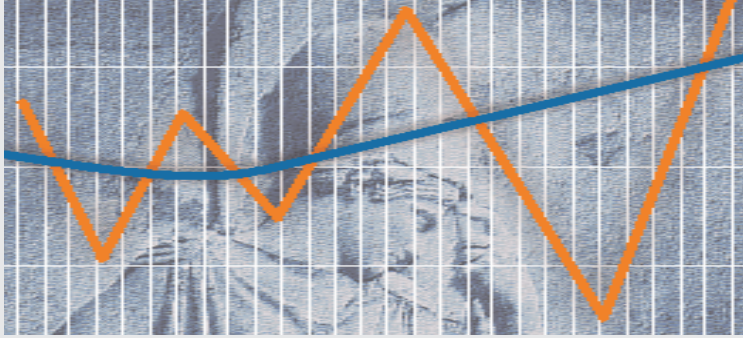


ifo Beiträge
zur Wirtschaftsforschung

44

Three Empirical Essays on the Long-Run Consequences of Early-Life Living Conditions

Sven Neelsen

ifo Institut
Leibniz-Institut für Wirtschaftsforschung
an der Universität München e.V.

Herausgeber der Reihe: Hans-Werner Sinn
Schriftleitung: Chang Woon Nam

ifo Beiträge zur Wirtschaftsforschung
44

Three Empirical Essays on the Long-Run Consequences of Early-Life Living Conditions

Sven Neelsen

ifo Institut
Leibniz-Institut für Wirtschaftsforschung
an der Universität München e.V.

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation
in der Deutschen Nationalbibliografie; detaillierte bibliografische
Daten sind im Internet über
http://dnb.d-nb.de
abrufbar

ISBN-13: 978-3-88512-525-9

Druck: ifo Institut, München

ifo Institut im Internet:
http://www.cesifo-group.de

Preface

This study was prepared by Sven Neelsen while he was working with the ifo Institute for Economic Research. It was completed in December 2011 and accepted as a doctoral thesis by the Department of Economics at the University of Munich in May 2012. The study investigates long-run effects of early-life living conditions using micro-datasets from three countries. The employed empirical strategies aim to identify causal relationships between early-life living conditions and the outcomes of interest.

Chapter 1 analyzes effects of early-life exposure to the Greek 1941/2 famine on school and labor-market performance. I find famine exposure in the first and second years of life to be associated with lower educational attainment in four waves of the Greek decennial census between 1971 and 2001. Chapter 2 estimates long-run effects of fetal exposure to the 1918-19 influenza pandemic for Switzerland. Using data from the 1970 Swiss census, I find that the male 1919 cohort that had a strongly increased likelihood of fetal exposure to the pandemic performs significantly worse in terms of educational attainment and has a lower chance of marriage than the surrounding cohorts. Chapter 3 examines mid-run effects of early childhood exposure to a large-scale Indonesian midwife placement program. Using panel data from the 1993, 1997, 2000, and 2007 waves of the Indonesian Family Life Survey, I find improvements in height-for-age and cognitive skill for adolescents with exposure to the program during infancy.

Keywords: In utero, early childhood development, long-run effects, malnutrition, influenza, midwife, education, cognitive skill, human capital, labor market, Greece, Switzerland, Indonesia

JEL-Nr. H43, I0, I12, I15, I18, I21, J13, J24,

Gewidmet meinen Eltern

Danksagung

Mein Dank gilt: Meiner Familie für ihre liebevolle Unterstützung, zu jeder Zeit und in jeder Stimmung; meinem akademischen Lehrer und Ko-Autoren Prof. Thomas Stratmann, Ph.D.; meinem Doktorvater Prof. Dr. Joachim Winter; Dr. Thomas Strobel für viele beantwortete Fragen und Hinweise; Dr. Wolfgang Ochel; meinen Freunden.

Three Empirical Essays on the Long-Run Consequences of Early-Life Living Conditions

Inaugural-Dissertation
zur Erlangung des Grades
Doctor oeconomiae publicae (Dr. oec. publ.)
an der Ludwig-Maximilians-Universität München

2011

vorgelegt von
Sven Neelsen

Referent: Prof. Dr. Joachim Winter
Korreferent: Prof. Thomas Stratmann, Ph.D.
Promotionsabschlussberatung: 16. Mai 2011

Table of Contents

List of Tables

List of Figures

Introduction

It is now well established in the medical and biological literatures that early-life living conditions are crucial to physical and cognitive development (see, for instance, Barker 1992 and Shonkoff and Phillips 2000). Optimal development requires adequate early-life nutrition, protection from disease, and stimulating psychosocial environments (Walker et al. 2007). For over 200 million of today's children under the age of five, one or several of these requirements are not met, be it through undernutrition or unbalanced diets, lacking access to basic healthcare, or inadequate parental stimulation (Grantham-McGregor et al. 2007).

Such inadequate living conditions caused over six million deaths of children under five in 2008, equaling to two thirds of global deaths in this age group (Black et al. 2010). The first two chapters of this dissertation are concerned with the many million who survive early-life hardship. Chapter 1 examines education and labor market consequences of early-life exposure to the Greek 1941-42 famine and Chapter 2 long-run effects of fetal exposure to the 1918-19 influenza pandemic for Switzerland. Chapter 3 evaluates an intervention to improve early-life living conditions in a developing country. Specifically, I investigate effects of early-life exposure to an Indonesian midwife placement program on adolescent height and cognitive ability.

Earlier work on long-run consequences of early-life malnutrition and disease separates into two methodological strands. The first strand compares long-run outcomes of individuals with early-life malnutrition or disease – typically proxied by early-life growth retardation – to long-run outcomes of individuals without. This approach is most common in the medical literature. Victoria et al. (2008) review five such studies from low- and middle-income countries and contextualize their findings with the current state of the literature. The sum of evidence, they conclude, suggests positive associations of birth weight and early-life height with adult height, weight, educational attainment, income, wealth, and offspring body size. Further, Victoria et al. (2008) find consistent evidence for negative associations with adult type II diabetes and constrained lung functioning. Walker et al. (2007) summarize results of studies on later-life effects of more specific types of malnutrition. The authors conclude that there are consistent relationships between early-life iodine and iron deficiency and impairments of mental health and cognitive development whereas the results from studies for zinc and micronutrient deficiencies, and short breastfeeding duration are more mixed. Walker et al. (2007) furthermore conclude that infant HIV infection and severe or repeated early-life malaria impair the development of cognitive abilities. In comparison, studies on long-run effects of early-life diarrhea and intestinal helminthes infection have yielded mixed results.

One methodological issue with studies that compare individuals with early-life malnutrition or disease to individuals without is that omitted variable bias (OVB) can lead to an overstatement of long-run effects. This occurs when the investigator cannot observe and thus control for harmful circumstances like poor parental education and low parental income that typically coincide with early-life malnutrition and disease. As a result, such

studies identify composite effects of poor living conditions but not the specific effects of the treatment of interest.

Scholars, in particular in the field of economics, have respondent to this issue with different methodological approaches (Almond and Currie 2011). One approach uses monozygotic twin samples to identify long-run birth-weight effects. Because monozygotic twins have identical genetic endowments and typically grow up in the same households, the differencing out of twin fixed effects excludes genetic and most family-level sources of OVB. The twin studies find consistent associations of low birth-weight with lower educational attainment (Behrman and Rosenzweig 2004 and Oreopoulos et al. 2008 for the US, Black et al. 2007 for Norway, and Lin and Liu 2009 for Taiwan) and lower adult height (Behrman and Rosenzweig 2004, Black et al. 2007). Unlike Behrman and Rosenzweig (2008) and Royer (2009) for the US, Black et al. (2007) moreover find negative income effects for Norway. Finally, Royer (2009) finds evidence for negative effects on the health of children but not on the health of adults for the US.

Despite its power in addressing the OVB issue, the twin fixed effects approach has its boundaries. First, scholars have questioned the transferability of findings from twin to non-twin populations (see Black et al. 2007 for a discussion and test). Second, large twin datasets with follow-up into adulthood exist only for a limited number of developed countries. Therefore, there are no twin studies for developing countries where early-life malnutrition and disease are most prevalent.

An alternative approach to avoid the OVB-issues that arise in comparisons of individuals with and without early-life malnutrition and disease is that of instrumental variables. To qualify as an instrument, a variable needs to be correlated with the endogenous treatment variable, e.g. early-life malnutrition or disease, but uncorrelated with unobserved determinants of long-run outcomes that correlate with the treatment (see Wooldridge 2002, chapter 5). Various studies have used famines or epidemics affecting entire regions for limited periods of time as such instruments. In this approach, the treatment groups consist of cohorts that lived in the affected areas early in life. The control group can include cohorts from the affected areas that were older during the famine, cohorts that were conceived in the affected areas after the famine, and cohorts of the same age as the treated cohorts who lived in non-famine-affected areas early in life.

The Chinese 1959-61 famine is the most prominent case in the economic literature with at least 30 studies on long-run effects as of 2011. The sum of evidence indicates that the surviving members of the early-life exposed cohorts have worse adult physical and mental health, and perform worse in terms of socioeconomic outcomes like education and wealth (see, for instance St. Clair et al. 2005, Meng and Qian 2006, Chen and Zhou 2007 and Huang et al. 2010). The most researched case in the medical literature is the Dutch 1944-45 famine with over 160 articles published to date. Long-run effects of fetal exposure depend on gestational age and include increased rates of obesity, type II diabetes, coronary heart disease, poor self-reported health, schizophrenia and depression in adulthood (see, for instance, Roseboom et al. 2011). Rather than instrumenting for

early-life malnutrition with the timing and location of birth, Macchini and Yang (2009) use variations in rainfall intensity in the early-life phases of Indonesian women. They find that infant exposure to drought adversely affects adult height, self-reported health, educational attainment, and economic status.

Chapter 1 contributes to this literature by examining long-run effects of early-life exposure to the Greek 1941-42 famine. Exploiting both the famine's timing and regional differences in intensity, my study finds adverse education and labor market effects for cohorts with fetal, infant and second year of life famine exposure. In addition to being the first to consider long-run effects of the Greek famine, my study's contribution is twofold.

First, the famine's short duration – only eight months – permits the estimation of long-run effects of malnutrition over a narrow time period, e.g. for malnutrition limited to the first life-year. In longer-lasting famines like that of China 1959-61, it is impossible to identify the effects of first year of life exposure. This is because the cohorts with first year of life exposure were in addition either exposed as fetuses (born 1961) or during the second year of life (born pre-1961).

