$ and $E(H_{ih1}\varpi_h) \neq 0$ in (2.2). The former problem might arise if parents make nutritional investments in their children in Period 1 conditional on future learning ability. Parents who expect their children to do badly in school might provide less nutrition, or more to make sure their children will be useful as labour. The resulting bias is indeterminate. Genes or other biological processes in utero might also potentially determine both height and cognitive development. Berger (2001) suggests that the role of insulin-like growth factors could affect both in early life. To account for the potential simultaneity of A_{ih2} and H_{ih1}, we will employ an instrumental variables approach.

Our second problem with OLS relates to the fact that $E(H_{ih1}\varpi_h) \neq 0$. The attitude and preferences of parents could determine early childhood health, both in utero and after birth, and cognitive development. As we know from the literature, early childhood- and in utero nutrition is particularly important in determining early childhood growth[15]. To account for the potential bias this issue creates, we will control for household- and maternal characteristics. Our sample contains some

[15] see Barker (1998)

siblings and we will use this smaller sample to run household fixed effects. Community characteristics could partly determine both the early childhood nutritional status through the quality of health care or through the community pooling of resources to insure against for example famine. Likewise, communities might differ in the quality of education. Our data set is not extensive enough to run community fixed effects. However, most siblings have grown up in the same communities and attended the same schools.

Addressing the first estimation problem requires valid instruments. We need an instrument which has a sizeable effect on H_{ih1}. It also needs to be transitory enough to vary over the children in our sample and not affect A_{ih2} directly, but only through its effect on H_{ih1}. Chapter 1 documented the negative effect of the South-East Asian financial crisis on the z-score of Indonesian children. We employed a cohort difference-in-difference approach and found that children born in the years 1998-2000, had 0.28 points lower z-score compared to children who were born prior to the crisis. Our instrument is a dummy variable indicating whether a child was born during what we define as the crisis period. As is common in the literature, H_{ih1} will be measured using height-for-age z-score. This measure is found by comparing the height of a child to the appropriate age and gender group of a healthy, American reference group[16]. Z-score is the number of standard deviations difference from this reference population. The advantage of this measure is that we purge the data of most of the age and gender specific components. The crisis in Indonesia started in the fall of 1997 and by 2002, GDP per capita had fully recovered[17]. We exploit this period of crisis (1998-2000) to identify differences in cognitive ability by comparing children born during the crisis to those born prior.

It is difficult to determine the bias of OLS estimates ex ante. If we are not able to control for all household characteristics directly using proxies for material resources and the level of nurturing, we would expect the OLS estimates to be biased upwards. It is likely that a nurturing family environment is positively correlated with both height and cognitive ability. If there is measurement error in our height measurement, we would expect to see a downward bias in OLS estimates. The provision of nutrition conditional on future learning potential, could bias the estimates either way. Parents could provide children who are expected to do well

[16] US National Center for Health Statistics' (NHCS) height-for-age tables
[17] World Bank Development Indicators

academically less food compared to children who were expected to contribute more to the household income throughout their childhood. As we measure H_{ih1} prior to age four, it is possible that some parents will feed their children based on their view of the cognitive potential of that child. The overall effect on OLS estimates is not clear and only the results will provide an answer.

2.3 Data

The data used in this paper is exclusively collected from the Indonesian Family and Life Survey[18] (IFLS) of households and communities. IFLS is an ongoing longitudinal survey with data collection rounds in 1993 (IFLS1), 1997 (IFLS2), 2000 (IFLS3) and 2007 (IFLS4). Excluding early childhood z-scores, all the data used here is from IFLS4. The IFLS questionnaire consists of questions covering socioeconomic, health status and community variables. The health information was collected by a trained nurse and height measurements were taken lying down for the youngest children to allow the spine to be fully extended. The IFLS4 covered 13 535 households and interviewed a total of 44 103 individuals. These households were drawn from 14 of the 27 Indonesian provinces containing about 83% of the country's population.

The full sample used here, consisted of 2250 children in primary school education. We need information on the following six variables for a child to be included in our full sample:

* Cognitive test score, IFLS4
* Age when interviewed, IFLS4
* Gender, IFLS4

[18] IFLS1 (1993): Frankenberg, E. and L. Karoly. "The 1993 Indonesian Family Life Survey: Overview and Field Report." November, 1995. RAND. DRU-1195/1-NICHD/AID

IFLS2 (1997): Frankenberg, E. and D. Thomas. "The Indonesia Family Life Survey (IFLS): Study Design and Results from Waves 1 and 2". March, 2000. DRU-2238/1-NIA/NICHD

IFLS3(2000): Strauss, J., K. Beegle, B. Sikoki, A. Dwiyanto, Y. Herawati and F. Witoelar. "The third Wave of the Indonesia Family Life Survey (IFLS3): Overview and Field Report". March 2004. WR-144/1-NIA/NICHD

IFLS4 (2007): Strauss, J., F. Witoelar, B. Sikooki and A.M. Wattie. "The Fourth Wave of the Indonesia Family Life Survey (IFLS4): Overview and Field Report". April 2009. WR-675/1-NIA/NICHD

* Household per capita expenditure, IFLS4

* Mother's highest level of education, IFLS4

* Height measurement from a previous IFLS survey prior to the age of five years, IFLS1-3

The latter variable proved to be the main qualifier for inclusion in our sample. Children of school age in IFLS4 could have their preschool height measured in either of the preceding waves or none of these. Of our full sample, 192 children had their preschool height (z-score) measured in IFLS1, 1078 in IFLS2 and 980 in IFLS3. The full sample does not consist of siblings only and the usable sample for household fixed effects is significantly smaller. Following Alderman et al. (2006) and Hoddinott & Kinsey (2001) we excluded all children with a z-score of more than six in absolute value to remove extreme outliers and potentially faulty recordings. In the full sample, we have included all children who have had height measurements taken up to the age of 5 years. The official primary school enrolment age in Indonesia is six years, but some children in our sample started at five. The height measurement should capture the stature of the children prior to school age. Table 2.1 describes the data used.

Panel A describes the data used for non fixed effects analysis, while Panel B consists of sibling groups. The restricted sample accounts for the average differences in age at height measurement between the crisis and the non-crisis children. To construct the restricted sample, we constrain the non-crisis group to children who had their height measured within the same age range as the crisis group. In addition, we drop all children measured in IFLS1 (1993) to limit the influence of time trends in z-scores. This means that only non-crisis children who had height measurements taken between the ages of 0-35 months are included in our non household fixed effects sample and 0-33 months in the fixed effects sample. Some of the older children at height measurement did not have siblings in the restricted sample which explains the difference in age range at measurement between the two samples. The major differences between the treated and untreated children are their age when completing the cognitive test and the age when height was measured. Crisis children were born in the period 1998-2000, while the non-crisis children were born prior to 1998. As long as we restrict childhood z-scores to be measured up to the age of 60 months as in the full sample, the crisis children will be younger on average when measured since these children had a maximum age at

height measurement of 35 months. We can see from both panels that the average age of crisis children at height measurement is approximately half of the untreated children. The restricted sample limits the non-crisis children in our sample to be within the same age range as the crisis children at height measurement. We discuss further below why this might be an important issue to deal with in this context and we will use the restricted sample in what follows. Crisis children on average have more educated mothers which may reflect the overall increase in education for women over this time period.

2.3.1 Outcome variable

We are primarily interested in the causal interaction between early childhood health and schooling outcomes. The data provided on formal tests taken throughout school, is patchy and does not provide enough data points for a thorough analysis. The use of the crisis variable as an instrument further limits the range of outcome variables as the children who underwent the Indonesian experience of the South-East Asian crisis in utero, are at a maximum age of ten at the time of IFLS4.

Cognitive test scores RAND has since IFLS2, included a cognitive test which most interviewed children have completed. The test consists of 12 questions for children between the age of 7 and 14. The individual questions all consist of adding a missing part to a geometric shape. The test is based on the Raven Progressive Matrices (RPM) developed by Dr. John C. Raven in 1936. RAND delivered a test which included some of the questions from the original test, but not all due to time constraints in the interview situation. The questions got increasingly difficult throughout the test. RPM was developed to test general intelligence and is widely used in similar research. The test data from IFLS2 is not directly comparable to the test given in the subsequent phases. The test score data used in our analysis is all taken from IFLS4 and there is a close relationship between age and score as can be seen from Figure 2.1.

The concave relationship between test scores and age necessitates the inclusion of age variables in the structural equation since crisis children in IFLS 4 are younger on average than non-crisis children at the time of cognitive testing. This concave relationship is present both in the IFLS3 and IFLS4 data.

2.3.2 Instrumental variable

Our instrument is a dummy variable indicating whether the child had in utero exposure to the South-East Asian financial crisis. Many commentators place the start of the crisis at the time when the Indonesian central bank decided to float the Rupiah in August 1997. This decision was made after the currency had come under severe pressure and the Rupiah quickly depreciated as a result. Chapter 1 describes the time line of the crisis in more detail. The real GDP of Indonesia fell by 13.1%[19] in 1998 and did not recover to pre-crisis levels until 2002. We follow Chapter 1 and set the cut-off point at January 1st 1998. Children in our sample who were born after this date, are labelled crisis children, while everyone born prior to this date, were not affected for the purposes of this analysis. The cut-off point is set to allow crisis children to experience at least one trimester of crisis in utero. Children with less than 4 months of in utero exposure or no in utero exposure at all, are assigned to the non-crisis group.

Chapter 1 documented by using a cohort difference in difference approach, that children born during the crisis lost approximately 0.28 z-score points. In the raw data, we find very little absolute difference in z-scores between children born during the crisis and children born in 1993. Our results from Chapter 1, suggest that average z-scores seem relatively stable between 1993 and 2000 for children that are of crisis age. The problem with ending the analysis there is that there are clear trends in children's z-scores over the period leading up to the crisis. This is a period where the Indonesian economy was growing at a rate of 7.1% (real GDP) and the Indonesian government was investing heavily in education and health infrastructure (Lanjouw et al., 2001). The absolute effect of the crisis might not be noticeable in the raw data, but the decline relative to pre-crisis trend is. In Chapter 1 we showed this by comparing the changes in z-score for crisis age children between 1993 and 2000 compared to the development over the same period for children of an older age group that even though they experienced the crisis, were too old to have had their z-scores affected (Martorell, 1995 and Hoddinott & Kinsey, 2001). Chapter 1 also shows that there was little effect of the crisis on children born just prior to the crisis, indicating that a large part of the crisis effect comes through intrauterine effects. Most of the effects of the crisis affected urban children, but we found some effect on rural children towards the end of the crisis.

[19] World Bank Development Indicators

The inclusion of a quadratic in age in our structural equation is justified by Figure 2.1 and creates a quadratic time trend in our first stage since age in 2008 is highly correlated with year of birth. It is not perfectly correlated since the interviews in IFLS4 were conducted over almost a year. Two children born at the same date could therefore have been interviewed and tested at different ages in IFLS4. The correlation between age and birth year is 98.54% in the restricted sample. Our first stage shows that there is a clear association between the crisis dummy and early childhood z-score conditional on the included covariates.

Children in the IFLS are not all measured at the same age. This further complicates our analysis if z-scores depend on the age when it was measured. Z-score measures the difference between a child's actual height and the height of a healthy reference population (NHCS) of the same sex and age measured in standard deviations. If there is a systematic relationship between z-score and the age of measurement, we could well obfuscate the relationship between the crisis variable and z-score. Figure 2.2 shows that there seems to be such a relationship for children in the IFLS4 sample. These children were born years after the 1998 crisis and should not have been directly affected by it. There is a tendency for Indonesian children to be closer to its reference counterparts just after birth. The vertical axis measures age in months. The z-score-age relationship seems to disappear somewhere after the age of 18 months. This finding echoes the results in Chapter 1 which showed that the coefficients on monthly age dummies in a regression with z-score as the independent variable, stabilize after the age of 22 months and that older children on average have lower z-scores. The age at height measurement will therefore be inversely related to z-scores and crisis children are on average younger than their untreated counterparts. This will tend to decrease the significance of our instrument and the overall statistical power. We try to solve this issue by running our analysis on the restricted sample which includes untreated children who were measured within the same age range as the crisis children.

Table (2.2) shows the result of estimating the first stage as in Equation (2.3).

(2.3) $Z_{ih1} = \beta_1 X_{ih2} + \alpha_3 crisis_{i1} + \varepsilon_{ih2}$

Z_{ih1} denotes the z-score of child i in household h in Period 1. X_{ih2} contains all the controls included in the estimation of Equation (2.2). $crisis_{i1}$ is a dummy variable of value one if the child was born during the crisis period, and zero otherwise.

The crisis variable is negative and statistically significant. Our results indicate that conditional on the exogenous covariates, children who were born during the crisis suffered a z-score loss of approximately one point. The F-statistic on the crisis instrument is large and statistically significant at 1%. The coefficients on the age variables reflect the positive trend in z-scores of children over the period of analysis. We find that urban children and children from wealthier households are healthier on average. Children with more educated mothers have higher z-scores, and it seems like there is a dichotomy between mothers with schooling from the two lower categories and mothers with senior high school or a higher degree. The last group have children with better health. The coefficient on the crisis dummy is smaller (0.795) when using the full sample as we would expect and the F-statistic is less than half of what we found in Table 2.2. The coefficients on the other variables are similar in size and significance.