Second, unlike other famine datasets, mine includes birthplace information which permits me to include birthplace fixed effects in my empirical model. Their inclusion reduces the famine effect estimates. Because my data indicate negative birthplace selection in the treated cohorts, I suspect that these reductions are at least in part the result of the fixed effects removing OVB from negative sample selection. This points to an overstatement of famine effects in specifications without birthplace control and highlights the threat of misspecification in famine studies. Future work may therefore prioritize the modeling of sample selection mechanisms, i.e. try to shed more light on differences in the backgrounds of the treatment and control groups.

The validity of the cohort membership instrument is, however, not the only methodological issue in the famine literature. Every famine is unique in that it occurs in a very specific environment: some famines coincide with political unrest or war, others with droughts, pests, or outbreaks of infectious disease. This *noisiness* limits the generalizability of findings from different famines. Moreover, it reduces the usefulness of famine studies for investigating the effects of malnutrition as a standalone treatment.

Summarizing recent developments in the famine literature, senior scholar Angus Deaton concludes that while famines remain a field of study in their own right, future investigations of long-run malnutrition effects should focus on less noisy instruments (Duh 2011). He identifies a recent study by Almond and Mazumder (forthcoming) who use differences in fetal Ramadan-exposure to instrument for early-life malnutrition as a possible blueprint for such future work.

Chapter 2 uses the 1918-19 influenza pandemic to instrument for fetal exposure to an infectious disease. Showing evidence for Switzerland, the study contributes to a literature that has found adverse effects on adult health and socioeconomic outcomes for the US (e.g. Almond 2006), Brazil (Nelson 2010), and Taiwan (Lin 2008). For several reasons, the issues of sample selection and noisiness in the instrument are less severe for the Swiss

influenza pandemic than for famines. First, the pandemic came suddenly and lasted only a few months, limiting the scope for behavioral responses. Second, unlike in famines, influenza mortality was not or not strongly associated with socioeconomic backgrounds, thus there is less concern about OVB from sample selection by mortality. Third, for Switzerland that was neutral in the First World War, the pandemic did not coincide with noisy events like war, severe political unrest or nutritional crisis. Finally, because Switzerland also remained neutral in the Second World War, there is no threat of later-life war-related sample selection like in the US influenza studies.

My empirical analysis exploits both over-time and regional differences in fetal influenza exposure. I find evidence for adverse effects on educational attainment, labor market performance and the likelihood to marry, in particular for males. The effects are substantially smaller than those in work for the US, despite similar infection rates. Future research may address this puzzle, for instance by collecting evidence from other European countries. Sweden that was also neutral in both World Wars may be a suitable candidate. Moreover, more evidence is needed on long-run socioeconomic effects of early-life exposure to other infectious diseases, in particular for those most prevalent in the developing world like malaria, diarrhea, or helminthes infection. A recent study by Barreca (2010) who instruments for early-life malaria infection with exposure to temperatures conducive to malaria in the US and finds reductions in years of schooling, may serve as an example.

The long-run consequences of poor early-life living conditions reflect on the macroeconomic level, as forgone human capital and rising healthcare costs reduce the competitiveness and growth potential of developing countries (Heckman and Masterov 2007, Naudeau et al. 2011 Section 1 for a review). Consequently, leading economists argue that successful early-life interventions yield high economic returns (Lomborg et al. 2004, 2009, Naudeau 2011 Section 1).

To shape such effective and cost-efficient interventions, more knowledge is needed about what works – and what does not. The evidence from developed countries is mixed. US home visit programs with regular consultations on nutrition, health and parenting practices for underprivileged families show consistent benefits on educational attainment, while impacts on labor market outcomes and risky behaviors are more varied (Almond and Currie 2011). The variation in net present benefits for these programs is large and costs per participant often exceed $30,000 (Almond and Currie 2011). This limits their transferability to developing countries. However, because the unmet needs of children in developing countries are much more basic than in the developed world, large improvements in early-life conditions may be achieved at much lower cost: the World Health Organization estimates that a basic set of interventions to reach the Millennium Development Goals of improving nutrition, maternal and child health, and access to essential medicines while fighting back HIV, malaria, and tuberculosis costs between $15 and $30 per year and person in 49 low income countries (World Health Organization 2010).

For developing countries, various studies have examined effects of early-life health and nutrition interventions on the health and cognitive ability of young children. Walker et al. (2007) review this evidence that mainly comes from randomized controlled trials. They conclude that food and iron supplementation in the first life years, and iodine supplementation during pregnancy consistently improve child outcomes. In comparison, the evidence for effects of food supplements for pregnant women and zinc supplements for young children is mixed.

Food supplementation interventions can form part of Conditional Cash Transfer programs. In these programs, families receive financial support upon fulfilling requirements like regularly taking their children to clinics for checkups, immunizations, or nutritional supplementation, enrolling them in school, and/or participating in parenting practice lessons. Such integrated programs have been shown to increase child height (Fernald at al. 2008), health (Gertler 2004), and cognitive ability (Grantham-McGregor et al. 1991, Macours et al. 2008).

In contrast to the growing number of studies on the short-run effects of early-life interventions, the evidence on mid- and long-run effects is scarce and inconsistent (Alderman 2011, Walker et al. 2007): Pollitt et al. (1997) and Walker et al. (2005) find limited or no effects of early-life supplemental feeding on the cognitive ability of older children in Indonesia and Jamaica, whereas Pollitt et al. (1993) find large such benefits for food supplementation throughout the first two years of life for Guatemala. Similarly, Hoddinott et al. (2008) find that a randomized nutrition intervention in Guatemala increased male wages by 46 percent if received before age three but had no effects when received later.

Chapter 3 contributes to the closing of this knowledge gap by examining effects of infant exposure to an Indonesian midwife program on adolescent height and cognitive skill. The midwife program is a large scale intervention that, from the beginning of the 1990s, dispatched over 50,000 trained midwifes to needy communities throughout the country (Frankenberg et al. 2005). The program's primary goal was to reduce maternal mortality through prenatal care and skilled birth attendance but the midwives were also trained and equipped to provide food supplements, early-life preventive and curative care, and advice on oral rehydration, breastfeeding, and child nutrition. Moreover, their responsibilities included the raising of awareness for public health issues like sanitation and personal hygiene in the community. Earlier evaluations of the program have shown positive short-run effects. Hatt et al. (2009) find reductions in neo-natal and Shresta (2010) in neo- and post-natal mortality, and Frankenberg et al. (2005) increases in height-for-age among young children with program exposure. My study adds to this literature that midwife presence in infancy increases adolescent height-for-age and cognitive ability. The effects are stronger for girls and restricted to children of poorly educated mothers.

To facilitate a comprehensive cost-benefit-analysis of the program, future work should investigate if the benefits I measure for adolescents also extend into improved adult outcomes. Once the full scope of program costs and benefits is known, comparisons with

other interventions, for instance those addressing the living-conditions of children of older age, may enable assessments of the program's relative efficiency and its potential to serve as a blueprint for similar programs in other developing countries. To facilitate such broad-based cost-effectiveness comparisons, practitioners underscore the need for more studies on mid- and long-run effects of early-life interventions in developing countries (Alderman 2011). Future work should address this need.

In its 1946 constitution, the World Health Organization declares the enjoyment of the highest attainable standards of health a basic human right (World Health Organization 1946). In this spirit, this dissertation aims to contribute to a growing understanding that investments in child development are not only ethically right – but also the smart thing to do.

References

Alderman, H. 2011. "Introduction". In: Alderman, Harold (ed.): No Small Matter - The Impact of Poverty, Shocks, and Human Capital Investments in Early Childhood Development, The World Bank, Washington D.C.: 9-50.

Almond, D. and J. Currie. 2011. "Human capital development before age 5". In: Orley Ashenfelter and David Card (eds.): Handbook of Labor Economics, vol. 4B, Elsevier, Amsterdam: 1315-1486.

Almond, D. 2006. "Is the 1918 influenza pandemic over? Long-run effects of in utero influenza-exposure in the post-1940 U.S. population". *Journal of Political Economy*, 114(4), 672-711.

Almond, D. and B. Mazumder. Forthcoming. "Health capital and the prenatal environment: the effect of Ramadan observance during pregnancy". *American Economic Journal: Applied Economics*.

Barker, D. (ed.). 1992. Fetal and infant origins of adult disease: Papers. British Medical Journal, London.

Barreca, A. 2010. "The Long-run Economic Impact of In Utero and Postnatal Exposure to Malaria". *Journal of Human Resources*, 45(4): 865-892.

Behrman, J. and M. Rosenzweig. 2004. "Returns to birthweight." *The Review of Economics and Statistics*, 86(2): 586-601.

Black, R., S. Cousens, H. Johnson, J. Lawn, I. Rudan, D. Bassani, P. Jha, H. Campbell, C. Fischer Walker, R. Cibulskis, T. Eisele, L. Liu, and C. Mathers. 2010. "Global, regional,

and national causes of child mortality in 2008: a systematic analysis". *Lancet*, 375: 1969-1987.

Black, S., P. Devereux, and K. Salvanes. 2007. "From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes". *The Quarterly Journal of Economics*, 122(1): 409-439.

Chen, Y. and L. Zhou. 2007. "The long-run health and economic consequences of the 1959-61 famine in China." *Journal of Health Economics*, 26: 659-681.

Duh, J. 2011. "Fetal origins, early childhood exposure, and famine: conference summary". Unpublished manuscript, Center for Health and Wellbeing, Princeton University.

Fernald, L., P. Gertler, and L. Neufeld. 2008. "Role of cash in conditional cash transfer programmes for child health, growth, and development: an analysis of Mexico's Oportunidades". *Lancet*, 371(9615): 828-837.

Frankenberg, E., W. Suriastini, and D. Thomas. 2005. "Can expanding access to basic healthcare improve children's health status? Lessons from Indonesia's 'midwife in the community' programme". *Population Studies*, 59(1): 5-19.

Gertler, P. 2004. "Do conditional cash transfers improve child health? evidence from progresa's control randomized experiment". *Health, Health Care, and Economic Development*, 94(2): 336–341.

Grantham-McGregor, S., C. Powell, S. Walker, and J. Himes. 1991. "Nutritional supplementation, psychosocial stimulation, and mental development of stunted children: the Jamaican Study". *Lancet*, 338: 1-5.

Grantham-McGregor S., Y. Cheung, S. Cueto, P. Glewwe, L. Richter, B. Strupp, and International Child Development Steering Group. 2007. "Developmental potential in the first five years for children in developing countries." *Lancet*, 369: 60-70.

Hatt, L., C. Stanton, C. Ronsmans, K. Makowiecka, and A. Adisasmita. 2009. "Did professional attendance at home births improve early neonatal survival in Indonesia?". *Health Policy and Planning*, 24: 270–278.