Our crisis children are the youngest when having their cognitive score measured. Since there is a clear relationship between age and cognitive score in both IFLS3 and IFLS4, the inclusion of age controls is crucial for our exclusion restrictions to hold. Given age, there are no obvious reasons why a child born in the crisis period should have lower test scores other than through the potential effect of early childhood z-score. As we are using a single instrument, we cannot test this assumption directly in the data. With the exception of age, the gender dummy is the only exogenous covariate in our specification of Equation 2.2 which will not be filtered out by employing household fixed effects. We did run probits for both samples with crisis status as the dependent variable and all the included instruments in Column (1) of Table 2.1, and there is no significant association between gender and crisis status. Our crisis children started school 4-7 years after the crisis ended and it seems unlikely that the crisis would have differentially have affected the schooling quality of these children compared to the older children in our sample. For this to be the case, we would have to believe that there was a clear deterioration in schooling quality of the youngest kids, but not the older ones, starting 4-7 years after the crisis ended. It seems like if the crisis would have affected the educational system in Indonesia negatively, it is more likely that children of all ages would have been impacted equally, not just the ones born after the onset of the crisis. None of the children in the restricted sample were enrolled in primary school during the crisis.

Table 2.3 shows the reduced form results as in Equation (2.4).

(2.4) $A_{ih2}= \beta_2 X_{ih2}+\alpha_4 crisis_{i1}+\epsilon_{ih2}$

A_{ih2} denotes the IFLS test score of child i from household h in Period 2. X_{ih2} contains all the controls included in estimating Equation (2.2). $crisis_{i1}$ is our crisis dummy. The coefficient on this variable is negative as we would expect. The magnitude of α_4 implies that children born during the crisis had 15% of a standard deviation lower test scores than children born prior to the crisis. This coefficient is 67% larger in absolute value than the corresponding coefficient using the full sample. There is a more continuous and clear association between the maternal educational variables in the reduced form compared to the first stage.

2.3.3 Selection bias

The available sample of children found in IFLS1-3 with height recorded at preschool age, is not complete in IFLS4. Of our potential sample of 7566 children, 2712 of them are missing in IFLS4. To qualify for our potential sample, the child would need to have been at preschool age during one of the three first rounds of the IFLS data collection. Most of these missing children had their height recorded in 1993 and have not been re-interviewed. Only spilt off household from the original 1993 survey which included the main respondent from IFLS1 were followed after 1993. Of the 2712 missing children in 2007, 2155 are from IFLS1. Earlier IFLS respondents were not re-interviewed if they moved out of the designated IFLS areas. 13 out of 27 provinces were excluded mainly for cost-effectiveness reasons. The remaining provinces contained approximately 83% of the population. It seems likely that most children missing from the 2007 sample lived in households which either moved to a non-IFLS province or split from one of the original IFLS1 households between IFLS1 and IFLS2 without the main IFLS1 respondent moving with them. We are not able to tell from the data whether mortality explains a non-trivial portion of the sample attrition.

Whatever the reasons for this rather large loss of observations, we need to worry whether selective sample attrition might bias our estimates. Table 2.4 presents the results of running a probit regression of a dummy which indicates whether a child part of our potential sample, was present in the IFLS4 sample. A child is part of the potential sample if their preschool height was recorded in IFLS1 (1993), IFLS2

(1997) or IFLS3 (2000). The dependent variable takes the value one if the child who was part of our potential sample was missing from the IFLS4 sample. Column (1) shows the result of regressing this dummy on early childhood z-score only. The result seems to indicate that children with higher early childhood z-scores are more likely to be part of our sample than more stunted children. However, the picture changes when we add more controls in Column (2). The coefficient on z-score changes sign and becomes insignificant. As early childhood z-score is our main variable of interest, we would like to make sure that our sample does not suffer from selective sample attrition along this dimension. Column (2) seems to provide some reassurance that there is little systematic selection depending on early childhood z-score.

We included dummies for all the pre-IFLS4 phases of the survey as controls in Column (2). This survey round fixed effects regression shows that we are less likely to find children measured in 1993 and 1997 in the sample compared to the excluded dummy indicating the 2000 survey. This effect is much more pronounced for IFLS1 and we mentioned above why this is the case. Besides the natural log of per capita household expenditure (logpce), none of the other controls are statistically significant. Our results seem to indicate that we are more likely to observe children from less affluent families in the 2007 survey. We are unaware of any good reason for observing this particular pattern.

The results from Table 2.4, goes some way to alleviate concerns that our results are driven by sample selection. Children observed in an earlier round are most likely to disappear from our IFLS4 sample if their family moved prior to IFLS4. Either they moved out of the IFLS areas or they split off from one of the original IFLS1 households leaving the original survey respondent behind. The reason for these moves could be many, but they do not seem to be systematically associated with the stature of their young children.

2.4 Results

As we saw in Table 2.1, there is a large gap in the average age between crisis and non-crisis children at the time of early childhood height measurement. Crisis children have been measured at an earlier age on average which leads to inflated z-score of these children due to the age-z-score relationship illustrated in Figure 2.2. There are several ways to go about correcting for the issues that this relationship

presents. We could include controls for age at height measurement directly into the structural equation. The children in our full sample had their heights measured over the first 3 IFLS phases, i.e. either in 1993, 1997 or 2000. There is no a priori reason why children measured in the last phase should have been measured at an earlier age than children in earlier phases, but seems to be a feature of the data as suggested by Figure 2.3. Figure 2.3 shows the relationship between age of cognitive testing and the average age at height measurement in the full sample. There is a clear relationship between mean age at height measurement and age in the IFLS4 survey in our full sample. Older children in IFLS4 tend to have had their early childhood z-score recorded at a later age. This leads to a lower early childhood z-score ceteris paribus. This correlation is strong enough that if we choose to attempt to solve the issue by including age at height measurement controls directly, we experience the common problems with multicollinearity; inflated standard errors. The coefficient on z-score is very large (1.37), but is imprecisely estimated (t-statistic of 0.56).

The second potential solution, involves using age at measurement variables as instrumental variables. However, there is no reason to expect these variables to be powerful in the reduced form, leading to weak estimates of the main variable of interest in the structural equation. We have attempted to deal with the issue by constraining the non-crisis group to children who had height measured within the same age range as their crisis counterparts. As we can see from Table 2.1, the exclusion of older children at the time of measurement does not seem to alter the non-crisis group significantly along the other dimensions listed in Table 2.1. The new, restricted non-crisis group is younger on average at the time of height measurement, but seem similar to the full sample non-crisis group based on all other covariates.

Table 2.5 shows the results of estimating Equation (2.2) by OLS (Column 1) and using instrumental variables (Column 2). The OLS coefficient on z-score is one fourth the size of the corresponding coefficient using IV. We see that even though the effect of early childhood z-score is only statistically significant at 10%, the effect is quite large. A fall of one standard deviation in z-score (1.80) in the restricted sample would lead to a 0.69 points fall in cognitive score on average as measured by the IFLS test. This constitutes an approximate 8% decrease in test score compared to the mean. Case & Paxson (2008) found that a one standard

deviation change in early childhood z-score would have the same impact on verbal test scores as a two standard deviation change in log per capita household income. Our results indicate a larger effect. We found that the corresponding number given our data is almost a four standard deviation change in log per capita household expenditure.

Column (1) and (2) shows the age pattern we found in Figure 2.1. All Columns seem to indicate that boys do better than girls on the IFLS cognitive test. Others have found that boys often outperform girls when tested on spatial cognition[20]. Children from more affluent households tend to do better on these tests on average. Most of the research on the association between family income and the cognitive test scores of children has been done in a developed country setting. Smith et al. (1997) found that U.S. children living in relatively poor households had lower cognitive test scores both in early and middle childhood after controlling for maternal education, family structure and child specific covariates. There are many potential reasons for this relationship. These include the lack of provision of a stimulating home environment of relatively poor parents. A poor home environment may encompass both material resources and negative neighbourhood conditions since poor families are likely to live in poor communities. Poor parents are also more likely to send their children to less equipped schools and nursing schools both in terms of material resources and teacher quality. The inclusion of father's education in Column (2) does not diminish the relationship between household income and cognitive test scores and the variables are insignificant. Our estimates confirm the findings from developed countries in a third world country setting.

We find in both columns that urban children do better on average compared to their rural counterparts. Some of this relationship may be accounted for by better developed educational infrastructure in urban centres. The exclusion of the urban dummy in Column (2) does not increase the size of the coefficients on the maternal educational variables which suggests that the urban dummy does not pick up some of the effects from the higher educational levels in urban areas on average. There seems to be a positive, monotone and close to linear relationship between the cognitive scores of children and the level of education of their mothers. This effect is quite large. Children born of mothers with college or university education, score on average 1.064 points, or almost half a standard deviation, higher on IFLS'

[20] see ex. Geary et al. (2000)

cognitive test compared to children born of mothers with elementary education only.

The estimate on early childhood z-score is about 40% smaller in the full sample and insignificant. The estimates on the controls in Column (2) are similar to those for the full sample. Using this sample, we find more precise estimates on the age variables while the magnitudes are similar. The full sample includes children who had height measured in IFLS1. The inclusion of these children increases the variation in the age variables and let us estimate those with more precision.

2.4.1 Fixed effects

Table 2.6 presents the first stage, reduced form and household fixed effects estimates of (2.2) combined with IV. The coefficient on the crisis dummy in Column (1) is approximately 60% larger in absolute value than what we found in Table 2.2. We might worry that the coefficient on early childhood z-score in Column (3) might suffer from weak instruments critique due to the somewhat low F-statistic (9.42) of the excluded instrument. However, since we are using a single instrument, we should expect our estimator to be approximately unbiased.

The reduced form estimate on the crisis variable in Column (2) is about 3 times the size of what we fund in Table 2.3 and statistically significant. Both this estimate and the estimate in Column (1) are larger in absolute value than the corresponding IV estimates. The FE-IV estimate contains the IV estimate in the 95% confidence interval. Figure 2.4 shows the distribution of age at measurement in months for the crisis and non-crisis group in our restricted, household fixed effects sample. Even though the sample used for FE is quite small and the estimates are not precise, some of the difference between the results presented in Table 2.6 and the IV only equivalents (Table 2.5) may be a result of differences in age at height measurement distributions in the two samples. Table 2.1 shows that the difference in median age at measurement between crisis- and non-crisis children is two months in the FE sample, while it is double in the non-FE sample. The closer the distribution of the two groups when it comes to age at height measurement, the larger coefficient we would expect to find in both the reduced form and the first stage. This is due to the z-score-age relationship documented in Figure 2.2 and which has been discussed above.

Case & Paxson (2008) found that adding mother fixed effects reduced the co-

efficient on z-score by one-third in a similar analysis of British youngsters. We do not find a similar effect. The FE-IV estimate in Column (3) is 28% larger than the IV equivalent. This may be the result of the change in age at height measurement distribution as discussed. However, the estimates presented in Table 2.6 are not adequately precise to conclude that the use of fixed effects changes the estimate significantly compared to instrumental variables only. The estimate in Column (3) implies that a one standard deviation in early childhood z-score (1.73) would increase test scores by approximately 40% of a standard deviation.

The fixed effects estimates on the control variables are all insignificant. In Column (3), the age variables have the wrong sign. When employing fixed effects, the sample size is greatly decreased and there does not seem to be enough variation in the data to estimate these coefficients precisely.

2.4.2 Discussion

The latter half of the preceding section analyses what we consider to be the main challenge with the data and estimation strategy employed here. We find that the effect of making our crisis- and non-crisis group more homogeneous with respect to the age at height measurement, changes the results notably in the direction we would expect. Since the non-crisis children are still three months older on average in our restricted sample, we believe that the results are likely to be even stronger if we had the luxury of measuring the children at exactly the same age. A measurement at the day of birth would not only overcome this problem in the data, but would also make it possible to assign all the effect to in utero effects. In our restricted sample, crisis children lived through the crisis in utero and through the age of 0-2.7 years. The non-crisis group had no exposure in utero and entered the crisis at the age of 3 months to 3 years. Due to the limited sample size, we are unable to finely separate groups according to the amount of exposure and the timing of the exposure. This question demands a lot from the data and is relatively unexplored in the literature.

The crisis group were born later than their non-crisis siblings. Not all of the potential crisis children had entered school when they took the IFLS cognitive test. The lack of any period of primary school education for some of the crisis children could have biased the results in the direction we have reported. To account for that, we excluded children who had not started school. Even though Indonesian children

are required to start school at the age of six, we see that a significant portion (28%) of the IFLS4 sample started late. Since the IFLS cognitive test is only performed on children who have reached the age of seven, there is some potential for testing children who have not yet started primary school education. We therefore only included children who had started school in our sample. This step could potentially have excluded the part of our sample that might have performed the absolute worst if there is a correlation between bad performance on cognitive test scores and a high enrolment age. The data however, shows that only 0.9% (32 children) of all children who could potentially have been included in our restricted sample, had yet to enrol. Employing household fixed effects would have effectively eliminated all but one pair of siblings from the sample and it is unlikely that this leads to a large underestimation of the true effect of childhood height on cognitive development.

Crisis children are by the nature of our sample, born later than their included siblings. Horton (1998) showed that higher order children are likely to have worse nutritional status due to competition over resources, maternal depletion and a greater likelihood of experiencing infection. If maternal depletion is a real concern in our data, we would expect to find that the exclusion restriction would be violated as maternal depletion could also have a direct effect on cognitive development. It is difficult to adjust for birth order directly in our analysis as almost all siblings are consecutive siblings in our data. A birth order variable would therefore capture largely the same information as the instrumental variable. In a sibling pair, the higher birth order sibling would also be the crisis child. We have run all our non FE regressions including a log household size variable which turns out to be insignificant in estimations of both the structural equation and the first stage and does not change our main results. This variable does not capture exactly the same information as a birth order variable as adults may move in and out of the household, but gives some reassurance that our results are not driven by competition for resources.

A related issue to the latter is the question of birth spacing. If potential mothers delayed having children during the crisis, the crisis children would have on average been born into households with less competition for maternal and material resources. If this were the case, our estimates could underestimate the true effects. With the data at hand, it is difficult to test this hypothesis directly due to the same difficulties of including a birth order variable. A back of the envelope exercise

is to calculate the number of births divided by the number of households in each wave up to and including, IFLS3 (2000). The number of children registered in each wave is very stable compared to the number of households interviewed. The relative short duration and sudden onset of the crisis, also makes it less probable that changes in birth spacing would affect our results significantly.

2.5 Conclusion

Mapping out the full effects of low early childhood health is important for policy decisions. We know that people who experience periods of poor nutrition early in life, will suffer health consequences throughout their life. There has been relatively less attention paid to the question of whether early life health will affect cognitive development. This question got some attention in the 70s and 80s, but most of the research focused on correlations. Intuitively we would expect parents who will not or cannot provide a proper diet to their youngster, to also neglect their cognitive development. Later papers by ex. Case & Paxson (2008) control for this using maternal fixed effects and find that early childhood z-score has a positive effect on later cognitive ability.

Most other research in this field employs fixed effects only. In another, related field IV coupled with FE has been the accepted approach since the middle of the 90s. Glewwe & Jacoby (1995) suggested this approach to the question on how early childhood nutrition affected school enrolment age. If parents make nutritional investments in their children conditional on future learning potential, instrumental variables are necessary to find the causal effect. Similarly, if there are genetic or other biological connections between height and intelligence in early life, instrumental variables could be necessary to show the true causal effects. We have analysed the question of how early childhood nutrition causes school age cognitive development using both IV and household fixed effects.

Finding the causal effect is important for policy development. If the height-intelligence correlation is all due to quality of parenting, policy should be aimed at increasing the awareness of parents. Even though this is undoubtedly a valuable effort, emphasis must also be put on improving the situation for pregnant women and infants. High quality parenting will according to the results presented here, not necessarily compensate for periods of poverty early in life. Our results show that the effects of early childhood health on later cognitive development can be

large even after controlling for household material resources and time invariant parental- and household characteristics. The use of IV increases the size of the coefficient on early childhood z-score by a factor of four compared to OLS. We find that an increase in early childhood z-score of one point causes an increase in school age cognitive test score of approximately 16% of a standard deviation. Our results are qualitatively very similar in effect to those found by Alderman et al. (2006).

As noted by Glewwe & Jacoby (1995), the proper estimation strategy demands a lot from the data. This is likely to be the reason why other papers have argued that the issues listed above are unlikely to be very important. We find the opposite to be the case.

Figure 2.1: Relationship between age and cognitive score for IFLS3 and IFLS4

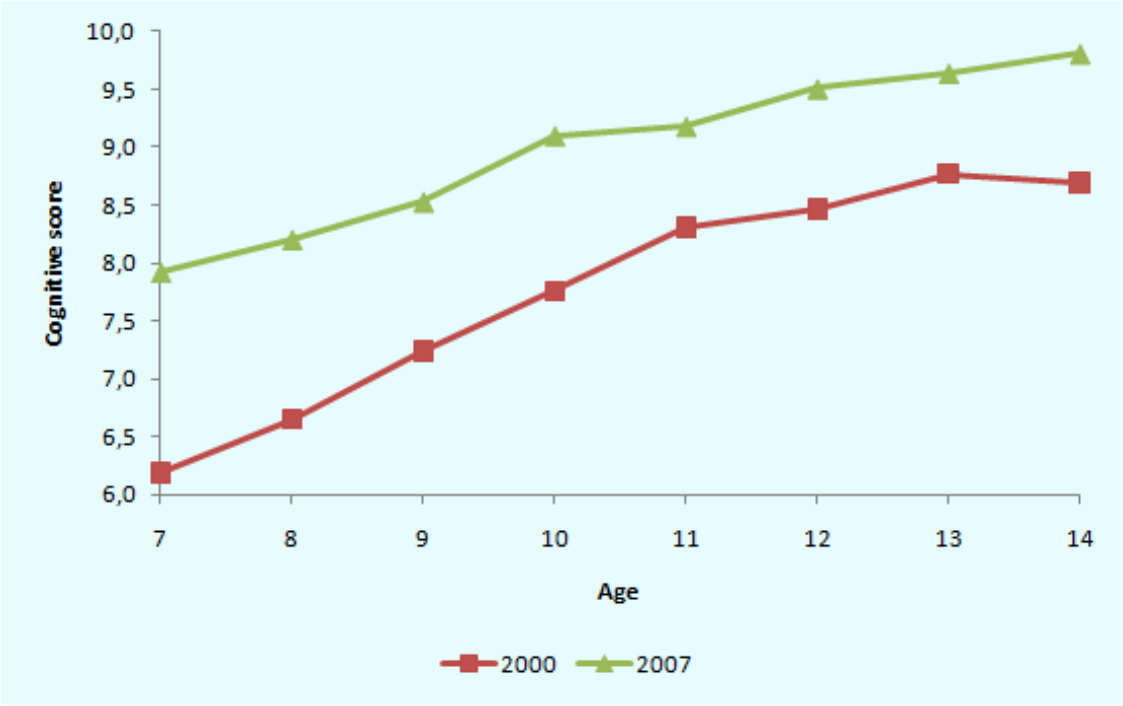


Figure 2.2: Relationship between age and z-score in IFLS4

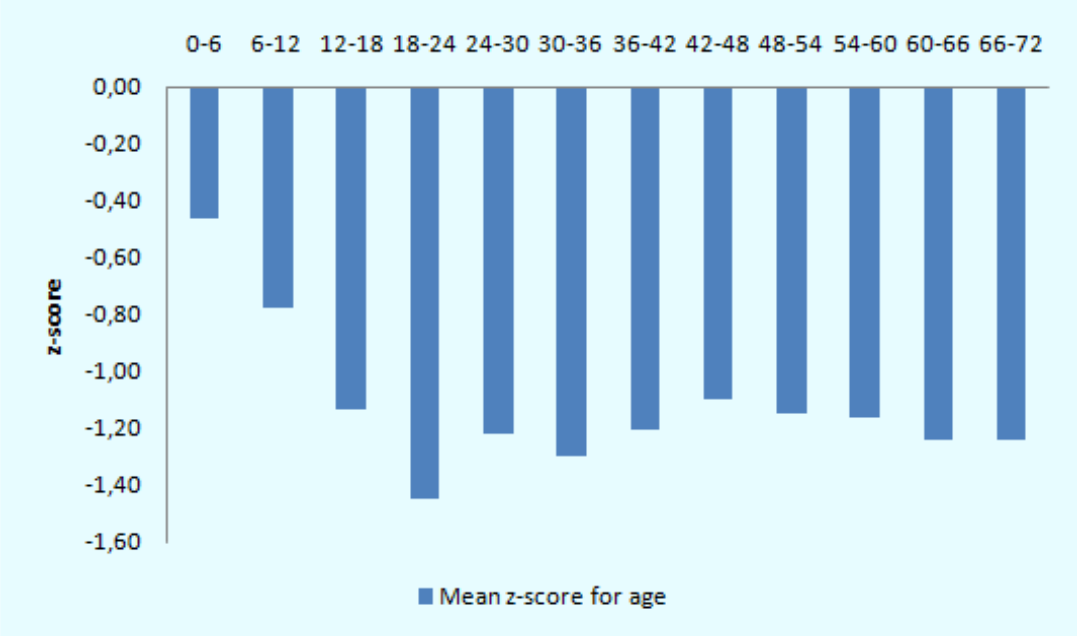


Figure 2.3: Relationship between average age at height measurement and age at cognitive testing in full sample

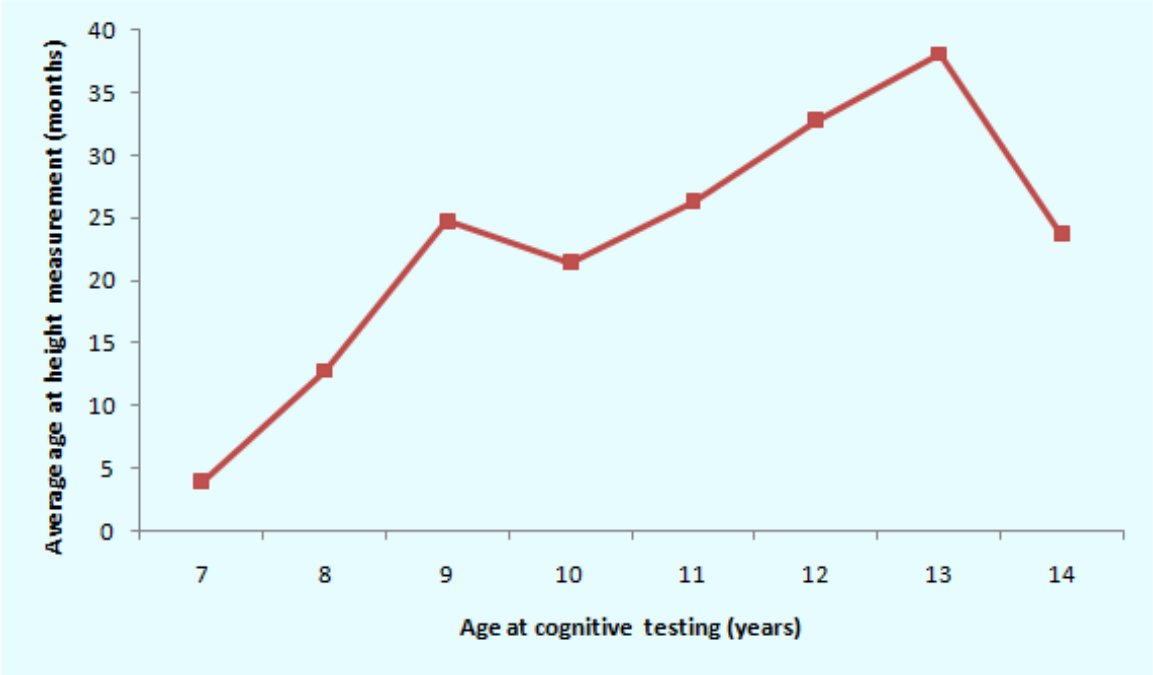


Figure 2.4: Density plots of crisis and non-crisis children in the restricted fixed effects sample

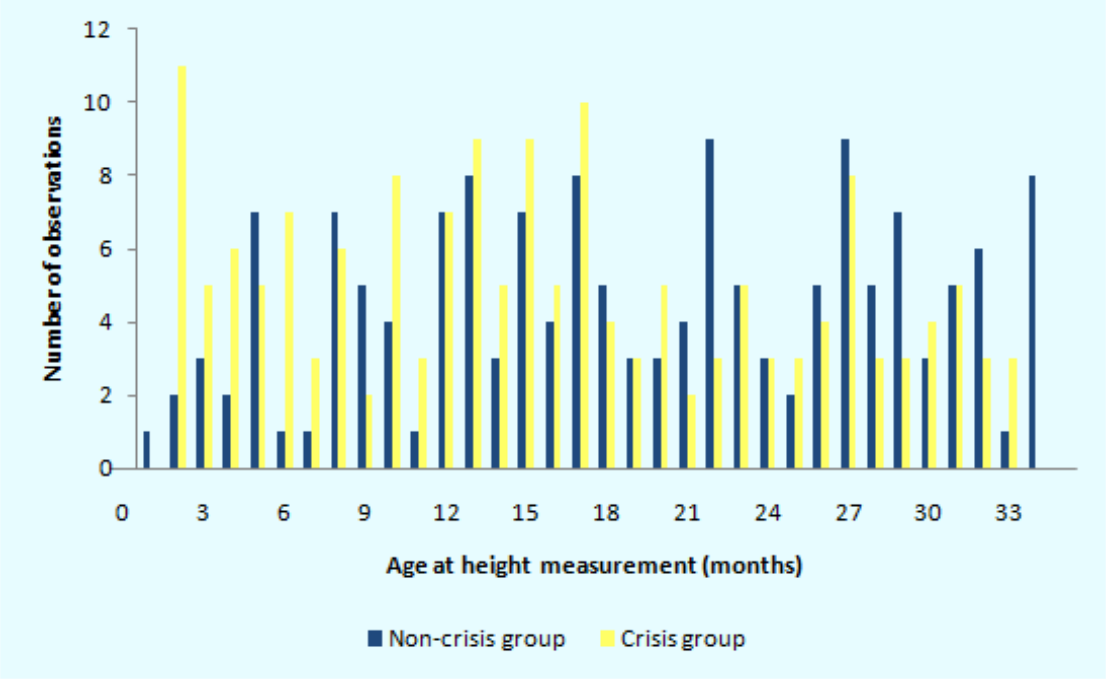


Table 2.1: Summary statistics of full- and restricted sample

Panel A: Non household fixed effects sample

	(1)	**(2)**	**(3)**
Sample	**Restricted**	**Full**	**Both**
Crisis status	Non-crisis	Non-crisis	Crisis
Observations	813	1550	698
Male	49.45%	50.00%	50.86%
Urban	50.92%	49.97%	51.00%
Logpce	13.26	13.29	13.29
Age at testing (years)	11.21	11.84	8.26
Age range testing	10-14	10-14	7-10
Average at measurement (months)	18.99	27.00	16.31
Age range at measurement	0-35	0-60	0-35
Median age at measurement	19	27	15
Mother's education			
Elementary school	54.49%	54.26%	45.70%
Junior high schooll	18.57%	19.03%	22.23%
Senior high school	19.31%	19.55%	25.64%
College/University	7.63%	7.16%	6.30%

Notes: The *restricted sample* includes *non-crisis* children who had their height measured within the same age range as *crisis* children. The *full sample* consists of all children who completed cognitive tests in IFLS4 and had their height measured prior to the age of 5 years. *Crisis* children were born 1998-2000. *Logpce* is the log of monthly per capita household expenditures. *Testing* refers to cognitive testing while *measurement* refers to height measurement.