Heckman, J. and D. Masterov. 2007. "The Productivity Argument for Investing in Young Children." *Review of Agricultural Economics*, 29(3): 446–493.

Hoddinott, J., J. Maluccio, J. Behrman, R. Flores, and R. Martorell. 2008. "Effect of a Nutrition Intervention during Early Childhood on Economic Productivity in Guatemalan Adults". *Lancet*, 371: 411-6.

Huang, C., Z. Li, M. Wang, and R. Martorell. 2010. "Early-life exposure to the 1959-61 Chinese famine has long-run health consequences". Journal of Nutrition, 140: 1874-1878.

Lin, M. 2008. “Does "in utero" pandemic exposure matter? Evidence from the long term effect of the 1918 influenza on educational attainments in Taiwan”. National Taiwan University Working Paper.

Lin, M. and J. Liu. 2009. “Do lower birth weight babies have lower grades? twin fixed effect and instrumental variable evidence from Taiwan”. *Social Science and Medicine*, 68(10): 1780-1787.

Lomborg, B., ed. 2004. Global Crises, Global Solutions. Cambridge University Press, Cambridge.

Lomborg, B., ed. 2009. Global Crisis, Global Solutions: Costs and Benefits. Cambridge University Press, Cambridge.

Maccini, S. and D. Yang. 2009. “Under the Weather: Health, Schooling, and Economic Consequences of Early-Life Rainfall”. *American Economic Review*, 99(3): 1006-1026.

Macours, K., N. Schady, and R. Vakis. 2008. “Cash transfers, behavioral changes, and cognitive development in early childhood: Evidence from a randomized experiment”. Working Paper 4759, The World Bank: Development Research Group, Human Development and Public Services Team.

Meng, X. and N. Qian. 2009. “The long-run consequences of famine on survivors: Evidence from a unique natural experiment using China’s Great Famine”. NBER Working Paper No. 14917.

Naudeau, S., N. Kataoka, A. Valerio, M. Neuman, and L. Elder. 2011. Investing in young children: an early childhood development guide for policy dialogue and project preparation”. The World Bank, Washington DC.

Nelson, R.. 2010. “Testing the fetal origins hypothesis in a developing country: Evidence from 1918 Influenza pandemic”. Health Economics, 19(10): 1181-1192.

Oreopoulos, P., M. Stabile, R. Walld, and L. Roos. 2008. “Short, medium, and long-run consequences of poor infant health: An anlysis using siblings and twins.” Journal of Human Resources, 43(1): 38-138.

Pollitt, E., K. Gorman, P. Engle, R. Martorell, and J. Rivera. 1993. “Early supplementary feeding and cognition: effects over two decades”. *Monographs of the Society for Research in Child Development*, 58: 1-99.

Pollitt E., W. Watkins, and M. Husaini. 1997. “Three-month nutritional supplementation in Indonesian infants and toddlers benefits memory function 8 y later”. *American Journal of Clinical Nutrition*, 66: 1357-1363.

Roseboom, T., R. Painter, A. van Abeelen, M. Veenendaal, and S. de Rooij. 2011. "Hungry in the womb: what are the consequences? Lessons from the Dutch famine". *Maturitas*, 70(2): 141-145.

Royer, H. 2009. "Separated at girth: Us twin estimates of the effects of birth weight". *American Economic Journal: Applied Economics*, 1(1): 49–85.

Shonkoff, J. and D. Phillips. 2000. From Neurons to Neighborhoods: The Science of Early Childhood Development. National Academy Press, Washington, DC.

Shresta, R. 2010. "The Midwife Program and Infant Mortality in Indonesia". *Bulletin of Indonesian Economic Studies*, 46(2): 193-211.

St. Clair, D., M. Xu, P. Wang, Y. Yu, Y. Fang, F. Zhang, X. Zheng, N. Gu, G. Feng, P. Sham, and L. He. 2005. "Rates of adult schizophrenia following prenatal exposure to the Chinese Famine of 1959-1961". *JAMA*, 294: 557-562.

Victoria, C., L. Adair, C. Fall, P. Hallal, R. Martorell, L. Richter, and H. Singh Sachdev. 2008. "Maternal and child undernutrition: consequences for adult health and human capital". *Lancet*, 371: 340-357.

Walker S., S. Chang, C. Powell, and S. Grantham-McGregor. 2005. "Effects of early childhood psychosocial stimulation and nutritional supplementation on cognition and education in growth-stunted Jamaican children: prospective cohort study". *Lancet*, 366: 1804-1807.

Walker S., T. Wachs, J. Meeks Gardner, B. Lozoff , G. Wasserman, E. Pollitt, J. Carter, and the International Child Development Steering Group. 2007. "Child development: risk factors for adverse outcomes in developing countries." *Lancet*, 369: 145-157.

Wooldridge, J. 2002. Econometric Analysis of Cross Section and Panel Data. MIT Press, Cambridge, Massachusetts.

World Health Organization. 1946. "Constitution of the World Health Organization". Accessed online 4th December 2011 at www.who.int/governance/eb/constitution/en/index.html.

World Health Organization. 2010. Constraints to Scaling Up the Health Millennium Development Goals: Costing and Financial Gap Analysis. Background Document for the Taskforce on Innovative International Financing for Health Systems.

1 Long-Run Effects of Early-Life Malnutrition: Evidence from the Greek Famine

1.1 Introduction[1]

Several hundred million of today's children under five suffer from hunger (Victoria et al. 2008). Recent research indicates that such early-life malnutrition is not only associated with impairments of adult health but also with worse educational and labor market outcomes (Walker et al. 2007). A causal interpretation of these associations is, however, complicated by the fact that early-life malnutrition is typically not independent of other determinants of long-run outcomes like parental health, wealth, and cognitive skill. Failure to control for these background variables can thus cause bias in the estimates of early-life malnutrition effects. In an extreme case, the entire malnutrition effect estimate may in reality be driven by omitted variables.

To address this issue, scholars have used famines affecting entire regions for limited periods of time as instruments for individual-level early-life malnutrition. The rationale for this is that because famines affect entire cohorts rather than selected individuals, there is limited scope for unobserved heterogeneity between the individuals with early-life famine exposure that form the treatment group and the individuals without that form the control group.

The 1944-45 Dutch famine and the 1959-61 Chinese famine are the most prominent cases for famine cohort studies. This chapter adds estimates of long-run effects for the Greek famine of 1941-42 to this literature.[2,3]

Examining the Greek famine has several advantages. Relative to the three-year Chinese famine, it was short in duration (6-8 months). This limits the scope for behavioral responses, in particular in terms of fertility decisions that may otherwise cause unobserved heterogeneity between the treatment and control cohorts. Because child mortality in the famine was in addition relatively low, it can be assumed that the Greek cohorts with famine exposure in crucial phases of development are rather similar to the cohorts without. More formally, sample selection issues are likely less severe for the Greek than for other famines.

An additional advantage of the famine's short duration is that it allows me to identify effects of malnutrition at specific ages, that is, of malnutrition during the second year of life, during the first year of life, and in the fetal stage. This is impossible in multiyear famines like that of China. For instance, the cohorts with first year of life exposure to the

[1] This chapter bases on joint work with Thomas Stratmann (Neelsen and Stratmann 2011).

[2] The Greek famine has received little attention in the famine and other academic literatures. Moreover, Hionidou (2006, p. 30-31) observes "a complete absence of reference to the food crisis of the occupation years in the collective memory of the Greek population, no collective of even official memory of the famine, let alone collective trauma such as that relating to the Irish famine."

[3] Except for Valaoras' (1946) analysis of short-run somatometric famine impacts on Athenian children (height, weight, weight-for-height) there are no quantitative studies on famine effects for Greece to date.

1959-61 Chinese famine were in addition either exposed as fetuses (born 1961) or during the second year of life (born pre-1961).

Finally, unlike many more recent famines, Greece's did not coincide with severe outbreaks of infectious disease (Hionidou 2006). This is important because early-life disease exposure has its own long-run consequences (Walker et al. 2007). The cohort-membership instrument cannot distinguish these from the long-run consequences of malnutrition if both occur simultaneously.

My results indicate that malnutrition in the fetal stage, during infancy, and in the second year of life impairs the development of human capital: early-life famine exposure is associated with lower rates of literacy, upper secondary schooling, fewer years of education, and lower job status in 1971-2001 Greek census data. The inclusion of birthplace controls in my empirical model reduces my estimates of long-run effects for all three age groups and the effects for the fetally exposed cohort are no longer statistically significant. The latter finding indicates substantial birthplace selection in this cohort. Finally, in accordance with contemporary sources that suggest that the famine hit hardest in urban areas, I find larger long-run effects on urban- than on rural-born cohorts.

The rest of this chapter is organized as follows. The next section reviews earlier work on later-life health and socioeconomic consequences of early-life health shocks. Section 1.3 describes the Greek 1941-42 famine and Section 1.4 my dataset and empirical strategy. Section 1.5 presents results and Section 1.6 concludes.

1.2 Early-life Malnutrition and Later-Life Outcomes

Barker's *fetal origins* hypothesis suggests a causal relationship between fetal conditions and later-life health (e.g. Barker 1998). Empirical tests of this hypothesis often use birth-weight to proxy fetal conditions. One of the most consistent results in this literature is an association of low birth weight with a higher risk of adult chronic disease, especially after the primary reproductive age. For the underlying mechanism, scholars suspect that in order to increase the chance of survival in the face of fetal malnutrition, blood and nutrients are diverted to the brain rather than other vital organs. This impairs the development of these organs or permanently adapts them to nutritional deprivation, the latter creating so called *thrifty phenotypes* (Hales and Barker 1992, Barker and Hanson 2004). Under improved nutritional conditions later in life, the thrifty phenotypes face a higher chance of metabolic disorders such as type-II diabetes (Hales 1997).[4] Fetal malnutrition has also been shown to impair cardiac health (Hoet and Hanson 1999) and kidney functioning (Brenner and Chartow 1994). Moreover, Case et al. (2002) and Almond et al. (2005) show that the negative long-run effects of poor fetal development do

[4] Stocker et al. (2005) argue that metabolic disorders are relatively uncommon in African countries because fetal malnutrition is typically followed by nutritional scarcity after birth.

not stop at health, but can translate into lower educational attainment, employment and income.[5]

Individuals that experience severe malnutrition after birth equally suffer long-run health impairments. In their review of studies from Asian, African, and South American countries, Walker et al. (2007) and Victoria et al. (2008) show that, like for malnourished fetuses, malnutrition of children in the first 24 months after birth is associated with lower adult height, higher blood glucose concentrations, increased blood pressure, harmful lipid profiles, deficits in cognitive skills, and a higher likelihood of mental illness. Also like for malnourished fetuses, the negative effects from severe malnutrition early after birth can translate into inferior socioeconomic outcomes, such as lower educational attainment and reduced income (Walker et al. 2007). This transmission can occur by both direct and indirect pathways (Case and Paxson 2006). On the direct pathway, early-life malnutrition impairs cognitive development which reduces school success and, subsequently, labor market outcomes. On the indirect pathway, early-life malnutrition translates into poor child health which reduces school attendance and attainment with adverse consequences for labor market performance.