Table 2.1 cont.: Summary statistics of full- and restricted sample

Panel B: Household fixed effects sample

	(1)	(2)	(3)	(4)
Sample	**Restricted**		**Full sample**	
Crisis status	**Crisis**	**Non-crisis**	**Crisis**	**Non-crisis**
Observations	162	153	354	407
Male	47.53%	52.94%	52.26%	52.83%
Urban	47.53%	46.41%	51.41%	51.11%
Logpce	13.19	13.08	13.19	13.12
Age at testing (years)	8.09	11.29	8.14	11.92
Age range testing	7-10	10-13	7-10	10-14
Average at measurement (months)	14.70	18.09	15.05	30.38
Age range at measurement	1-33	0-33	1-33	0-60
Median age at measurement	14	16	14	33

Notes: The *restricted sample* includes *non-crisis* children who had their height measured within the same age range as *crisis* children. The *full sample* consists of all children who completed cognitive tests in IFLS4 and had their height measured prior to the age of 5 years. *Crisis* children were born 1998-2000. *Logpce* is the log of monthly per capita household expenditures. *Testing* refers to cognitive testing while *measurement* refers to height measurement.

Table 2.2: First stage regressions with restricted sample

	(1)
LHS variable	Early childh. z-score
Sample	**Restricted**
Crisis	-0.919**
	(-5.14)
Age	-1.105**
	(-3.42)
Age squared	0.035**
	(2.20)
Male	0.051
	(0.57)
Logpce	0.106**
	(2.31)
Urban	0.294**
	(3.10)
Mother's education	
Junior high school	-0.049
	(-0.41)
Senior high school	0.281**
	(2.31)
College/University	0.219
	(1.16)
Constant	5.024**
	(2.86)
F-test of instrument	F(1,1501)=36.717
Observations	1511
R-squared	0.067

Notes:*z-score* is the difference between actual height and the height of a same age and gender reference population, measured in standard deviations. The *restricted sample* includes *non-crisis* children who had their height measured within the same age range as *crisis* children. The *full sample* consists of all children who completed cognitive tests in IFLS4 and had their height measured prior to the age of 5 years. *Crisis* children were born 1998-2000. *Logpce* is the log of monthly per capita household expenditures. *Urban* is a dummy of value 1 if the household resides in an urban area. Elementary school is the excluded educational category. **p<0.05, *p<0.10. t-stats in parentheses.

Table 2.3: Reduced form with restricted sample

	(1)
LHS variable	Cog. score
Sample	**Restricted**
Crisis	-0.353
	(-1.59)
Age	0.480
	(1.17)
Age squared	-0.012
	(-0.59)
Male	0.361**
	(3.17)
Logpce	0.195**
	(3.35)
Urban	0.486**
	(4.03)
Mother's education	
Junior high school	0.337**
	(2.22)
Senior high school	0.715**
	(4.62)
College/University	1.137**
	(4.73)
Constant	6.779**
	(3.92)
Observations	1511
R-squared	0.119

Notes: The *restricted sample* includes *non-crisis* children who had their height measured within the same age range as *crisis* children. The *full sample* consists of all children who completed cognitive tests in IFLS4 and had their height measured prior to the age of 5 years. *Crisis* children were born 1998-2000. *Logpce* is the log of monthly per capita household expenditures. *Urban* is a dummy of value 1 if the household resides in an urban area. Elementary school is the excluded educational category. **p<0.05, *p<0.10. t-stats in parentheses.

Table 2.4: Probits testing for sample selection

	(1)	**(2)**
LHS variable	Not in IFLS4	Not in IFLS4
z-score	-0.010**	0.008
	(-2.32)	(1.49)
Height recorded 1993		2.588**
		(45.95)
Height recorded 1997		0.262**
		(4.79)
Male		-0.033
		(-0.87)
Logpce		0.136**
		(5.98)
Urban		0.131
		(3.33)
Constant	-0.377**	-2.977**
	(-23.61)	(-10.73)
Observations	7557	7497
R-squared	0.001	0.455

Notes: *Not in IFLS4* is a dummy indicating a child with observations in earlier IFLS surveys which is missing in IFLS4. *z-score* is the difference between actual height and the height of a same age and gender reference population, measured in standard deviations. *Logpce* is the log of monthly per capita household expenditures. *Urban* is a dummy of value 1 if the household resides in an urban area. **p<0.05, *p<0.10. t-stats in parentheses.

Table 2.5: The effects of childhood z-score on cognitive score using restricted sample

	(1)	**(2)**
	OLS	IV
LHS variable	Cog. score	Cog. score
Sample	**Restricted**	**Restricted**
z-score	0.091**	0.384*
	(2.80)	(1.68)
Age	0.719*	0.849**
	(1.83)	(2.05)
Age squared	-0.019	-0.024
	(-0.95)	(-1.14)
Male	0.357**	0.344**
	(3.14)	(2.97)
Logpce	0.184**	0.160**
	(3.17)	(2.54)
Urban	0.465**	0.388**
	(3.85)	(2.78)
Mother's education		
Junior high school	0.341**	0.354**
	(2.24)	(2.29)
Senior high school	0.688**	0.621**
	(4.45)	(3.71)
College/University	1.121**	1.064**
	(4.46)	(4.28)
Instruments		Crisis
Constant	0.646	0.519
	(0.32)	(0.25)
Observations	1511	1511
Adj. R-squared	0.122	

Notes: The first stage and reduced form corresponding to the results presented in *Column (2)* are shown in *Column (1)* of Table 2.2 and 2.3. *IV*=Instrumental Variables. *Cog. score* is score from the IFLS cognitive test for children aged 7-14, maximum score is 12. The *restricted sample* includes *non-crisis* children who had their height measured within the same age range as *crisis* children. The *full sample* consists of all children who completed cognitive tests in IFLS4 and had their height measured prior to the age of 5 years. *z-score* is the difference between actual height and the height of a same age and gender reference population, measured in standard deviations. *Crisis* children were born 1998-2000. *Logpce* is the log of monthly per capita household expenditures. *Urban* is a dummy of value 1 if the household resides in an urban area. Elementary school is the excluded educational category. **p<0.05, *p<0.10. t-stats in parentheses.

Table 2.6: The effects of childhood z-score on cognitive score using household fixed effects with restricted sample

	(1)	**(2)**	**(3)**
	First stage	Reduced form	IV-HHFE
LHS variable	Early childh. z-score	Cog. score	Cog. score
z-score			0.493*
			(1.81)
Crisis	-1.602**	-1.059**	
	(-3.07)	(-2.02)	
Age	-0.400	-0.565	-0.300
	(-0.47)	(-0.67)	(-0.43)
Age squared	-0.009	0.030	0.036
	(-0.24)	(0.71)	(0.99)
Male	-0.059	0.212	0.251
	(-0.24)	(0.89)	(1.22)
Instrument			Crisis
Constant	3.997	12.269**	10.897**
	(0.91)	(2.64)	(3.21)
F-test of instrument	F(1,159)=9.42		
Observations	315	315	315
Adj. R-squared	0.592	0.546	0.666

Notes: *IV*=Instrumental Variables, *HHFE*=Household Fixed Effects. *Cog. score* is score from the IFLS cognitive test for children aged 7-14, maximum score is 12. *z-score* is the difference between actual height and the height of a same age and gender reference population, measured in standard deviations. *Crisis* children were born 1998-2000. *Logpce* is the log of monthly per capita household expenditures. *Urban* is a dummy of value 1 if the household resides in an urban area. Elementary school is the excluded educational category. **p<0.05, *p<0.10. t-stats in parentheses.

3 Health effects of on the job promotions: A Whitehall II study

3.1 Introduction

Richard G. Wilkinson's paper 'Income distribution and life expectancy' (British Medical Journal) from 1992, sparked an increased interest in the relationship between income distribution and a number of social indicators. Wilkinson and others have shown that there is little relationship between average income and health among rich countries. However, within each and every country they find a strong health gradient in income. This gradient works its way from the bottom of the income distribution and all the way to the top. As Pritchett and Summers suggested, it seems like wealthier is indeed healthier. But there seems to be a disconnect between the wealthier is healthier notion between countries and within countries. An American CEO is on average healthier than an American shop assistant, but an American is not much healthier on average than a Greek even though the PPP adjusted GDP per capita in the U.S is more than twice as large as that in Greece (Wilkinson, 2005).

These observations raise the possibility of reasons other than financial resources for the gap in health between relatively richer and poorer people in the richest 25-30 developed countries. Differences in health behaviours has been launched as a possible explanation for the health advantage of the rich. Lower paid workers do indulge in more unhealthy behaviours like smoking and bad eating on average, but this effect is not large enough to explain the entire health gap. Smith et al (1990) found using the Whitehall II data that these behavioural risk factors could only explain about one-third of the increased risk of deaths caused by heart disease.

Wilkinson has written several books[21] that deal with the idea that it is a person's subjective social status which explains the difference in health and other social indicators like crime and teenage pregnancies, between people at different positions in the income distribution. People care about how others view them and their status in a community is important for self-worth. Wilkinson thinks of income as

[21] 'The Impact of Inequality: How to Make Sick Societies Healthier', 'Social Determinants of Health' and 'The Spirit Level: Why More Equal Societies Almost Always Do Better'.

a signal of a person's social status. Social status is potentially closely connected to the occupation of a person and higher paying occupations are more valued in a social context. The central idea is that people in lower paying occupations will feel less valued and compare themselves to people higher up the social ladder. It is somewhat problematic to determine the suitable group of comparison. Do people compare themselves to their neighbourhood, people in their city, their state or the country? Evidence is mixed (Wilkinson, 2005) and it seems like the smaller you make the area of study, the harder it is to document a health gradient in income. However, the health gradient has been documented for all the U.S. States, the U.S. as a whole and for a number of other countries (Wilkinson, 2005).

In this paper we attempt to test whether a job promotion is associated with physical or mental health. We use data from the Whitehall II study of civil servants working in London, UK. These people work in an environment where the grade system is uniform and thus can be evaluated by people around them. If job grade signals status, we could expect that working in a low grade job leads to lower social status, more stress and worse health. The clear health gradient in job grade has been found repeatedly in this data[22] and it was this discovery from the original Whitehall study that prompted the Whitehall II follow-up. All the subjects in these studies are in employment and should have access to adequate financial resources to enjoy a healthy diet and good health care. The initial finding of a clear health gradient in job grade in this population was therefore somewhat surprising at the time of Whitehall I.

If it is the case that civil servants in London evaluate their standing in the community based on their job grade and this could cause stress and worse health, we should expect that a job promotion would at least temporarily promote better health. How long this effect lasts, depends on whether people have a tendency to always compare themselves to others higher up the ladder. In the extreme, this would imply that we all would suffer some degree of stress as long as we are not at the top. It seems likely that someone working as a shop assistant would at least for a while, enjoy a feeling of progress and improvement in their relative social status when promoted to shop manager.

A promotion is associated with higher salary. It will be difficult to completely isolate the effects of the pay increase and the changes in stress levels resulting from

[22] See ex. Marmot et al. (1999)

a job promotion. Increased financial resources could buy you better health, but it seems implausible that a modest increase in salary for this group would lead to significant changes in health status. These workers already had a job with a reasonable salary before the promotion and health care in the UK is largely free. The Whitehall II study has not collected income data in all phases and in the data we have been given access to, such data has only been recorded in one of the six phases. Since we will exploit the time series nature of the data, we cannot control directly for the potential health effects of increased pay.

There are to our knowledge, two other papers which attempt to analyse the health effects of job promotions. These are different in methodology, use different data and reach opposite conclusions. Anderson and Marmot (2007) use data from the Whitehall II study. They employ an instrumental variables approach to address the causality issues related to job promotions and health. It is possible that healthier people have a greater likelihood of being promoted than less healthy workers. They use the different departmental promotion rates in Phase 5 as instruments for promotions. These promotion rates are determined by cohort sizes and levels of wastage in each department. Anderson and Marmot (2007) test whether the probability of heart disease is caused by job promotions for workers in the civil service in London. The results suggest that the probability of coronary heart disease is reduced by 6-18 % by a job promotion in the Phase 5 cross-section. These results are qualitatively similar to the results presented here.

Boyce & Oswald (2008) employs a difference in difference approach using data from the British Household Panel Survey. This survey includes workers from a number of different industries and has a standardized rating of employment grade. They use a treatment group of people who underwent a job promotion and a control group of workers who did not. The results show that promoted workers did not experience better physical or mental health post promotion compared to the control group. Their results suggest that the treatment group was healthier prior to the promotion and that promoted workers experienced worse mental health after the promotion.

We are facing the danger of reverse causality when analysing the impact of job promotions on health. It is not unlikely that healthier workers are more likely to be promoted. This is also the result of Boyce & Oswald (2008). Anderson & Marmot (2007) attempts to deal with this issue by employing instrumental variables. We

will use a fixed effects (FE) approach to alleviate the reverse causality issues. Although it is likely that healthier workers are more frequently promoted, it seems less likely that a given worker has a higher probability of promotion in a period of especially good health. This should be even truer for generally healthy workers. We might think of examples with workers suffering from a debilitating disease and then recovering as being more likely to be promoted post recovery. However, for workers who are of average to good health throughout their career, this is unlikely to be an issue.