The abovementioned studies use individual-level variation in fetal and postnatal malnutrition to identify long-run effects. One issue with this approach is the danger of omitted variables bias (OVB). If early-life malnutrition correlates with determinants of long-run outcomes that remain unobserved and thus not accounted for in the estimation of long-run effects, the estimates contain bias and a causal interpretation is no longer warranted. Typical candidates for omitted variables are the health, cognitive skill, and wealth of parents and their parenting practices.

Twin studies are one approach to address this issue. Because (monozygotic) twins share the same genetic endowments and, given that they grow up together, similar family environments, the differencing out of twin fixed effects rules out OVB from these factors. While not entirely unambiguous, the sum of evidence from twin studies supports the hypothesis that low birth-weight has long-run health and socioeconomic consequences (Behrman and Rosenzweig 2004, Almond et al. 2005, Oreopoulos et al. 2008, Black et al. 2007).

Besides concerns about the transferability of findings from the twin to the non-twin population, a limitation of the approach is that sufficiently large twin datasets only exist for a small number of developed countries. Hence, no twin studies can be undertaken for countries where early-life malnutrition is most prevalent today.

An alternative approach uses famines as instruments for individual-level early-life malnutrition. The rationale for this is that because famines affect entire cohorts rather than selected individuals, there is limited scope for unobserved heterogeneity between individuals with early-life famine exposure (the treated cohorts) and individuals without

[5] The evidence is, however, not unambiguous. Maccini and Yang's (2009), for example, find no long-run effects of poor fetal conditions on adult health, education and socioeconomic outcomes for Indonesian women when using rainfall in the year before birth as an instrument.

(the control cohorts). In these studies, the treatment group consists of cohorts that lived in famine-affected areas early in life. The control group can include cohorts from the affected areas that were older during the famine, cohorts that were conceived in the affected areas after the famine, and cohorts of the same age as the treated cohorts who lived in non-famine-affected areas early in life.

The famine cohort studies have produced mixed evidence on long-run effects of fetal famine exposure. St. Clair et al. (2005) find that fetal exposure to the Chinese famine of 1959-61 is associated with higher levels of adult schizophrenia. For the same famine, Luo et al. (2006) show associations with higher rates of female obesity, and Meng and Qian (2009) adverse effects on adult height, but none on coronary or metabolic conditions. Almond et al. (2007) provide evidence that the male Chinese cohorts with fetal famine exposure are less likely to be literate, to work, and to be married.

Stanner et al. (1997) and Stanner and Yudkin (2001) examine long-run consequences of fetal exposure the German siege of Leningrad (1941-44) which led to severe starvation of the city's inhabitants. They do not find effects on adult metabolic or cardiac conditions. Similarly, Kannisto et al. (1997) find no negative effect on longevity for cohorts conceived during or shortly before the Finnish famine of 1866-68. A possible explanation for the absence of long-run effects in these studies is the length and severity of the Leningrad and Finnish famines. Both lasting three years, the former killed up to one-third and the latter eight percent of the population. Against this background, the survivors of fetal famine exposure likely form a strongly positively selected group, i.e. they (the treatment group) can be assumed to have substantially better unobserved characteristics than the cohorts without fetal famine exposure (the control group). In this scenario, the famine effect estimates understate the true famine effects.

The scope for such selection was smaller in the shorter and less deadly Dutch famine of 1944-45 in which nutritional crisis was caused by a six month Nazi blockade of the western Netherlands. In a series of articles, Roseboom et al. (e.g. 1999, 2000a, 2000b, 2001) present evidence that fetal exposure to the famine is associated with impairments of the central nervous system, worse self-reported health, and coronary heart disease. Furthermore, Neugebauer et al. (1999) find a higher prevalence of adult antisocial personality disorders and Ravelli et al. (1998, 1999) increased glucose resistance and higher rates of obesity for the fetally exposed cohorts.

In contrast to the large number of studies on fetal famine exposure, only a few have examined long-run consequences of famine exposure after birth. Exceptions are Meng and Qian (2009), Chen and Zhou (2007) and Gørgens et al. (2007) who find negative effects on height, weight, weight-for-height, education, labor supply, income, and housing space for children with postnatal exposure to the Chinese famine. With empirical approaches similar to those of the famine cohort studies, Alderman et al. (2006) find that postnatal exposure to drought and civil war lowers adult height and the number of grades of schooling completed in a Zimbabwean sample, and Maccini and Yang (2009) show that

higher rainfall in the year after birth improves the health and increases the height, educational attainment and wealth of female Indonesian adults.

1.3 The Greek 1941-42 Famine

On April 30th 1941, only twenty-four days after Nazi Germany had joined the invasion of Greece by Italian forces, open warfare ended with the Greeks' unconditional surrender and the country's occupation by German, Italian, and Bulgarian troops.[6] The Allied forces responded with a full naval blockade, cutting off all imports to Greece, including foods.

Immediately following victory, the occupying forces divided the country into 13 zones between which any movement of goods and people was strictly prohibited. Also within the zones, the confiscation of fuels and all means of transportation including fishing boats and pack animals reduced mobility to a minimum. The occupiers seized strategic industries, and appropriated or bought all stocks of commodities like tobacco, olive oil, cotton, and leather and transferred them to their home countries.

The occupying forces ordered a newly installed central government to reorganize the food supply to the Greek civil population. Farmers had to pay a 10 percent in kind tax on their produce and sell to the government at fixed prices all production above the government-determined subsistence level. Moreover, the food price controls and rationing that had been in place before the Greek defeat were now tightened. With the low government prices and newly imposed taxes, farmers went to great lengths to hide their produce from the officials and traders pulled their merchandise from the shelves. Additionally, the naval blockade and warfare in surrounding countries severed the foreign trade routes on which Greece traditionally depended for food imports.

The nutritional situation became critical in the summer of 1941 and in the fall turned into a famine. In the Greater Athens area, the calorific value of rations and food provided by public or charity soup kitchens deteriorated from 600 calories per day per person in July of 1941 to 320 in November of 1941. In many places, civil registration records were discontinued during the occupation (Valaoras 1960). Where they were not, the data suggest mortality increases between 300 and 1000 percent in the winter of 1941-42 compared to pre-war years. Estimates of a country-wide death toll of the famine vary between 100,000 and 200,000 (Hionidou 2006) or 1.4 to 2.8 percent of the Greek population, the large majority of which occurred between October 1941 and March 1942 (Helger 1949).

Not all parts of Greece experienced equal levels of food scarcity. The available evidence indicates that with the severe movement restrictions, the proximity to agricultural production and the level of urbanization became crucial determinants of

[6] The following account is based on Hionidou (2006) who provides a detailed history of the famine from a socio-historic perspective and on Mazower (1993) who focuses on the political and military implications of the 1941-44 occupation period.

famine severity: while certain isolated islands and mountain villages suffered greatly, the urban population was most affected by the famine.[7]

Because of the efforts of the Greek diasporas in the US and Britain, the situation of the Greek civilian population soon became widely known in the Allied countries, and increasing public pressure led to the lifting of the naval blockade in February 1942. Wheat shipments soon began and together with the rising temperatures of springtime, this brought down mortality rates. The international relief focused mainly on children. In February 1942, the Red Cross started to provide daily milk rations[8], medical services and clothing to Athenian children under three. From March 1942 onwards, pregnant women and breastfeeding mothers received extra food supplies which were further increased for women that were temporarily unable to breastfeed because of malnutrition.

Also in March 1942, the occupiers and Allied forces agreed to the establishment of the Swedish-run Joint Relief Commission to reorganize the public food supply system. The occupiers moreover committed to replace all appropriated agricultural produce with food imports of equal calorific value[9] and relaxed the harshest mobility restrictions and price regulations. As a result, fresh produce from the June harvest and foods that had hitherto been hoarded entered the markets and food shipments further grew in volume. Towards the end of 1942, the nutritional situation had returned to acceptable levels in most parts of the country (Hionidou 2006).

1.4 Data and Empirical Approach

My data come from the 1971, 1981, 1991, and 2001 Greek National Population Housing Censuses from the IPUMS website (Minnesota Population Center 2009). Each wave represents a 10 percent sample of the Greek population. In my samples, I include the 11 cohorts born between 1936 and 1946.[10] The sample individuals are therefore 25-36 years old in the 1971 wave, 35-46 years old in the 1981 wave, 45-56 years old in the 1991 wave, and 55-66 years old in the 2001 wave.

The medical literature suggests that early-life malnutrition is particularly harmful to long-run outcomes in the period from conception to the first 24 months after birth (Bryce et al. 2008). Therefore, my treatment group consists of the 1940 birth cohort of which the majority was in the second year of life when the 1941-42 famine struck, the 1941 birth cohort the majority of which experienced the famine in their first year of life, and the 1942 birth cohort of which the majority experienced it as fetuses.[11,12] In the following, I

[7] Hionidou (2006) estimates that mortality increases in urban areas were on average twice as large as in the countryside.

[8] The form in which milk rations were provided required immediate consumption. This method improved hygiene in the preparation of the rations and assured that the milk actually reached young children instead of being bartered for other goods (Helger 1949).

[9] According to sources cited by Hionidou (2006) this agreement was mostly adhered to.

[10] The large number of observations in the Greek censuses permits me to use a rather short 11 cohort sample. By limiting the sample to 11 cohorts, I seek to increase homogeneity in unobserved factors across birth cohorts.

[11] The Greek census does not contain month of birth information. With the 1940-42 birth cohorts I therefore consider as my treatment group those who experienced the most severe episode of nutritional deprivation

refer to those exposed in the second year of life as one-year-olds, to those exposed in the first year of life as infants, and to those exposed in the fetal stage as fetuses.

The cohorts in my control group did not experience comparable levels of malnutrition during these crucial phases of development: the cohorts born before 1940 were exposed to the famine at older ages and those born in 1943 or later had no direct famine exposure.

I exploit this discontinuous change in the likelihood of early-life malnutrition exposure to estimates cohort-level long-run effects by the model

$$y_i = cons + \beta_1 1940 + \beta_2 1941 + \beta_3 1942 + sex_i + yob_i + yob_i^2 + \varepsilon_i \ . \quad (1)$$

The dependent variable y_i represents a set of educational outcomes and, for the working subsample, a measure of socioeconomic status for person *i*. *cons* is a constant and *1940*, *1941* and *1942* indicators that equal 1 if the individual was born in the respective year and 0 otherwise. The variable *sex* is a gender dummy. *yob* denotes the year of birth, thereby controlling for linear trends in the outcome variables *y*. Moreover, to account for non-linearities in the outcome trends I include the squared year of birth yob^2. In this model, the coefficients β_1, β_2 and β_3 represent the 1940-42 cohorts' departures from a secular 1936-46 outcome trend that I interpret as the cohort-level famine effects.

The first dependent variable in the vector y_i is an indicator whether an individual is literate.[13] The second dependent variable indicates whether an individual has completed upper secondary or technical school. In the following, I refer to this variable as upper secondary schooling. I compute this variable using the census data on the highest level of education.[14] The third dependent variable is the total years of education that I compute using the regular time required to gain the highest level of education obtained.[15,16]

While the Greek census does not report income or wages, it has detailed occupation data that permit me to approximate a working individual's socioeconomic status as my

(October 1941 - March 1942) as fetuses or up to their 27th month of life. Specifically, the oldest members of the 1940 birth cohort (born in January) were exposed during the 22nd and 27th month of life and the youngest (born in December) during the 11th and 16th month of life. Accordingly, the 1941 cohort received treatment between the 10th and 15th month of life (born in January) and the last three months of gestation up to the 4th month after birth (born in December). Finally, for the 1942 birth cohort, those born in January experienced the famine during the last 4 months before birth and the first three months after birth. Those born later than March 1942, i.e. the majority of the cohort, were exposed to malnutrition only as fetuses.

[12] The treated cohorts were thus between 29 and 31 years of age in 1971, between 39 and 41 in 1981, between 49 and 51 in 1991, and between 59 and 61 in 2001.

[13] The method to determine literacy status is not the same across the four census waves. While the 1971 census obtained it through direct questioning, the 1981-2001 waves do not contain a literacy question but instead consider individuals literate if they ever attended school.

[14] Until 1975, compulsory schooling in Greece comprised of six years of primary school. Optional higher schooling tracks included three years of lower secondary (gymnasium) and three years of upper secondary schooling (lyceum).

[15] The census's highest educational degree variable is more detailed in later census waves. I exploit this in my construction of the years of education variable. As a result, the estimates for this variable are not necessarily comparable across the four census waves.

[16] My approximation of years required to complete Greece's different education levels comes from European Commission (2009) and is available on request.

final dependent variable. The over 400 occupational categories in the census are in many cases equal to those of the 4-digit ISCO88 scheme. I map these categories into socioeconomic status scores that base on Ganzeboom et al.'s (1992) International Socio-Economic Index of Occupational Status (ISEI). The index uses an occupation's educational requirements and income to assign a score that lies between 16 for occupations with the lowest, and 90 for occupations with the highest socioeconomic status. My coding of occupations follows that by Ganzeboom and Treiman (1996). For the occupational categories in the Greek censuses that do not precisely match the 4-digit ISCO88 codes, I apply the coding for the more aggregated 3-digit or 2-digit ISCO88 categories.[17]

The estimates I obtain from model (1) are unbiased if the error term ε is uncorrelated with the independent variables, including the 1940-42 cohort dummies that form the treatment. This requires that the unobserved outcome determinants captured in ε are equally distributed across the 1936-46 birth cohorts, that is, that membership in the treated cohorts is random with regards to ε. This assumption is, however, unlikely warranted due to non-random mortality and conception during the famine.

With respect to famine mortality, I suspect that individuals with adverse genetic endowments and worse initial health had a lesser chance of survival in the treated 1940-42 cohorts than in the surrounding control cohorts. If such selection into survival in fact occurred more in the treatment than in the control group, and if bad genes and initial health correlate with adverse educational and labor market outcomes, the surviving members of the treated cohorts would on average have better such unobserved endowments than the control cohorts. As a result, OVB would bias the coefficients β_1, β_2 and β_3 towards zero, causing an understatement of the true magnitude of adverse early-life famine exposure effects. The same logic would apply if selective famine mortality led to an overrepresentation of individuals with poor and poorly-educated parents in the treated relative to the control cohorts.

With respect to fertility, I assume positive selection in the famine-conceived 1942 cohort. Valaoras (1946) estimates that during the famine, up to seventy percent of females in the Greater Athens area stopped menstruating because of severe malnutrition (famine amenorrhea). It is likely that such lost births were more frequent among women of low

[17] The code for conversion to the ISEI-scores is available on request. Because the 2001 census used a different job classification method than the prior census waves, estimates of ISEI-score-effects are not necessarily comparable between the 2001 and the 1971-1991 samples in Table 1.2.

Table 1.1: Variable means for 1940-42 cohorts and the two surrounding cohorts

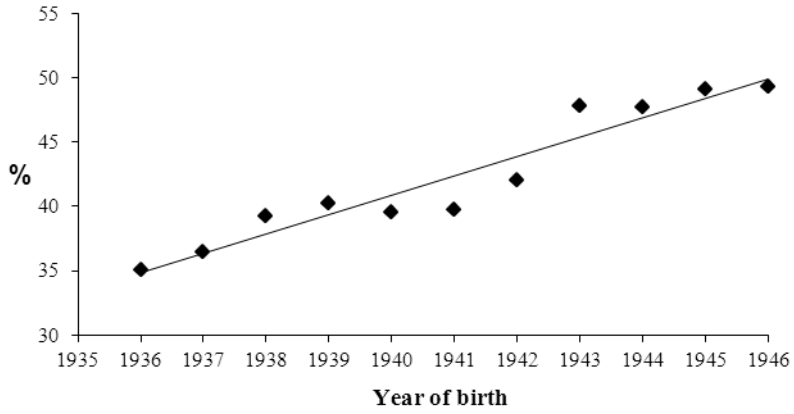
	1971		1981		1991		2001	
Cohorts	1940-42	Surr.	1940-42	Surr.	1940-42	Surr.	1940-42	Surr.
Literate	.950	.950	.958	.962	.953	.954	.961	.961
SE	[.219]	[.219]	[.200]	[.191]	[.210]	[.208]	[.1943]	[.192]
N	29,419	19,616	33,114	21,134	32,998	20,850	32,059	20,467
Upper sec. school.	.218	.237	.233	.253	.255	.273	.263	.285
SE	[.413]	[.425]	[.422]	[.434]	[.436]	[.445]	[.4403]	[.451]
N	28,999	19,342	33,114	21,134	32,998	20,850	32,059	20,467
Years of education	7.244	7.380	7.405	7.616	7.697	7.838	7.735	7.919
SE	[3.654]	[3.720]	[4.040]	[4.126]	[4.246]	[4.321]	[4.3239]	[4.443]
N	28,999	19,342	33,114	21,134	32,998	20,850	32,059	20,467
ISEI-score	34.501	34.763	36.895	37.430	37.170	37.242	36.454	37.637
SE	[14.646]	[14.523]	[15.525]	[15.823]	[16.209]	[16.343]	[16.703]	[17.231]
N	17,314	11,542	20,431	13,467	18,979	11,940	10,812	6,647
Sex	.521	.519	.515	.497	.507	.499	.527	.517
SE	[.450]	[.4950]	[.500]	[.500]	[.500]	[.500]	[.499]	[.499]
N	29,452	19,630	33,114	21,134	33,009	20,857	32,059	20,467
Urban born							.296	.315
SE							[.456]	[.464]
N							31,299	19,935

My samples form subsamples of the 1971-2001 Greek censuses in that I only include Greek citizens in my calculations. The surrounding cohorts are 1939 (famine exposure in the third year of life) and 1943 (conceived after the famine, no direct famine exposure). Standard errors are reported in brackets.

socioeconomic status and bad health.[18] Combining the effects of selective mortality and fertility, I therefore suspect that OVB-related downward bias in the famine effects estimates is largest for the 1942 famine-conceived cohort.

[18] An alternative hypothesis regarding selection into fertility during the famine is that couples with better socioeconomic backgrounds adapted their fertility decisions more quickly to the new nutritional situation than couples with worse backgrounds. This would lead to an overrepresentation of individuals with inferior parental backgrounds in the 1942 cohort and failure to control for such selection would cause upward OVB. However, evidence by Valaoras (1946) indicates that the fertility reductions in most cases came through hunger-related physical impairments to conceive rather than deliberate family planning decisions. As food deprivation was more prevalent among the disadvantaged, I consider an upward bias from selective fertility in the 1942 cohort unlikely.

Figure 1.1: Percent of 1936-46 Greece-born cohorts with upper secondary schooling and linear trend (2001 census)



However, OVB in model (1) may not unambiguously bias my famine effects estimates downward. If, for instance, famine mortality was low in places with poor educational and labor market prospects, individuals from these disadvantaged areas would be overrepresented in the treated 1940-42 cohorts. Not controlling for such selection in model (1) would cause upward bias in the famine effect estimates. Because the 2001 census provides detailed birthplace information I can address this issue by including 52 birth prefecture dummy variables in my model.[19] These birth prefecture *fixed effects* control for all determinants of education and labor market outcomes that are specific to the birth prefecture and constant across the 1936-46 period. Among other things, I thereby rule out OVB from differences in the birthplace distribution between the treatment and control cohorts. The inclusion of birthplace controls, however, bears the danger of over-controlling. The inclusion of the fixed effects removes variation in outcomes between birth prefectures. To the degree that these between-prefecture differences are caused by differences in famine severity, the birthplace controls remove more variation than is intended and the estimates form lower bounds of the true famine effects.

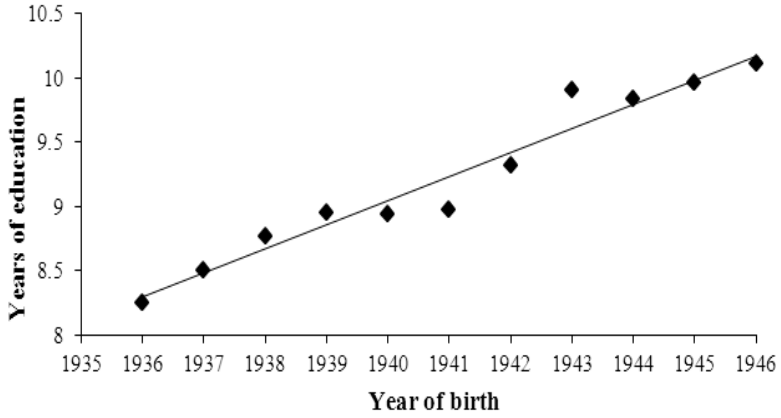
1.5 Results

Descriptive Statistics (Table 1.1)

Table 1.1 presents descriptive statistics for the Greek citizens in the 1971, 1981, 1991, and 2001 census waves. For each wave, the table provides variable means for the 1940-42 cohorts that form the treatment group and for the surrounding 1939 and 1943 cohorts that form part of the control group. The results show that there are no large differences in literacy rates between the 1940-42 and the surrounding cohorts. However,

[19] In 2005, the average prefecture population was 200,000, ranging from the most populous Athens prefecture with almost 2,700,000 million to 23,000 in the Lefkada prefecture. My results are robust to the inclusion of 156 sub-prefecture level birthplace controls.