By using fixed effects, we effectively follow individual civil servants and observe whether their health responds to job promotions. As we alluded to earlier, it seems plausible that workers who have been promoted, but subsequently work for a long time without further promotions, might increasingly suffer from stress as time goes by. Our results show that there seems to be a substantial Self-Assessed Health (SAH) benefit from a job promotion. We found some evidence that suggests that this benefit decreases with the length of tenure until the promotion was gained. There was little evidence of a mental health gain from job promotions.

Our paper analyses a specific situation which is likely to change the social status of a person. Although a promotion is not an exogenous shock to the social status of the recipient, we employ an econometric strategy that will alleviate the concerns regarding endogeneity. There are very few papers which deal with the question posed here and none to our knowledge that uses time series data coupled with a fixed effects approach. We also find that fixed effects tend to lower the traditional health gradient estimations which are widely use and sited other where.

We will start by presenting the theory formulated by Richard G. Wilkinson in the next section. The section condenses a lot of the medical literature which makes it plausible to suggest a promotion-health link. Section 3.3 presents the Whitehall II data and our estimation strategy. The following section presents the results and discusses potential alternative explanations for the observed results. The final section concludes.

3.2 Stress and health implications

Wilkinson proposes a chain of mental and physiological responses that run from social status to stress and the health implications of prolonged stress. The first

link in the chain, hypothesises that low social status will lead to mental stress. This is also the part of the argument which is harder to document. We cannot experimentally change a person's social status and record the health consequences for obvious reasons. These experiments have been done with primates and could potentially teach us something about the stress responses of humans. Shively and Clarkson (1994) is a frequently quoted experiment of this type. They studied captive macaques that were divided into separate groups where they naturally formed a social hierarchy. Social status was experimentally changed by moving animals that had attained high- or low status into groups where their social status would change. As a result some subordinate animals became dominant and vice versa. All animals were fed the same diet to eliminate stress from material resources. They found a relationship between low social status and central obesity, insulin resistance and a worsening ratio of high- to low density blood fats.

Sapolsky (1993) found among baboons in the wild that lower ranking monkeys had a higher level of cortisol than higher ranking animals. Cortisol is an important stress related hormone. When under severe stress, humans also produce more cortisol as a physiological attempt to retain physiological equilibrium in the face of external stress. We should be careful to infer too much from these results, but since such experiments are infeasible on humans, we do not have any direct ways of investigating this relationship.

The chronic activation of the stress system has been found to have several negative health implications. McEwen (1998) found that the response of the nervous system increased the risk of diabetes, hypertension, cancer and cardiovascular disease. Dickerson & Kemeny (2004) writes in their synthesis of research on stress and cortisol response that 'we would predict that individuals who possess uncontrollable characteristics that result in social rejection (e.g. those who are stigmatized) could persistently experience uncontrollable social-evaluative threat; this could lead to chronic activation of this physiological system with its possible health implications.' By social-evaluative threat, they refer to situations where humans posses characteristics or are asked to perform tasks that can be judged by people around them. Such situations are found to particularly increase stress. Vitetta et al (2005) state that '...the most important factor in why a person becomes ill lays in the brain. Stress and pleasure play a critical role in wellness and disease, with stress contributing significantly to the risk of disease.'

There is an abundance of evidence which shows that stress will lead to poorer health in humans. Whether the social position of someone could create stress, is only possible to test indirectly. We will attempt to show that the theory of Wilkinson is a reasonable explanation of the results presented.

3.3 Data and estimation strategy

3.3.1 The sample

All the data used is from the Whitehall II[23] study. The study commenced in 1985 when 10 308 non-industrial civil servants between the ages of 35 and 55 years where interviewed and completed medical examinations. This group has been followed over time and at the time of writing, phase 9 has been collected. We have been given access to the initial six phases. Three of these phases included medical examinations, while the respondents completed questionnaires only in the three remaining.

Due to privacy issues, we have not had access to all the variables at the most detailed level. Particular note should be given to the grade variable. The civil service grade structure is standardized and the Whitehall researchers have coded these into six grade levels. From higher to lower rank, these are Unified Grades 1-6 (Administrative), Unified Grade 7 (Administrative), Senior Executive Officer, Higher Executive Officer, Executive Officer and Clerical/Support staff. We have been given access to a broader breakdown of only three levels. Our grade variable combines the Unified Grade levels into grade 3, the three executive officer levels into grade 2, and clerical/support staff is level 1. This is likely to create measurement error in our promotion variable. However, the more expansive grade variable in Whitehall II is also condensed compared to the number of actual grades in the civil service. There are as far as we can see no feasible method of correcting for this problem. We will proceed bearing this in mind and the fact that the potential measurement error could tend to bias our results toward zero.

[23] The Whitehall II study has been supported by grants from the Medical Research Council, the Economic and Social Research Council, the British Heart Foundation, the Health and Safety Executive, the Department of Health, the National Heart Lung and Blood Institute/National Institutes of Health (NIH), the National Institute on Ageing/NIH, and the John D. and Catherine T. MacArthur Foundation Research Networks on Successful Midlife Development and Socio-Economic and Health and Agency for Health Care Policy Research Grant HS06516

We took a number of steps to create our final baseline sample. First, we excluded everyone who joined the civil service prior to 1970 to make our sample somewhat more homogeneous with respect to general health trends. The 1970 cut-off is somewhat arbitrary, but was chosen to keep the starting time frame relatively tight without discarding too many observations. This step removed 5 168 of the original 10 308 respondents. Second, we discarded all respondents who joined the civil service at rank 3 (recoded from the original Whitehall rank 1) or entered the sample at this rank. This is the highest rank possible and would therefore make these respondents ineligible of receiving a promotion. The second step lost us an additional 1 238 from our starting sample. Third, we removed everyone who was observed for one period only as we wish to analyse how a promotion between the previous- and current period affects current health. This step removed another 545 respondents. In the final step, we excluded everyone who was not in the sample in the first phase, but who entered later. All these steps, lead to a baseline sample of 2866 respondents and 11978 data points.

We observe a total of 702 promotions in our sample. Phase 1 have no promotions by construction as we do not observe our respondents prior to this phase and thus cannot know whether they have been promoted just prior to this phase. Some respondents (55) moved out of the survey group and back in at a later stage. This reflects the fact that they for a period of time were not employed by the civil service.

Table 3.1 summarises the most pertinent information of our baseline sample. The number of observations decreases as we move from the earlier phases to the latter ones. Table 3.1 indicates that a majority of the promotions happen earlier in a respondents run. Overall this tendency is not extreme, but it emphasises the need to control for age not only because health usually decreases with age, but also because a person is somewhat more likely to receive a promotion earlier in their career.

Our baseline sample is predominantly male and white. The gender composition is fairly even as 47.6% of our sample is female. The average SAH score is 3.64 which correspond most closely to answering 'good' to the question of how the respondents health has been the last year. The average General Health Questionnaire score is 3.95 which is quite high considering a person with a score of five and higher is considered a case. A random respondent in a random phase is labelled a case 28%

of the time. This cut-off point is standard in the literature, but it seems unlikely such a high number of respondents in this sample have a serious psychological issue.

3.3.2 Dependent variables

The Whitehall II study contains few clinical health variables. Blood pressure is an exception, but was only measured in phases one, three and five. We will not attempt to analyse the impact of promotions on blood pressure. A person's blood pressure is volatile and it takes multiple measurements to achieve a credible measurement. The Whitehall medical screenings made one measurement without knowledge of what the person had done just prior to the measurement. In the absence of good clinical health measurements, we will rely on a subjective health assessment score and a psychological health assessment score as our main dependent variables. The former score is known as Self-Assessed Health (SAH) and is widely used in the literature while the latter is the General Health Questionnaire score. The respondent is asked to evaluate their physical health on a scale of 1-5, 1 being very good and 5 being very poor. In our analysis we have re-coded this variable to make the higher value refer to better health.

Self-Assessed Heath Self-Assessed Health is a subjective evaluation of one's health and might not reflect actual health. We have limited options when it comes to correlating this variable with objective health measures. Other researchers have investigated this correlation.

Johnston et al. (2009) compared the difference in self-reported hypertension and clinical measures of the same medical condition using data from Health Survey for England (HSE). Each respondent was asked to list any long-standing illnesses and was within a two week period also repeatedly tested for high blood pressure by a qualified nurse. They document a large, systematic under reporting of hypertension among English adults. 7% of the sample reported hypertension while almost 35% had clinical hypertension which is defined as either systolic blood pressure of more than 139 or diastolic blood pressure in excess of 89. The fraction of respondents who reported a false negative was 87%. Moderate and even high levels of hypertension are often asymptomatic. At very high levels of hypertension, it is common to experience symptoms like headache, dizziness, nausea, shortness of

breath, heart palpitations and fatigue. It is therefore possible that even though a person is unaware that he has hypertension, this condition would lead to lower general health scores due to the symptoms describes. Johnston et al. (2009) finds no relationship between undetected hypertension and general health scores. However, they do not differentiate between modest and high levels of hypertension with the latter group being much more likely to experience these symptoms. It seems likely that hypertensive people experiencing these symptoms would also rate their general health as worse than people who do not.

Bentsen et al. (2008) assessed 100 patients with chronic obstructive pulmonary disease, more commonly known as chronic bronchitis and emphysema. This is a condition where the airways become narrower leading to shortness of breath. They found that patients with reduced lung capacity or with more breathlessness reported lower Self-Assessed Health. This result is not surprising, but shows that SAH does reflect illness symptoms.

SAH has been associated with mortality in a number of studies. Mossey & Shapiro (1982) even found that SAH was a better predictor of seven year survival than medical records for elderly Canadians. Idler & Benyamini (1997) reviewed 27 American studies on the relationship between SAH and mortality. These studies universally report an increased risk of dying when the subject report lower Self-Assessed Health even after accounting for known health risks. They conclude their article in the following way: 'We would argue that the global rating represents an irreplaceable dimension to health status and in fact that an individual's health status cannot be assessed without it.' Huisman et al. (2007) investigated whether the predictive ability of SAH on mortality varied by education. They found a somewhat higher predictive power of SAH for men with tertiary education compared to the lowest educational group. This effect was not large and no such effect was found for women. It seems like people with different levels of education emphasise somewhat disparate elements when evaluating their own health, but they all do a good job at predicting mortality. A Whitehall II study, Sing-Manoux et al. (2007), found that SAH predicts mortality and that men and women predict equally well.

Blood pressure was measured in phases 1, 3 and 5 and is the only clinical measure of health we have access to in our data set. The crude correlation coefficient between a marker for high blood pressure (systolic blood pressure>139) and SAH was negative and highly statistically significant. This could reflect an age effect

where blood pressure rises and SAH decreases by age. We therefore ran a fixed effects regression with SAH as the dependent variable and high systolic blood pressure as the independent variable while controlling for a quadratic in age. The results show that the coefficient on blood pressure is negative and statistically significant at the 5% level. This analysis was performed using the whole Whitehall II sample.

General Health Questionnaire Score D.P. Goldberg introduced the General Health Questionnaire (GHQ) in 1978 and it has since been translated into almost 40 languages. It now exists in several incarnations varying by the number of questions asked. Whitehall II uses the 30 item version, while 12-, 28- and 60 items versions are also commonly used. The questionnaire is used as an assessment of psychological well-being. It consists of questions like 'Have you recently lost much sleep over worrying', 'Have you recently felt on the whole you were doing things well' and 'Have you recently been feeling unhappy and depressed'. These questions are all answered using a four point scale ranging from not at all to much more than usual. Whitehall II attaches a score of 1 or 0 to each of the questions. A score of one reflects the answer of one of the two alternatives which would indicate psychological distress. The GHQ score is a summation of the individual question scores and ranges from 0-30. A person with a GHQ score of five and above is considered a case. The GHQ has been shown to have a high degree of validity in a number of environments[24].

By using both SAH and GHQ score, we will attempt to analyse to what extent a job promotion will impact physical health through SAH and mental health through GHQ score. It is likely that these scores will overlap to a certain degree since ex. being depressed might make a person score himself worse on SAH as well as GHQ, but by construction, these two items are designed to measure different aspects of a person's health.

[24] See Aderibigbe & Gureje (1992) for Nigerian women after childbirth, Andersen et al. (2002) in Danish prisons and Romans-Clarkson et al. (1989) for female New Zealanders.

3.3.3 Estimation strategy

The major challenge in estimating the effect of job promotions on health is one of reverse causality or simultaneity. It is possible and even likely that healthier workers are promoted more often than similar, but less healthy colleagues. It makes sense for a firm to promote and thereby give more responsibility to individuals who will miss fewer days of work due to illness and could potentially pass away later. We have chosen a fixed effects approach partially to account for the problem of simultaneity. In the absence of good instrumental variables, we believe that using fixed effects will go some way to alleviate the concerns described above. By employing a fixed effects approach, we follow individuals instead of making cross-sectional comparisons. For reverse causality to be an issue in this context, we would have to believe than an individual worker is more likely to be promoted in a period of good health than in a period of less good health. This could be the case if a worker had a serious illness that he recovered from, but it seems much less plausible for workers with relatively small variation in health. Of our sample of 2867 respondents, approximately 74% had a variation of one point or less in the health variable throughout the sample period. Equation (3.1) below will be our default regression equation.

$$h_{it} = \alpha_1 promotion_{it} + \beta_1 X_{it} + \alpha_i + \epsilon_{it} \quad (3.1)$$

h_{it} denotes health for person i in period t. Our health variables are SAH and GHQ score. Our promotion variable is a dummy with value 1 if the respondent has been promoted between the previous period and the current. X_{it} is a matrix of time varying individual characteristics. Among these will be age. The fixed effects will be captured in α_i. This coefficient will encompass all individual time invariant characteristics. In this context, unobservable ability is likely to be an important asset when a promotion decision is made. OLS without fixed effects would not account for ability and could attribute positive health effects of a promotion when both these variables could be caused by an omitted variable like ability, ambitions or IQ. These unobserved characteristics are likely to be fixed over the time period we follow these civil servants and can therefore be controlled for directly by fixed effects. ϵ_{it} is a random, individual error.