Figure 1.2: Average years of education in the 1936-46 Greece-born cohorts and linear trend (2001 census)



for upper secondary schooling, the 1940-42 cohorts have a 2 percentage point lower completion rate across all census waves. Further, the 1940-42 cohorts have between 1.4 and 2.4 months less education and with respect to the ISEI-score, jobs with lower socioeconomic status than the surrounding cohorts.

Table 1.1 moreover shows that across all census waves, the share of females is higher in the treated than in the two control cohorts, with the differences ranging between 0.2 and 1.8 percentage points. These differences likely reflect that, like in other famines, male children were more likely to die during the Greek famine (Helger 1949, Jakobovits 1991).

Finally, the last row of Table 1.1 shows that in the treated cohorts, the share of individuals from urban areas is lower than in the two surrounding cohorts. This is consistent with evidence that the famine was more severe in urban than in rural areas (Helger 1949).[20]

Figures 1.1 and 1.2 plot the share of individuals with upper secondary schooling and the average years of education for each of the 1936-46 cohorts in the 2001 census. Both figures include a linear outcome trend line that shows an upward trend in both measures of educational attainment. Outcomes for the treated 1940-42 birth cohorts, however, lie below this trend, providing initial descriptive evidence for harmful long-run effects of early-life famine exposure.

Basic Regressions (Table 1.2)

For my basic model (1) Table 1.2 shows the 1940-42 cohort effect estimates for each of the four 1971-2001 census waves. For the 1940 cohort, that was exposed to the famine as one-year-olds, I find 0.6 and 0.5 percentage point declines in literacy in the 1981 and 1991 censuses, respectively. The reductions in upper secondary schooling range

[20] When I regress the urban birthplace indicator on the explanatory variables in model (1), I find that the likelihood of being born in an urban area is significantly reduced by 0.9 percentage points for the 1940 cohort, by 2.2 percentage points for the 1941 cohort, and by 4.4 percentage points for the 1942 cohorts.

Table 1.2: OLS estimates of departures from 1936-46 cohort trend for the 1940-42 cohorts; 1971, 1981, 1991 and 2001 census waves

		(1) 1971 census		(2) 1981 census		(3) 1991 census		(4) 2001 census	
	Cohort (exp. in)	Coeff.	SE	Coeff.	SE	Coeff.	SE	Coeff.	SE
Literate	1940 (1-year-old)	-.002	[.002]	-.006**	[.002]	-.005**	[.002]	.000	[.002]
	1941 (infant)	-.004*	[.002]	-.011***	[.003]	-.007**	[.002]	-.007***	[.002]
	1942 (fetus)	.001	[.002]	-.003	[.002]	-.001	[.002]	.002	[.002]
N		119,368		130,321		129,372		126,214	
Upper sec. schooling	1940 (1-year-old)	-.016***	[.004]	-.013***	[.004]	-.013***	[.003]	-.018***	[.004]
	1941 (infant)	-.022***	[.005]	-.024***	[.005]	-.018***	[.004]	-.023***	[.004]
	1942 (fetus)	-.026***	[.005]	-.023***	[.004]	-.021***	[.004]	-.023***	[.004]
N		117,638		130,321		129,372		126,214	
Years of education	1940 (1-year-old)	-.147***	[.037]	-.177***	[.048]	-.159***	[.045]	-.148***	[.043]
	1941 (infant)	-.215***	[.043]	-.293***	[.055]	-.173***	[.053]	-.235***	[.050]
	1942 (fetus)	-.181***	[.041]	-.200***	[.052]	-.157***	[.051]	-.173***	[.048]
N		117,638		130,321		129,372		126,214	
ISEI-score	1940 (1-year-old)	-.503**	[.215]	-.700***	[.211]	-.296	[.229]	-.456	[.332]
	1941 (infant)	-.521**	[.242]	-.533**	[.245]	-.235	[.263]	-.839**	[.348]
	1942 (fetus)	-.478**	[.235]	-.583**	[.241]	-.279	[.253]	-1.111***	[.324]
N		69,998		80,993		73,852		43,177	

The results in Table 1.2 are for the subsamples of individuals in the respective census waves that are Greek citizens. In addition to the indicators for whether born in 1940, 1941, or 1942, all specifications include year of birth, year of birth squared and a sex indicator. Huber-White robust standard errors are reported in brackets.

between 1.3 percentage points in the 1981 and 1991 censuses, and 1.8 percentage points in 2001. Years of education are also fewer for one-year-olds with reductions ranging between 1.2 and 1.4months across all four census waves.[21] Finally, the ISEI-score reduces by 0.5 points (1.5 percent of the sample average) in the 1971 census and by 0.7 points (1.9 percent) in the 1981 census. The ISEI reductions in the 1991 and 2001 censuses are not statistically significant.[22]

The 1941 birth cohort with famine exposure during infancy appears to be the most famine-affected in the long-run. It is the only cohort that shows significant reductions in the share of literate individuals for all four census waves, ranging from 0.4 percentage points in 1971 to 1.1 percentage points in 1981.[23] The negative effects on upper secondary

[21] As I mention above, the method by which I obtain the years-of-education variable differs between census waves. This limits the comparability of coefficient magnitudes across census waves. The estimates, however, change only slightly when I use harmonized years-of-education variables. Results are available on request.

[22] As I mention above, the 2001 census uses a different occupation coding, limiting the comparability of ISEI-score effect estimates between the 1971-91 waves and the 2001 wave.

[23] The large increase in the literacy effect between the 1971 and 1981 census waves may in part be the result of the change in the method of obtaining the variable that I discuss in Section 1.4.

schooling are statistically significant in all censuses, ranging between 1.8 percentage points in 1991 and 2.4 percentage points in 1981. A similar picture emerges for years of education where being born in 1941 is associated with a reduction between 1.4 and 2.4 months in 1991 and 1981, respectively. I also find reductions in the ISEI-score that are statistically significant in the 1971, 1981 and 2001 census waves. They range between 0.5 and 0.8 points, or between 1.4 and 2.5 percent of the sample averages.

In contrast to one-year-olds and infants, the 1942 cohort treated as fetuses does not show significant reductions in literacy.[24] Instead, in two of the four censuses, the point estimates have positive signs while not being statistically significant. I do, however, find statistically significant reductions in upper secondary schooling for fetuses in all censuses. They range between 2.1 percentage points in the 1991 wave and 2.6 percentage points in the 1971 wave. There are also reductions in years of education for fetuses in all waves that range between 1.3 and 1.7 months.[25] With respect to the ISEI-score, I find statistically significant reductions in the 1971, 1981 and 2001 waves. The magnitude of these effects lies between 0.5 points in the 1981 and 1991 waves (1.4 and 1.5 percent) and 1.1 points (3 percent) in the 2001 wave.

Summarizing the results in Table 1.2, I find support for the hypothesis that early-life exposure to famine worsens long-run outcomes. There are statically significant reductions in educational attainment, measured either as literacy, upper secondary schooling, or years of education for one or more of the treated cohorts in all census waves. When comparing the magnitudes of effects between cohorts with famine exposure at different ages, I find that the adverse education effects are largest for infants. For fetuses,the negative effect on upper secondary schooling is similar to that for infants while the reductions in years of education are slightly smaller.

Regressions with Birthplace Controls (Table 1.3)

The estimates in Table 1.2 may overstate the true famine effects because of an overrepresentation of individuals with inferior birthplaces in the treated compared to the control cohorts. As discussed above, I can limit the risk of such bias by including prefecture of birth indicators in model (1). Table 1.3 presents results for such specifications for the 2001 census wave for which birthplace information is available. In contrast to Table 1.2, the results in Table 1.3 are not for all Greek citizens but for the subsample that the 2001 census identifies as Greece-born.

For reference, the first column of Table 1.3 shows results for the basic model (1). Column 2 adds 52 dummy variables for the prefecture of birth. In column 3, I, in addition, include an indicator for whether a person was born in an urban area and interaction terms between the urban birthplace and birth prefecture indicators. Because of these interaction terms, this last specification not only controls for time-invariant prefecture and urban

[24] This finding contrasts with evidence from the Chinese Famine: Almond et al. (2007) find that the male (female) cohorts with famine exposure during the fetal stage are 9 (7) percent more likely to be illiterate.
[25] For the Chinese Famine, Meng and Qian (2009) find an 8.6 percent reduction in years of education for the fetally exposed cohorts but no effect on the cohorts exposed after birth.

Table 1.3: OLS estimates of departures from 1936-46 cohort trend for the 1940-42 cohorts; with and without birthplace controls; 2001 census

		(1) Basic		(2) Prefecture FE		(3) Prefecture & urban FE	
	Cohort (exp. in)	Coefficient	SE	Coefficient	SE	Coefficient	SE
Literate	1940 (1-year-old)	-.001	[.003]	.000	[.002]	-.001	[.002]
	1941 (infant)	-.007**	[.003]	-.007***	[.002]	-.007***	[.002]
	1942 (fetus)	.002	[.003]	.004*	[.002]	.004*	[.002]
N		123,793		123,793		123,793	
Upper sec. schooling	1940 (1-year-old)	-.018***	[.005]	-.012***	[.004]	-.012***	[.004]
	1941 (infant)	-.022***	[.006]	-.016**	[.007]	-.013**	[.006]
	1942 (fetus)	-.023***	[.006]	-.006	[.008]	-.003	[.006]
N		123,793		123,793		123,793	
Years of education	1940 (1-year-old)	-.149***	[.046]	-.092**	[.043]	-.096**	[.038]
	1941 (infant)	-.232***	[.052]	-.172**	[.065]	-.146**	[.056]
	1942 (fetus)	-.184***	[.050]	-.017	[.064]	.008	[.052]
N		123,793		123,793		123,793	
ISEI-score	1940 (1-year-old)	-.501	[.335]	-.282	[.329]	-.194	[.311]
	1941 (infant)	-.806**	[.354]	-.563	[.526]	-.496	[.486]
	1942 (fetus)	-1.145***	[.328]	-.493	[.283]	-.292	[.282]
N		42,285		42,285		42,285	

All specifications in Table 1.3 use data from Greece-born individuals in the 2001 census. In addition to the indicators for whether born in 1940, 1941, or 1942, all specifications include year of birth, year of birth squared and a sex indicator. Column 1 shows results for my basic model (1) for Greece born individuals. Column 2 adds 52 prefecture of birth dummies to the specification in column 1. Relative to the specification in column 2, column 3 adds an urban birthplace indicator and interactions of the 52 prefecture of birth dummies with this urban birthplace indicator. For column 1, Huber-White robust standard errors are reported in brackets. For columns 2 and 3, I report robust standard errors clustered at the prefecture of birth level in brackets.

birthplace effects, but also for prefecture-specific consequences of being born in an urban area. For all models that include birthplace controls, I cluster standard errors at the prefecture level.