To achieve unbiased results, two things need to be true. First, an individual civil servant is not more likely to get a promotion in a period of better health

compared to a period of worse health. This might seem like a somewhat optimistic assumption, but we need to keep in mind that most of the respondents in our sample are healthy individuals and whether their health is average or good seems unlikely to determine the timing of a promotion. Second, we assume that SAH and GHQ measures physical and mental health in a meaningful way. We have outlined earlier why we think that is a reasonable assumption.

3.4 Results

Table 3.2 summarises the results of estimating Equation (3.1) over different samples. Panel A uses SAH as the dependent variable, while the second panel estimates the effects of job promotions on GHQ score. All results come through using fixed effects estimation.

Column (1) in Panel A and B are the fixed effects equivalent of the standard gradient regression for SAH and GHQ score respectively. The Whitehall II project was spawned as a result of finding a clear health-grade gradient among civil servants in London in 1967. Their initial finding was that civil servants in lower grade jobs were more likely to die prematurely than their colleagues in higher ranking grades. Whitehall II was set up to explore this association and the health gradient was still a feature of the data 20-40 years later. Column (1) shows a clear health gradient in SAH. Higher ranking respondents are more likely to score higher on the SAH scale than civil servants in lower grade jobs on average. Our FE coefficient on grade is smaller than the traditional cross-section analysis done elsewhere. The cross-section coefficients of the six phases varies between 0.07 (phase 5) and 0.23 (phase 1) with an observationally weighted average of 0.18. There seems to be a general trend in the data that the SAH gradient is decreasing over time. The difference between FE estimation and cross-sectional OLS is likely a result of healthier people reaching higher positions. Whether this is a pure health effect or whether education, IQ or other unobservable variables drive this association, we cannot tell. Panel B shows no association between grade level and GHQ score independent of the sample used. With the exception of phase 1, we find no cross-sectional association between these two variables. In phase 1, we find a positive correlation which suggests that higher ranking officials are more likely to have mental issues on average which aligns with the findings of Boyce & Oswald (2008).

In Column (2), we introduce the promotion variable. Promotion is a dummy variable with value 1 if the respondent has moved up a grade level between the previous phase and the current. Column (2) shows a positive relationship between a job promotion and SAH. The relationship is statistically significant at the 5% level. A recent promotion increases the SAH score by almost 0.1 points.

Columns (3) run the identical analysis as in the previous two columns with a different sample. For this analysis, we have included only respondents who were in the sample for all six phases. The larger number of phases on average should increase the consistency of our fixed effects estimator. Our coefficient in Column (3) is almost identical to the one we found in Column (2). We find no significant association between any of the dependent variables and GHQ score.

Column (4) concludes our initial analyses. Here we use a much more homogeneous sample. In this sample we only allow respondents who joined the civil service between 1980 and 1985. We also only include people who started at the lowest job grade, grade 1 (grade 3 as commonly referred to in Whitehall II literature). These respondents are seemingly very similar when they enter the Whitehall II sample with respect to time of entrance and the qualifications they had when entering. They all started out at the bottom and we can avoid potential heterogeneous effects due to starting grade. Column (4) shows that the coefficient on promotions almost doubles compared to Column (2) and (3). The estimate is less precise due to the decreased sample size and is only statistically significant at the 10% level.

The overall picture from Table 2 tells a story of a significant physical health effect of a job promotion, but no mental health benefits. Ferrie et al. (2003) found a significant difference in GHQ scores between civil servants in the lowest grade versus in the highest grade, but they only studied the phase 5 cross-section. It could well be the case that this effect is more pronounced at the extremes and that it is more of a cross-sectional phenomenon than an individual effect.

We have argued that running FE estimation should limit the influence of reverse causality. In a cross-section, it is probable that workers of better health are more likely to receive a promotion than a similar co-worker of worse health. We ran a probit regression with promotion as the independent variable and lagged health as one of the exogenous variables. We found that the coefficient on lagged health was small, negative and statistically insignificant. It seems like a given worker is no more likely to be promoted after a period of good health.

Our results suggest that a promotion is likely to increase SAH in the next phase. Unfortunately we do not have access to good, direct measures of physical health. Blood pressure is the only candidate, but we have argued that the blood pressure measure included in the Whitehall II data set is not reliable for the purpose of this paper. It is also a variable which is often included as a binary variable to indicate the presence, or the lack of presence, of hypertension. As health is not related to blood pressure in a monotonic way, a hypertension variable is often the better way of operationalizing blood pressure. Running these regression with hypertension as the dependent variable gave insignificant results indicating that the probability of hypertension is unrelated to a promotion. However, the hypertension variable is a very crude measure of health and it is possible that a promotion does not give large enough improvements in health to change the likelihood of hypertension.

A criticism of our results is the lack of results using good, clinical measures of health. We lack such measures in the data set. The only one we have access to is quite crude and has not been taken at all phases. We have described studies which show a clear association between SAH and physical health measures. However, as SAH is self-reported such a measure could pick up changes at the individual level which are not related to physical health. Van Tilburg et al. (1999) found that people who are homesick reported worse mood and worse SAH than others. There is a substantial literature on the relationship between affect and SAH[25]. Positive affect has been found to be a predictor of SAH. A promotion could lead to increased positive affectivity or the reversal of negative affectivity and as such lead to higher reported SAH through this channel and not through an increase an actual physical health. We cannot dismiss that this channel is causing our results. However, the non-results of GHQ helps alleviate such concerns to some degree. If a promotion leads to better mood or improved affectivity, we would expect a promotion to be associated with GHQ. That is not the case in our sample. For mood, affectivity or other potential psychological channels to cause our results, it would need to affect SAH and not GHQ which seems less likely.

3.4.1 Tenure effects

A promotion may differentially affect people's health depending on when it is

[25] see Pettit et a. (2001) for a summary

achieved. As the empirical literature has little to say on this issue, we are unable to make clear predictions. The net present value of a promotion is higher the earlier a promotion is achieved. Whether the financial benefits of a promotion will affect health is unclear and a question we address below. As tenure increases so does age and it is possible that the importance of changes in social status (grade) would have different effects on an individual depending on the person's age or length of tenure. We have found no analysis of this question in the literature. We will attempt to address this issue by estimating Equation (3.2) below.

$$(3.2)\ h_{it} = \alpha_2 promotion_{it} + \beta_2 X_{it} + z_1 tenure_{it} + z_2 promotion_{it} * tenure_{it} + \alpha_i + \epsilon_{it}$$

Equation (3.2) is identical to Equation (3.1) except the tenure terms. z_2 captures to what extent the length of tenure determines the health effect of achieving a promotion. Table 3.3 summarises the results.

Our baseline sample only includes 57 respondents who have been promoted twice in our sample which is the maximum number of promotions due to our three point grade variable. This means that for the majority of our sample, promotion*tenure is also the length of time it took to gain the first and only promotion. It seems likely that stress would increase with the time a person is stuck at the same grade level. Whether the length of non-promotion should be positively correlated with the health effect when the promotion occurs, is unknown in the literature. The tenure variable is expressed in number of years since the person entered the civil service. A large value on the promotion*tenure variable indicates both that the respondent was older when the promotion occurred and in most cases, that it took a long time to gain the promotion.

Column (1) in Table 3.3, estimates Equation (3.2) for the baseline sample and finds a negative coefficient on the main variable of interest; promotion*tenure. The coefficient is statistically significant at 10%. This result indicates that the longer a person waits for a promotion the smaller is the health effect of a promotion. The same analysis with a balanced sample, reaches the same conclusion with the same level of significance. When we employ the restricted sample in Column (3), the results turn insignificant and the sign changes. Using this sample leads to a much smaller sample size and larger standard errors.

Table 3.3 seems to indicate that waiting a long time for a promotion, means a smaller health effect of the promotion. However, this effect could also be an

age effect. The increase in social status might be less beneficial from a health standpoint as age increases. Tenure and age will be highly correlated for our sample and running the same analysis with promotion*age, gave similar, but somewhat weaker results.

Table 3.4 shows the results of estimating (3.2) using a different tenure variable. This tenure variable expresses the number of years in the current grade. Since we have no knowledge of promotions occurring prior to data collection of respondents who joined the Civil Service at a different grade than the one they had when entering the Whitehall II sample, we have excluded all respondents who received a promotion prior to Phase 1. While the tenure variable used in Table 3.3 expresses whether a promotion happens early or late in a respondents career, the latter tenure variable measures the period of time a person has stayed at the same grade level. The coefficient on the interaction term is larger in absolute value than what we found in Table 3.3, and is statistically significant at 5% in Column (1) and (2). In Column (3), the coefficient is small, changes sign and becomes insignificant. The results in Column (1) indicate that waiting an extra 10 years for a promotion, more than halves the positive health effect of a promotion. We can only speculate to the reasons for observing this result. It could be that the negative stress response of no social upwards mobility over time makes it increasingly difficult to recover when a promotion is gained.

3.4.2 Income effect

We have throughout shown a significant and economically meaningful association between job promotions and health. Thus far we have attributed the association to psychological processes related to the individual's place in the relevant social hierarchy. For our sample, we have assumed that at least part of the relevant reference group is their co-workers. It is possible that all we are observing is the effect of increased purchasing power which a promotion is likely to bring.

The income-health gradient has been well documented in a number of settings. Very few would dispute that there is a strong relationship between income and health which runs all the way from the bottom of the income distribution to the very top. The most naive way of interpreting the association between health and income, would be to attribute a large portion of the association to a direct income effect. With more money comes the opportunity to buy better health. This might

certainly be the case in poor countries as shown by Case (2001) in her paper on the health effects of changes in the pension system in South-Africa. She showed that households who pooled their resources and had one or more recipients of the new and improved pension scheme, had health benefits from the additional income which was shared by all. However, the evidence is very mixed and weak when it comes to rich, developed countries. As we have noted throughout, there are potentially serious concerns of simultaneity when estimating the effect of income on health. A few recent papers have made some serious attempts at tackling these issues. Frijters et al. (2005) investigates whether there was an improvement in Self-Assessed Health of East Germans after the reunification of 1990. They view the reunification as a natural experiment where East Germans experienced rapidly increasing real household incomes while the former West Germans did not. Frijters et al. employ a fixed effects technique and find a positive, but very small health effect of increased income. They were able to attribute approximately 5% of the changes in SAH over the period to changes in income. Adda et al. (2008) construct an individual dynamic model where they test whether permanent income shocks affect a number of health measures. They use cohort wide income shocks related to changes in the macro economic environment which can plausibly be regarded as exogenous to cohort health. They find that negative shocks increase mortality, but have little or no effect on SAH, blood pressure, cardiovascular condition, respiratory condition or mental health.

At this stage is seems very uncertain whether there is a causal link from income to health in developed countries. There are a number of other potential explanations for the health-income relationship. It could manifest itself through omitted variables which determine both health and income. Such variables could be genes, IQ and ability. By using FE we effectively filter out these variables as long as they are time invariant. Another potential explanation could be education. Better educated people could be better equipped at gathering, understanding and efficiently acting on new health information. With our FE approach, this is not the cause of the association between promotions and health and has not been found to be the 'smoking gun' behind the health-income gradient even though education tends to weaken the relationship somewhat[26]. In our sample, we include only adults with a steady occupation and salary who live in England where health coverage

[26] See ex. Deaton & Paxson (1999)

is good and largely free. They are presumably all in a position where they can acquire sufficient nutrients and eat a healthy diet if so inclined. It seems unlikely that the increased salary a civil servant receives when promoted could possibly explain much or any of the associated health gains. Unfortunately, income data was only collected in Phase 5 of our panel and we cannot test this explicitly by including such data in our analysis. We have however run the analysis on the entire Whitehall II sample in Phase 5 to test how income is associated with health in this particular sample. We do, however, have to keep in mind the reverse causality- and omitted variables issues which we cannot control for in this cross section. The results are shown in Table 3.5.

We have included two new covariates, male and white, which are both dummies with value one if the variable labels are true and zero otherwise. These variables are generally positive and significant for our sample. The age variables are statistically insignificant and reflect the age distribution at Phase 5 which is quite narrow with a mean age around 56 years. The income variables are based on the respondent's own income. Income information in Whitehall II was divided into four groups and we have excluded the lowest. In Column (1) we run the analysis without the grade variable and the income variables are highly significant. We see that civil servants earning higher salaries rate their health higher than the lower earners. In Column (2) we include the grade variable and with the exception of the highest earners, the income variables lose all significance. This is also the case in Column (3) where we include a variable that measures the number of years of education. The variable itself is negative and statistically significant at 10%. This is the opposite of what we would expect and what others have reported. We prefaced this by pointing out the weaknesses of the estimation strategy employed here and the resulting bias might explain this counter-intuitive result. The main impression to take away from Table 3.5, is that it seems like the grade variable is globally positive and significant no matter the specification and once we include this variable, the income variables lose significance. The grade variable we include in most of our regressions seems to pick up most of the variation in personal income.

We also ran a regression where we substituted the promotion variable for the lagged equivalent. We found no statistical association between a promotion which occurred between the two periods prior to the current, and current SAH. It seems like if income explains a large part of the relationship between a promotion and

health, this effect would not be as transitory as we find here.