For the 1940 cohort of one-year-olds, the famine effects on literacy are neither statistically significant in column 1, nor in columns 2 and 3 which include birthplace indicators. The estimate of the famine's impact on upper secondary schooling drops from minus 1.8 percentage points in the basic specification in column 1 to minus 1.2 percentage points in columns 2 and 3. Similarly, the effect on years of education falls from 1.2 months to about 0.8 months after the inclusion of birthplace indicators. The estimated

impact on socioeconomic status in column 1 also reduces in columns 2 and 3 and continues to not be statistically significant.

For the 1941 cohort exposed during infancy, I find a negative and statistically significant effect on literacy regardless of whether I include birthplace controls. The estimated effects on upper secondary schooling and years of education are smaller in columns 2 and 3 than in column 1, but remain statistically significant. Similar as for one year-olds, the coefficients for socioeconomic status reduce for infants when including birthplace indicators and the point estimates are no longer statistically significant.

Finally, while the direction of the 1942 cohort effect estimates on upper secondary schooling and on years of education in columns 2 and 3 point to negative famine effects, none of the coefficients is precisely estimated. Further, in columns 2 and 3, fetal exposure is associated with statistically significant 0.4 percentage point increases in the likelihood to be literate. Unlike in the specification without birthplace controls, the reductions in the ISEI-score for the 1942 cohort are small and not statistically significant in columns 2 and 3.

The reduction in the famine effect estimates after the inclusion of birth prefecture controls is consistent with the hypothesis that birthplaces associated with inferior later-life prospects are overrepresented in the 1940-42 cohorts. Moreover, the fact that the reductions are largest for the 1942 cohort is consistent with this cohort experiencing additional negative birthplace selection through fertility.

However, as discussed above, the inclusion of birth prefecture fixed effects may remove more variation in the depended variables than is intended, leading to too large reductions in the famine effects. In addition, positive selection into survival with regards to unobserved background variables like parental health, wealth, and cognitive skill likely causes additional downward bias in my famine effect estimates. Therefore, while the estimates in Table 1.3 likely identify lower bounds for the true famine effects, the absence of effects on fetuses does not necessarily provide evidence against Barker's hypothesis.

Regressions for Urban, Rural, and Foreign Born Subsamples (Table 1.4)

Because famine severity differed between urban and rural areas, it is likely that any long-run effects differ between the urban- and rural-born cohorts. I account for this possibility by estimating separate models for the urban- and rural-born subsamples in the 2001 census. This permits me to distinguish the long-run effects of severe famine from those of milder forms.

For comparison purposes, Table 1.4, column 1 replicates the results for the 2001 sample of individuals born in Greece that I also show in column 2 of Table 1.3. Table 1.4, column 2 shows estimates for the urban-born subsample, and column 3 for the rural-born subsample.

For one-year-olds, the results in columns 2 and 3 neither indicate a famine effect on literacy for urban-born nor for rural-born individuals. In contrast, the reduction in upper secondary schooling is 2.1 percentage points (5 percent of the sample mean) in the

Table 1.4: OLS estimates of departures from 1936-46 cohort trend for the 1940-42 cohorts; subsamples with different birthplace; 2001 census

		(1) Greece-born		(2) Urban-born		(3) Rural-born		(4) Foreign-born	
	Cohort (exp. in)	Coeff.	SE	Coeff.	SE	Coeff.	SE	Coeff.	SE
Literate	1940 (1-year-old)	.000	[.002]	-.003	[.003]	.001	[.003]	-.006	[.012]
	1941 (infant)	-.007***	[.002]	-.007*	[.004]	-.006**	[.003]	-.016*	[.007]
	1942 (fetus)	.004	[.002]	-.001	[.003]	.005**	[.003]	-.018**	[.011]
	N	123,793		40,053		83,740		5,811	
Upper sec. schooling	1940 (1-year-old)	-.012***	[.004]	-.021***	[.007]	-.008*	[.004]	.021	[.021]
	1941 (infant)	-.016**	[.007]	-.033***	[.011]	-.004	[.005]	.013	[.027]
	1942 (fetus)	-.006	[.008]	-.012	[.013]	.001	[.007]	-.010	[.026]
	N	123,793		40,053		83,740		5,811	
Years of education	1940 (1-year-old)	-.092**	[.043]	-.188**	[.078]	-.050	[.043]	.287	[.249]
	1941 (infant)	-.172**	[.065]	-.342***	[.101]	-.058	[.063]	.107	[.185]
	1942 (fetus)	-.017	[.064]	-.034	[.096]	.027	[.064]	.056	[.306]
	N	123,793		40,053		83,740		5,811	
ISEI-score	1940 (1-year-old)	-.282	[.329]	-.159	[.529]	-.177	[.334]	-.378	[1.12]
	1941 (infant)	-.563	[.526]	-.962	[.896]	-.288	[.531]	-.807	[.980]
	1942 (fetus)	-.493*	[.283]	-.044	[.534]	-.377	[.362]	.802	[1.22]
	N	43,177		12,859		29,426		2,229	

In all specifications I use data from the 2001 census that contains birthplace information. In addition to the indicators for whether born in 1940, 1941, or 1942, all specifications include year of birth, year of birth squared and a sex indicator. Moreover, the specifications in columns 1-3 include prefecture of birth dummies and the specification in column 4 country of birth dummies. Column 1 reproduces column 2 of Table 1.3. Column 2 provides estimates for the same specification but for the subsample of individuals born in urban areas, column 3 for individuals born in rural areas, and column 4 for individuals born outside Greece. For columns 1, 2 and 3, I report robust standard errors clustered at the prefecture of birth level in brackets. For column 4, the robust standard errors reported in brackets are clustered at the country of birth level.

urban- and 0.8 percentage points (4 percent) in the rural-born subsample. The reduction in years of education for the urban-born subsample is 0.19 years (2 percent) whereas the reduction in the rural-born subsample is less than one third as large and not statistically significant. The 1940 cohort effect on the ISEI-score is not statistically significant in either subsample.

For infants, the likelihood of being literate falls by 0.7 percentage points (0.7 percent) in the urban-born subsample and 0.6 percentage points (0.6 percent) in the rural-born subsample. The reduction in upper secondary schooling is 3.3 percentage points (7.7 percent) for the urban-born cohorts in column 2 whereas it is much smaller and not statistically significant for the rural-born cohorts in column 3. Similarly, urban-born individuals that experienced the famine as infants have 0.34 years (3.7 percent) fewer

years of education than urban-born individuals whose infancy did not coincide with the famine. At the same time, there is no statistically significant difference in years of education between individuals with and without infant famine exposure in the rural-born subsample. Just like for one-year-olds, for infants there are no statistically significant famine effects on the ISEI-score in either subsample.

Finally, for fetuses, the only statistically significant effect in columns 2 and 3 is a 0.5 percentage point (0.6 percent) increase in literacy in the rural-born subsample in column 2. The direction of the effect is puzzling, since I expected positive selection into fertility to be stronger in the urban sample, and thus, any positive bias larger in this subsample than in the rural-born subsample.

Despite this contradictory effect for rural-born fetuses, the general picture that emerges from the comparison of results in columns 2 and 3 is that long-run famine effects are larger for the urban-born subsample that was exposed to more severe early-life malnutrition. The effects indicate that the negative famine effects on the surviving urban-born cohort members are larger than the downward bias introduced through positive selection into survival. Conversely, the smaller magnitude and lack of statistical significance of effects in the rural-born subsample likely is rather due to lesser exposure to malnutrition than to higher degrees of positive selection in famine survival.

Table 1.4, column 4 shows estimates for the subsample of individuals in the 2001 census that were born outside Greece. Because members of the 1940-42 foreign-born cohorts are unlikely to have experienced systematic malnutrition early in life, I predict no significant departure in outcomes from the cohort trend for this subsample.[26]

The results in column 4 indicate statistically significant reductions in literacy for those born in 1940 and 1941. I, however, find no effects on upper secondary schooling, years of education and the ISEI-score in any of the 1940-42 cohorts. The sum of findings in column 4 is therefore consistent with my prediction, suggesting that my results for the Greece-born cohorts in Tables 1.2, 1.3 and 1.4 are unlikely to be a mere statistical anomaly.[27]

A concern with my identification strategy may be that there were events other than the 1941-42 famine that affected socioeconomic outcomes exclusively for the 1940-42 birth cohorts. A candidate for such an event is the Greek civil war. While hostilities began as early as 1942, warfare was most intense between 1946 and 1949. The 1940-42 birth cohorts with early-life famine exposure experienced this period of destabilization between age 4 and 9. For the cohorts in my control group, the 1936-39 cohorts lived through the civil war between age 7 and 13, and the 1943-46 cohorts between age 0 and 6. It is likely that any civil war related disruptions in the educational system had more severe consequences on educational outcomes for children exposed to the war at school-age than

[26] This is unless they immigrated to Greece before age three.

[27] I obtain similar results for the subsample of Albanian-born individuals that form the largest single group of immigrants in the 2001 census. Results are available from the authors on request.

for children exposed earlier. Using this assumption, any negative civil war cohort effects would be the largest for the 1936-39 cohort.

I, however, do not find that educational outcomes for the 1936-39 cohorts diverge from the 1936-46 outcome trends. Furthermore, my results for the 1940-42 cohorts do not change, if, in an alternative specification, I limit the control group to individuals born 1930-39. This indicates that the negative divergences from trend in the 1940-42 cohorts are not driven by better educational opportunities for the later-born cohorts that experienced the civil war at a younger age.

1.6 Conclusions

This chapter examines long-run education and labor market effects of early-life exposure to the Greek 1941-42 famine for the first time.

In my basic specification without birthplace controls, I find reductions in upper secondary schooling and years of education for the cohorts exposed to the famine as fetuses, infants and one-year-olds in all four waves of the Greek decennial census between 1971 and 2001. In addition, I find reductions in literacy for infants in all four censuses and for fetuses in the 1981 and 1991 waves. Finally, I find reductions in job status for all three age groups in the 1971 and 1981 waves and for infants and fetuses in the 2001 wave. The education and job status effects are typically larger for infants and fetuses than for one-year-olds.