3.5 Conclusion

The main idea of Richard G. Wilkinson's work on health and social status is that people on the lower rungs of society are getting stressed and ill by their social position. He, and others, documented the health-income gradient within a number of countries and the lack of a similar gradient between relatively rich developed countries. We have attempted to investigate whether a change in grade level for civil servants in London is associated with better health. The grade level is likely to be an important input in these government officials's view of their social status. We found that a job promotion increases Self Assessed Health score by 0.1-0.3 points. This effect is quite large and statistically significant. We also found some evidence indicating that a promotion late in their tenure will have a smaller impact on the health of a respondent than one gained early and that the effect of a promotion is smaller the longer a person has worked without a promotion. There were no significant results when using GHQ score as our health variable.

A job promotion is likely to represent an improvement of the social status of the respondent. The job promotion usually also increases salary. Unfortunately we have not had access to salary figures over the sample period to analyse how much of the health effect is caused by increased income. However, the people in our sample all have a job with steady, relatively well paying jobs. A job promotion for this group is unlikely to bring a substantial improvement in the finances of the recipient. If the health effects of a promotion were largely an increased purchasing power effect, we would not expect the effect to be as transitory as it seems to be in our data. We found no health effect of a promotion gained between the two previous periods, but found a large effect of a promotion between the previous and current periods.

We have argued throughout that the FE approach we use here should help alleviate concerns of potential simultaneity. By employing a FE approach, we difference out some of the time invariant unobservables which might drive both health and the potential for promotions. IQ could be such a variable. We follow largely healthy individuals over time and argue that it is unlikely that an individual will be promoted according to their health at any one point in time.

It is possible that people who experience periods of better health also feel more

confident and as a result are more likely to get promoted. There are some features of the data and results which makes this explanation somewhat less likely. We observe relative little variation in SAH for given respondents over the sample period. Most of our sample is healthy making it less likely that a period of better than normal health would increase confidence to a great extent. We also find that respondents who achieved a promotion were not healthier on average in the prior phase. None of this is conclusive and we certainly cannot dismiss confidence as the link between promotion and health entirely.

The results presented here falls in line with the theories and results of Richard G. Wilkinson. We have only demonstrated the final step in the chain of arguments needed to form these theories. We have shown that a promotion is significantly associated with better health in the subsequent phase of the Whitehall II sample. Whether the grade level is important in the formation of civil servants view of their social status, is unknown. We do not know whether our respondents actually get stressed by a lower grade level. These questions are unresolved and could potentially create more research.

Table 3.1: An overview of the data and variables of interest

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
	Total	Phase 1	Phase 2	Phase 3	Phase 4	Phase 5	Phase 6
Observations	11978	2866	2866	2499	1706	1187	851
Promotions	702	N/A	299	159	82	72	90
Observations/Promotions	0.06	N/A	0.10	0.06	0.05	0.06	0.11
Survey years		1985-1988	1989-1990	1991-1994	1995-1996	1997-1999	2001
	Average						
Male	52.4%						
White	84.1%						
SAH	3.64						
GHQ score	3.95						
GHQ 'cases'	28.0%						

Notes: *SAH*=Self-Assessed Health, 1(poor) - 5(excellent). *GHQ*=General Health Questionnaire, Whitehall II uses the 30 item version and a score of 5 and higher is considered a 'case'.

Table 3.2: The health effect of promotions

Panel A: Results for Self-Assessed Health

	(1)	**(2)**	**(3)**	**(4)**
LHS variable	SAH	SAH	SAH	SAH
Sample	Baseline	Baseline	Balanced	Restricted
Age	-0.048**	-0.120**	-0.219**	-0.106*
	(-3.44)	(-5.69)	(-7.73)	(-1.91)
Age squared	-0.000	0.001**	0.002**	0.000
	(-0.97)	(2.30)	(5.52)	(0.62)
Grade	0.068**	0.004	-0.008	-0.047
	(2.60)	(0.09)	(-0.17)	(-0.47)
Promotion		0.099**	0.096**	0.180*
		(2.78)	(2.11)	(1.78)
Constant	6.231**	8.315**	10.278**	8.128**
	(18.86)	(16.40)	(14.68)	(5.82)
Observations	11778	8800	3894	1280
R-squared	0.158	0.180	0.208	0.177

Notes: *SAH*=Self-Assessed Health, 1(poor) - 5(excellent). *Baseline sample* is our default sample and includes respondents entering the Civil Service after 1969 and did not enter at the highest grade. *The balanced sample* consists of civil servants who stayed in the WII sample for all six phases. *The restricted sample* consists of individuals who entered the Civil Service between 1980-1985 and entered at the lowest grade (grade 1). *Grade* refers to the respondent's grade in the Civil Service. The WII Survey use three categories, 1-3, with 3 coded as the higher ranking. *Promotion* is a dummy of value 1 if the respondent has received a promotion between the last survey round and the current. **p<0.05, *p<0.10. t-stats in parentheses.

Table 3.2: The health effect of promotions

Panel B: Results for General Health Questionnaire score

	(1)	**(2)**	**(3)**
LHS variable	GHQ	GHQ	GHQ
Sample	Baseline	Baseline	Balanced
Age	0.000	0.148	-0.031
	(0.00)	(0.84)	(-0.13)
Age squared	0.000	-0.001	0.001
	(0.21)	(-0.74)	(0.26)
Grade	-0.129	-0.019	-0.213
	(-0.62)	(-0.05)	(-0.49)
Promotion		0.143	0.240
		(0.47)	(0.60)
Constant	3.532	0.012	3.533
	(1.31)	(0.00)	(0.61)
Observations	10009	7122	3102
R-squared	0.001	0.001	0.001

Notes: *GHQ*=General Health Questionnaire, the Whitehall II Survey (WII) uses the 30 item version and a score of 5 and higher is considered a 'case'. *Baseline sample* is our default sample and includes respondents entering the Civil Service after 1969 and did not enter at the highest grade. *The balanced sample* consists of civil servants who stayed in the WII sample for all six phases. *The restricted sample* consists of individuals who entered the Civil Service between 1980-1985 and entered at the lowest grade (grade 1). *Grade* refers to the respondent's grade in the Civil Service. The WII Survey use three categories, 1-3, with 3 coded as the higher ranking. *Promotion* is a dummy of value 1 if the respondent has received a promotion between the last survey round and the current. **p<0.05, *p<0.10. t-stats in parentheses.

Table 3.3: Effects of the timing of a promotion on Self-Assessed Health using a tenure variable which captures number of years in the Civil Service

	(1)	**(2)**	**(3)**
LHS variable	SAH	SAH	SAH
Sample	**Baseline**	**Balanced**	**Restricted**
Age	-0.107**	-0.263**	-0.017
	(-3.67)	(-6.37)	(-0.23)
Age squared	0.001**	0.001**	0.000
	(2.26)	(4.91)	(0.64)
Promotion	0.193**	0.282**	0.222
	(2.34)	(2.49)	(1.18)
Tenure	-0.013	0.058**	-0.092*
	(-0.63)	(1.98)	(-1.73)
Promotion*tenure	-0.008*	-0.011*	-0.003
	(-1.79)	(-1.82)	(-0.17)
Grade	0.042	0.027	-0.043
	(1.03)	(0.51)	(-0.37)
Constant	7.825**	11.554**	4.530*
	(9.60)	(10.75)	(1.83)
Observations	8363	3657	1215
R-squared	0.185	0.219	0.180

Notes: *SAH*=Self Assessed Health, 1(poor) - 5(excellent). *Baseline sample* is our default sample and includes respondents entering the Civil Service after 1969 and did not enter at the highest grade. *The balanced sample* consists of civil servants who stayed in the Whitehall II (WII) sample for all six phases. *The restricted sample* consists of individuals who entered the Civil Service between 1980-1985 and entered at the lowest grade (grade 1). *Grade* refers to the respondent's grade in the Civil Service. The WII Survey use three categories, 1-3, with 3 coded as the higher ranking. *Promotion* is a dummy of value 1 if the respondent has received a promotion between the last survey round and the current. *tenure* measures the number of years since a respondent entered the Civil Service. **p<0.05, *p<0.10. t-stats in parentheses.

Table 3.4: Effects of the timing of a promotion on Self-Assessed Health using a tenure variable which captures the number of years without a promotion

	(1)	**(2)**	**(3)**
LHS variable	SAH	SAH	SAH
Sample	**Baseline**	**Balanced**	**Restricted**
Age	-0.141**	-0.211**	-0.072
	(-5.71)	(-6.26)	(-1.12)
Age squared	0.001**	0.001**	0.000
	(2.79)	(4.33)	(0.30)
Promotion	0.212**	0.235*	0.303
	(2.52)	(1.92)	(1.57)
Tenure	0.003	0.007	-0.016
	(0.92)	(1.42)	(-1.13)
Promotion*tenure	-0.012**	-0.015**	0.005
	(-2.34)	(-2.18)	(0.28)
Grade	0.047	0.089	-0.199
	(0.80)	(0.99)	(-1.35)
Constant	8.730**	10.343**	7.102**
	(14.70)	(11.62)	(4.31)
Observations	5972	2591	1023
R-squared	0.189	0.218	0.172

Notes: *SAH*=Self-Assessed Health, 1(poor) - 5(excellent). *Baseline sample* is our default sample and includes respondents entering the Civil Service after 1969 and did not enter at the highest grade. *The balanced sample* consists of civil servants who stayed in the Whitehall II (WII) sample for all six phases. *The restricted sample* consists of individuals who entered the Civil Service between 1980-1985 and entered at the lowest grade (grade 1). *Grade* refers to the respondent's grade in the Civil Service. The WII Survey use three categories, 1-3, with 3 coded as the higher ranking. *Promotion* is a dummy of value 1 if the respondent has received a promotion between the last survey round and the current. *tenure* measures the number of years the respondent has worked in the Civil Service since the last promotion. **p<0.05, *p<0.10. t-stats in parentheses.

Table 3.5: The effects of including income variables in a cross-sectional regression of Self-Assessed Health on the grade variable

	(1)	**(2)**	**(3)**
LHS variable	SAH	SAH	SAH
Sample	**Phase 5**	**Phase 5**	**Phase 5**
Age	0.018	0.038	0.016
	(0.50)	(0.50)	(0.20)
Age squared	-0.000	-0.000	-0.000
	(-0.17)	(-0.43)	(-0.14)
Male	0.139**	0.066*	0.065*
	(5.72)	(1.83)	(1.77)
White	0.483**	0.354**	0.360**
	(12.50)	(6.35)	(6.24)
Income			
15-25K	0.149**	0.030	0.052
	(5.29)	(0.46)	(0.78)
25K-50K	0.206**	0.070	0.081
	(6.84)	(0.91)	(1.03)
>50K	0.304**	0.158*	0.180*
	(7.28)	(1.69)	(1.90)
Grade		0.081**	0.097**
		(2.33)	(2.73)
Number of years of education			-0.007*
			(-1.77)
Constant	1.899*	1.619	2.283
	(1.86)	(0.79)	(1.10)
Observations	7029	3214	3146
R-squared	0.050	0.036	0.038

Notes: *SAH*=Self-Assessed Health, 1(poor) - 5(excellent). The Whitehall II Survey (WII) has 4 income categories and we have excluded the lowest earning category. *Grade* refers to the respondent's grade in the Civil Service. The WII Survey use three categories, 1-3, with 3 coded as the higher ranking. *white* is a dummy indicating a white respondent, while *male* indicates a male person. **p<0.05, *p<0.10. t-stats in parentheses.

References

Adda J., Banks J. & von Gaudecker H-M. **(2008)** 'The Impact of Income Shocks on Health: Evidence from Cohort Data'. IZA DP # 3329

Aderibigbe Y. A. & Gureje O. **(1992)** 'The validity of the 28-item general health questionnaire in a Nigerian antenatal clinic'. *Social Psychiatry and Psychiatric Epidemiology*, 27(6)

Alderman H., Hoddinott J. & Kinsey B. **(2006)** 'Long term consequences of early childhood malnutrition'. *Oxford Economic Papers*, 68: 450-474

Almond D. **(2006)** 'Is the 1918 Influenza Pandemic Over? Long-Term Effects of In Utero Influenza Exposure in the Post-1940 U.S. Population'. *Journal of Political Economy*, 114(4)

Almond D. & Chay K. **(2006)**. 'The Long-Run and Intergenerational Impact of Poor Infant Health: Evidence from Cohorts Born During the Civil Rights Era'. NBER, http://www.nber.org/~almond/chay_npc_paper.pdf

Andersen H. S., Sestoft D., Lillebaek T., Gabrielsen G. & Hemmingsen R. **(2002)** 'Validity of the General Health Questionnaire (GHQ-28) in a prison population: Data from a randomized sample of prisoners on remand'. *International Journal of Law and Psychiatry*, 25(6): 573-580

Anderson M. & Marmot M. **(2007)** 'The Effects of Social Status on Heart Disease: Evidence from Whitehall'. Working Paper

Anotov A. N. **(1947)** 'Children born during the siege of Leningrad in 1942'. *Journal of Pediatrics,* 1947(30): 250-259

Banerjee A., Duflo E., Poestel-Vinay G. & Watts T. M. **(2007)** 'Long Run health Impacts of Income Shocks: Wine and Phylloxera in 19th Century France'. NBER WP # 12895

Banks J., Marmot M., Oldfield Z. & Smith J. P. **(2006)** 'The SES Health Gradient on both sides of the Atlantic'. The Institute for fiscal studies, WP07/04.

Barker D.J.P. **(1992)** *Fetal and Infant Origins of Adult Disease: Papers.* London: British Medical Journal

Barker D.J.P. **(1998)** *Mothers, babies and health in later life.* Edinburgh: Churchill Livingstone, 1998

Bawah (2001) http://www.iussup.org/Brazil2001/s60/S69_01_bawah.pdf

Bentsen S. B., Henriksen A. H. & Wentzel-Larsen T., Hanestad, B. R. & Klopstad Wahl, A. **(2008)** 'What determines subjective health status in patients with chronic obstructive pulmonary disease: importance of symptoms in subjective health status of COPD patients'. *Health and Quality of Life Outcomes*, 6:115

Boyce C. J. & Oswald A. J. **(2008)** 'Do People Become Healthier after Being Promoted?' IZA DP # 3894

Berger A. **(2001)** 'Insuline-like growth factor and cognitive function' *British Medical Journal*, 322: 203

Brunner E., Marmot M., Canner R., Beksinska M., Davey Smith G. & O'Brien, J. **(1996)** 'Childhood social circumstances and psychosocial and behavioural factors as determinants of plasma fibrinogen'. *Lancet*, 347: 1008-1013

Cameron L. A. **(2002)** 'The Impact of the Indonesian Financial Crisis on Children: Data from 100 Villages Survey'. World Bank Policy Research Working Paper # 2799. Available at SSRN: http://ssrn.com/abstract=634425

Case A. **(2001)**. 'Does Money protect health status? Evidence from South African pensions'. NBER WP # 8495

Case A., Lubotsky D. & Paxson C. **(2002)** 'Economic Status and Health in Childhood: The origins of the Gradient'. *The American Economic review*, 92(5): 1308-1334

Case A. & Paxson, C. **(2008)** 'Stature and Status: Height, Ability, and Labor Market Outcomes'. *Journal of Political Economy*, 116: 499-532

Currie A., Shields M & Wheatley Price S. **(2004)** 'Is the Child Health/Family Income Gradient Universal? Evidence from England'. IZA Discussion Paper # 1328

Dancause K.N., Lapante D.P., Oremus C., Fraser S., Brunet A.. & King S. **(2011)** 'Disaster-related prenatal stress influences birth outcomes: Project Ice Storm'. *Early Human Development*, 2011: Jul 22

Dean R.F.A. **(1951)** 'The size of the baby at birth and the yield of breast milk. Studies of undernutrition'. Wuppertal 1946-9. London: HMSO, 1951, 346-378

Deaton A. S. **(1989)** 'Rice prices and income distribution in Thailand: A non-parametric analysis'. *The Economic Journal*, 99: 1-37

Deaton A. S. & Paxson C. **(1999)** 'Mortality, Education, Income and Inequality among American cohorts'. NBER WP # 7140

deCatanzaro D. & Macniven E. **(1992)** 'Psychogenic pregnancy disruptions in mammals'. *Neurosci. Biobehav. Rev.*, 1992(16): 43-53

Dickerson S. S. & Kemeny M. E. **(2004)** 'Acute Stressors and Cortisol Responses: A Theoretical Integration and Synthesis of Laboratory Research.' *Psychological Bulletin*, 130(3): 355-391

Flores-Lagunes A. **(2007)** 'Finites Sample Evidence of IV Estimators under Weak Instruments'. *Journal of Applied Econometrics*, 22: 677-694

Freeman H. E., Klein R. E., Townsend J. W. & Lechtig A. **(1980)** 'Nutrition and cognitive development among rural Guatemalan children'. *American Journal of Public Health*, 70: 1277-1285

Frijters P., Haisken-DeNew J. P. & Shields M. A. **(2005)** 'The causal effect of income on health: Evidence from German reunification'. *Journal of Health Economics*, 24: 997-1017

Geary C. G., Saults S. J., Liu F. & Hoard M. K. **(2000)** 'Sex Differences in Spatial Cognition, Computational Fluency, and Arithmetical Reasoning'. *Journal of Experimental Child Psychology,* 77: 337-353

Gigante D. P., Horta B. L., Lima R. C., Barros F. C. & Victoria C. G. **(2006)** 'Early Life Factors Are determinants of Female Height at Age 19 years in a Population-Based Birth Cohort (Pelotas, Brazil)'. *The Journal of Nutrition*, 136: 473-478

Glewwe P. & Jacoby H. **(1995)** 'An economic analysis of delayed primary school enrollment and childhood malnutrition in a low income country'. *Review of Economics and Statistics*, 77: 156-169

Glewwe P. & King E. **(2001)** 'The impact of early childhood nutritional status on cognitive development: does the timing of malnutrition matter?'. *World Bank Economic Review*, 15: 81-114

Grantham-McGregor S. M., Walker S. P., Chang S. M. & Powell C. A. **(1997)** 'Effects of early childhood supplementation with and without stimulation on later development in stunted Jamaican children'. *American Journal of Clinical Nutrition*, 66: 247-253

Hoddinott J. & Kinsey B. **(2001)** 'Child Growth in the Time of Drought'. *Oxford Bulletin of Economics and Statistics*, 63(4), 409-436

Horton S. **(1988)** 'Birth order and child nutritional status: evidence from the Philippines'. *Economic Development and Cultural Change*, 36: 341-354

Huisman M., van Lenthe F. & Mackenbach J. **(2007)** 'The predictive ability of self-assessed health for mortality in different educational groups'. *International Journal of Epidemiology*, 36: 1207-1213

Idler E. L. & Benyamini Y. **(1997)** 'Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies'. *Journal of Health and Social Behaviour*, 38(March): 21-37

Jhalani J., Goyal T., Clemow L., Schwartz J. E., Pickering T. G. & Gerin W. **(2005)** 'Anxiety and outcome expectations predict the white-coat effect'. *Blood Pressure Monitoring,* 10 (6): 317-319

Johnston D. W., Propper C. & Shields M. A. **(2009)** 'Comparing subjective and objective measures of health: Evidence from hypertension for the income/health gradient'. *Journal of Health Economics*, 28: 540-552

Karasek R. A. **(1979)** 'Job demands, job decision latitude, and mental strain: Implications for job redesign'. *Administrative Science Quarterly*, 24: 285-308

Lanjouw P., Pradhan M., Saadah F., Sayed H. & Sparrow R. **(2001)** 'Poverty, Education, and Health in Indonesia. Who Benefits from Public Spending?' World Bank Policy Research Working Paper # 2739

Lechtig A. & Klein R. E. **(1981)** 'Pre-natal nutrition and birth weight: Is there a causal association?' Dobbing J, ed. *Maternal nutrition in pregnancy-eating for two?* London: Academic Press, 1981, 131-174

Lynn R. **(1990)** 'The role of nutrition in secular increases in intelligence'. *Personality and Individual Differences*, 11(3): 273-285

Maccini S. & Yang D. **(2007)** 'Under the Weather: Health, Schooling, and Socioeconomic Consequences of Early-Life Rainfall'. Working paper, University of Michingan

Martorell R. **(1995)** 'Results and implications of the INCAP follow-up study'. *Journal of Nutrition,* 1995; 125(4S): 1127S-1138S

Martorell R. **(1999)** 'The nature of child malnutrition and its long-term implications'. *Food and Nutrition Bulletin*, 20(3): 288-292

Martorell R., Stein A., Wang M., DiGirolamo A., Grajeda R., Ramakrishnan U., Ramirez-Zea M. & Yount K. **(2008)** 'Nutritional Supplementation in Early Childhood, Schooling, and Intellectual Functioning in Adulthood'. *Arch Pediatric Adolescence Medicine*, 162(7): 612-618

Marmot M .G., Bosma H., Hemingway H., Brunner E. & Stansfield S. **(1997)** 'Contributions of job control and other risk factors to social variations in coronary heart disease incidence'. *Lancet*, 350: 235-239

McEwen B. S. **(1998)** 'Protective and damaging effects of stress mediators'. *New England Journal of Medicine*, 338:171-179

Meng X. & Qian N. **(2006)** 'The Long Run Health and Economic Consequences of Famine on Survivors: Evidence from China's Great Famine'. IZA Discussion Paper #2471

Van Tilburg M.A., Vingerhoets J. J., Van Heck G.L. & Kirschbaum C. **(1999)** 'Homesickness, mood and self-reported health'. *Stress Medicine*, 15: 189-196

Mossey J. M. & Shapiro E. **(1982)** 'Self-rated health: a predictor of mortality among the elderly'. *American Journal of Public Health*, 72: 800-808

North F., Syme S. L., Head J., Shipley M. J., Marmot M. G. **(1993)** 'Explaining socioeconomic differences in sickness absence: The Whitehall II study'. *British Medical Journal*, 306: 361-366

Pettit J.W., Kline J.P., Gencoz T., Gencoz F. & Joiner T.E. **(2001)** 'Are happy people healthier? The specific role of positive affect in predicting self-reported health symptoms'. *Journal of Research in Personality*, 2001

Persico N., Postlewaite A. & Silverman D. **(2004)** 'The Effect of Adolescent Experience on Labor Market Outcomes: The Case of Height'. *Journal of Political Economy,* 112: 1019-1053

Romans-Clarkson S. E., Walton V. A., Herbison G. P. & Mullen P. E. **(1989)** 'Validity of the GHQ-28 in New Zealand Women'. *Australian and New Zealand Journal of Psychiatry*, 23(2): 187-196

Rose S. A. **(1994)** 'Relation between Physical Growth and Information Processing in Infants Born in India'. *Child Development*, 65(3): 889-902

Pritchett L. & Sumarto S. **(2002)** 'Targeted Programs in an Economic Crisis: Empirical Findings from Indonesia's Experience'. CID Working Paper # 95

Sapolsky R. M. **(1993)** 'Endocrinology alfresco: Psychoendocrine studies of wild baboons'. *Recent Progress in Hormone Research*, 48:437-468

Sihvely C. A. & Clarkson T.B. **(1994)** 'Social status and coronary artery atherosclerosis in female monkeys'. *Arteriosclerosis and Thrombosis*, 14: 721-726

Singh-Manoux A., Gueguen A., Martikainen P., Ferrie J., Marmot M. G. & Shipley M. **(2007)** 'Self-Rated Health and Mortality: Short- and Long-Term Associations in the Whitehall II Study'. *Psychosomatic Medicine*, 63: 138-143

Smith C. A. **(1947)** 'The effect of wartime starvation in Holland upon pregnancy and it product'. *American Journal of Obstetrics and Gynecology*, 53: 599-608

Smith G. D., Shipley M. J. & Rose G. **(1990)** 'Magnitude and causes of socioeconomic differentials in mortality: further evidence from the Whitehall Study'. *Journal of Epidemiology and Community Health*, 44 (4): 265-270

Smith J. R., Brooks-Gunn J & Pamela K. **(1997)** 'The Consequences of Living in Poverty for Young Children's Cognitive and Verbal Ability and Early School Achievement'. *Consequences of Growing Up Poor*, ed. Greg J. Duncan and Jeanne Brooks-Gunn. New York: Russel Sage

Smith P. M., Frank J. W., Mustard C. A. & Bondy S. J. **(2008)** 'Examining the relationships between job control and health status: a path analysis approach'. *Journal of Epidemiological Community Health*, 62: 54-61

Strauss J. & Thomas D. **(1995)** 'Human Resources: Empirical Modelling of Household and Family Decisions'. *Handbook of Development Economics Vol. 3*, ed. Behrman, J & Srinivasan, T. N

Sumarto S., Wetterberg A. & Pritchett L. **(1999)** 'A National Snapshot of the Social Impact of Indonesia's Crisis'. *Bulletin of Indonesian Economic Studies*, 35(3): 145-152

Sundet J. M., Tambs J. R., Magnus P. & Torjussen T. M. **(2005)** 'Resolving the Genetic and Environmental Sources of Correlation Between Height and Intelligence: A Study of Nearly 2600 Norwegian Male Twin Pairs'. *Twin Research and Human Genetics*, 8(August): 307-311

Surono S. **(1999)** 'Food security and Social Safety Net in Indonesia'. Presented at International Seminar 'APEC Study Center Consortium', 1999 Conference, 31 May-2 June, Auckland, New Zealand

Suryahadi A., Sumarto S., Suharso Y & Pritchett L. **(2002)** 'The Evolution of Poverty during the Crisis in Indonesia, 1996-99'.The World Bank Policy Research Working Paper # 2435

Thomas D., Frankenberg E. & Beegle K. **(1999)** 'Health, Education, and the Economic Crisis in Indonesia'. Paper prepared for the 1999 Population Association of America meetings, New York, March 25-27

Torche F. **(forthcoming in Demography)** 'The Effect of Maternal Stress on Birth Outcomes: Exploiting a Natural Experiment'.

Victoria C. G., Adair L., Fall C., Hallal P. C., Martorell R., Richter L. & Sing Sachdev H. **(2008)** 'Maternal and child undernutrition: consequences for adult health and human capital'. *Lancet*, 371(9609): 340-357

Vitetta L., Anton B., Cortizo F. & Sali A. **(2005)** 'Mind-Body Medicine: Stress and Its Impact on Overall Health and Longevity'. *Annals of New York Academy of Science*, 1057

Waber D. P., Vuori-Christiansen L., Ortiz N., Clement J. R., Christiansen N. E., Mora J. O., Reed R. B. & Hererra M. G. **(1981)** 'Nutritional supplementation, maternal education and cognitive development of infant at risk of malnutrition'. *American Journal of Clinical Nutrition*, 34: 807-813

Watanabe K., Flores R., Fujiwara J. & Tran L. T. H. **(2005)** 'Early Childhood Development Interventions and Cognitive Development of Young Children in Rural Vietnam'. *The Journal of Nutrition*, 135(8): 1918-1925

Wilkinson R. G. **(1992)**. 'Income distribution and life expectancy'. *British Medical Journal*, 304:165-168

Wilkinson R. G. **(2005)** *The Impact of Inequality: How to make sick societies healthier*. Routledge

Wilkinson R. G. **(2006)** *Social Determinants of Health*. Oxford University Press

Wilkinson R. G. **(2009)** *The Spirit Level: Why More Equal Societies Almost Always Do Better*. Allen Lane