In an alternative specification, I exploit the availability of birthplace data in the 2001 census to control for birth prefecture fixed effects. While such controls may remove more variation in the outcomes than is intended, their inclusion limits the scope for OVB from negative birthplace selection in the early-life famine exposed cohorts. The birth prefecture controls reduce the cohort effects for all outcome variables. The education effects remain statistically significant for one-year-olds and infants but the job status effects are no longer statistically significant in any age group. Also, there are no statistically significant cohort effects for fetuses after the inclusion of birth prefecture controls.

Rather than providing evidence against Barker's hypothesis, these findings underscore the threat of non-random selection into treatment in famine studies. Because the fetally exposed cohort was conceived in a situation of severe nutritional strain, it was likely subject to stronger positive selection than the pre-famine conceived cohorts. Future famine studies may therefore prioritize the modeling of sample selection mechanisms, i.e. try to shed more light on possible differences in the backgrounds of the treatment and control groups.

In a final set of results, I show that long-run adverse famine effects are larger in the urban- than in the rural-born subsample. This is in line with anecdotal evidence that the famine affected urban populations more heavily than rural populations, as the latter often benefited from proximity to agricultural production.

The consistent result of adverse education effects permits me to form a rough estimate of the famine's long-run economic cost. To obtain this cost, I utilize the effects on years

of education from my basic model in Table 1.2 and assume that a work-life in Greece extends from age 16 to 61. My calculations suggest that the famine-related economic loss is about $1.4 billion (in 2005 US dollars), or 1.1 percent of Greece's 2003 GDP, the year in which the 1942 cohort turned 61.[28] This estimate underscores the role of appropriate early-life nutrition in the formation of human capital and indicates that developmental lags in countries that frequently experience nutritional crisis may in part be due to early-life malnutrition.

References

Alderman, H., J. Hoddinott, and B. Kinsey. 2006. "Long Term Consequences of Early Childhood Malnutrition." *Oxford Economic Papers*, 58(3): 450–474.

Almond, D., K. Chay, and D. Lee. 2005. "The costs of low birthweight." *The Quarterly Journal of Economics*, 120(3): 1031-1083.

Almond, Douglas, Lena Edlund, Hongbin Li and Junsen Zhang. 2007. "Long-run effects of the 1959-61 China famine: Mainland China and Hong Kong." *NBER Working Paper* No. 13384.

Barker, D. and M. Hanson. 2004. "Altered regional blood flow in the fetus: The origins of cardiovascular disease?" *Acta Paediatrica*, 93: 1559-1560.

Barker, D. 1998. "In utero programming of chronic disease." *Clinical Science*, 95: 111-128.

Behrman, J. and M. Rosenzweig. 2004. "Returns to Birthweight." *The Review of Economics and Statistics*, 86(2): 586-601.

Black, S., P. Devereux, and K. Salvanes. 2007. "From the cradle to the labor market? The effect of birthweight on adult outcomes." *The Quarterly Journal of Economics*, 122(1): 409-439.

[28] To arrive at this cost estimate, I first use the census data to calculate the size of the 1940-42 cohorts for each year in the 1956-2003 period that covers the work-lives of the 1940-42 cohorts. I then multiply the annual cohort sizes with the labor share of Greece's per capita GDP (in 2005 US dollars) in the respective year. Multiplication of the resulting figure with estimates of the return to additional years of education in Greece (0.07) and the famine effects on years of education from Table 1.2 gives the famine-related income loss over the work-life of each of the 1940-42 cohorts. GDP data for my calculation comes from USDA's (2009) International Macroeconomic Dataset, labor share ratio estimates from OECD (2011), and estimates of the returns to years of schooling in Greece from Cholezas and Tsakloglou (1999). Further details on the method of calculation and the underlying data are available from the authors on request.

Brenner, B. and G. Chartow. 1994. "Congenital oligonephropathy and the etiology of adult hypertension and progressive renal injury." *American Journal of Kidney Disfunctioning*, 23(2): 171-175.

Bryce, J., D. Coitinho, I. Danton-Hill, D. Pelletier, and P. Pinstrup-Andersen. 2008. "Maternal and child undernutrition: effective action at national level." *Lancet,* 317: 510-526.

Case, A., D. Lubotsky, and C. Paxson. 2002. "Economic status and health in childhood: the origins of the gradient." *American Economic Review*, 92(5): 1308-1334.

Case, A. and C. Paxson. 2006. "Children's Health and Social Mobility". *Opportunity in America*, 16(2): 151-173.

Cholezas, I. and P. Tsakloglou. 1999. Private returns to education in Greece: A review of the empirical literature". In R. Asplund and P. Telhado-Pereira (eds), *Returns to human capital in Europe*, ETLA, Helsinki: 147-162.

Chen, Y. and L. Zhou. 2007. "The long-run health and economic consequences of the 1959-61 famine in China." *Journal of Health Economics*, 26: 659-681.

European Commission. 2009. "National summary sheets on education systems in Europe and ongoing reforms – Greece." Accessed only 12-04-2009 at http://eacea.ec.europa.eu/education/eurydice/eurybase_en.php#greece

Ganzeboom, H. and D. Treiman. 1996. "Internationally comparable measures of occupational status for the 1988 International Standard Classification of Occupations." *Social Science Research*, 25: 201-239.

Ganzeboom, H., P. De Graaf, and D. Treiman. 1992. "A standard international socio-economic index of occupational status." *Social Science Research* 21: 1–56.

Gørgens, T., X. Meng, and R. Vaithianathan. 2007. "Stunting and selection effects of famine: a case study of the Great Chinese Famine." *IZA Discussion Papers* 2543.

Hales, C. and D. Barker. 1992. "Type 2 non-insulin-dependent diabetes mellitus: The thrifty phenotype hypothesis." *Diabetologica,* 35: 595-601.

Hales, C. 1997. "Fetal and infant growth and impaired glucose tolerance in adulthood: The "thrifty phenotype" hypothesis revisited." *Acta Paediatrica,* Suppl. 422: 73-77.

Helger, B. 1949. "Ravitaillement de la Grèce pendant l'occupation 1941-44 et pendant les premiers cinq mois apres la libéraction." Société Hellénique d'Éditions, Athens.

Hionidou, V. 2006. “Famine and death in occupied Greece 1941-44.” Cambridge University Press, Cambridge.

Hoet, J. and M. Hanson.1999. “Intrauterine nutrition: its importance during critical periods for cardiovascular and endocrine development.” *The Journal of Physiology*, 514: 617-627.

Jakobovits, L. 1991. “Sex ratio of spontaneously aborted fetuses and delivered neonates in the second trimester.” *European Journal of Obstetrics and Gynecology and Reproductive Biology*, 40: 211-213.

Kannisto, V., K. Christensen, and J. Vaupel. 1997. “No increased mortality in later life for cohorts born during famine.” *Epidemiology*, 145(11): 987-994.

Luo, Z., R. Mu, and X. Zhang. 2006. “The long term health impact of the Chinese Great Famine.” Michigan State University Working Paper.

Maccini, S. and D. Yang. 2009. “Under the Weather: Health, Schooling, and Economic Consequences of Early-Life Rainfall”. *American Economic Review*, 99(3): 1006-1026.

Mazower, M. 1993. “Inside Hitler's Greece. The Experience of Occupation 1941-44.” Yale University Press, New Haven.

Meng, X. and N. Qian. 2009. “The long-run consequences of famine on survivors: Evidence from a unique natural experiment using China’s Great Famine.” *NBER Working Paper* No. 14917.

Minnesota Population Center. 2009. Integrated Public Use Microdata Series, International: Version 5.0 [Machine-readable database]. Minneapolis: University of Minnesota. Online at https://international.ipums.org/international/citation.shtml

Neelsen, S. and T. Stratmann. 2011. “Effects of prenatal and early-life malnutrition: Evidence from the Greek famine*”. Journal of Health Econo*mics, 30(3), 479-488.

Neugebauer, R., H. Hoek, and E. Susser. 1999. “Prenatal exposure to wartime famine and development of antisocial personality disorder in early adulthood.” *JAMA*, 282(5): 455-462.

OECD. 2011. “OECDStat: Dataset: Unit Labour Costs - Annual Indicators: Labor Income Share.” OECDStat. [accessed 17 February 2011].

Oreopoulos, P., M. Stabile, R. Walld, and L. Roos. 2008. "Short-, medium, and long-run consequences of poor infant health: an analysis using siblings and twins." *Journal of Human Resources*, 43(1): 88-138.

Ravelli, A., J. van der Meulen, C. Osmond, D. Barker, and O. Bleker. 1999. "Obesity at the age of 50 in men and women exposed to famine prenatally." *American Journal of Clinical Nutrition*, 70: 811-816.

Ravelli, A., J. van der Meulen, C. Osmond, D. Barker, C. Hales, and O. Bleker. 1998. "Glucose tolerance in adults after prenatal exposure to famine." *Lancet*, 351: 173-177.

Roseboom, T., J. van der Meulen, A. Ravelli, C. Osmond, D. Barker, and O. Bleker. 2001. "Effects of prenatal exposure to the Dutch famine on adult disease in later life: an overview." *Molecular and Cellular Endocrinology*, 185: 93-98.

Roseboom, T., J. van der Meulen, A. Ravelli, G. van Montfrans, C. Osmond, D. Barker, and O. Bleker. 1999. "Blood pressure in adults after prenatal exposure to famine." *Journal of Hypertension*, 17: 325-330.

Roseboom, T., J. van der Meulen, C. Osmond, D. Barker, A. Ravelli, and O. Bleker. 2000a. "Plasma lipid profiles in adults after prenatal exposure to the Dutch famine." *American Journal of Clinical Nutrition*, 72: 1101-1106.

Roseboom, T., J. van der Meulen, C. Osmond, D. Barker, A. Ravelli, J. Schroeder-Tanka, G. van Montfrans, R. Michels, and D. Barker. 2000b. "Coronary heart disease after prenatal exposure to the Dutch famine 1944–45." *Heart*, 84: 595-598.

St. Clair, D., M. Xu, P. Wang, Y. Yu, Y. Fang, F. Zhang, X. Zheng, N. Gu, G. Feng, P. Sham, and L. He. 2005. "Rates of adult schizophrenia following prenatal exposure to the Chinese Famine of 1959-1961." *JAMA*, 294: 557-562.

Stanner, S. and J. Yudkin. 2001. "Fetal programming and the Leningrad Siege Study." *Twin Research*, 4(5): 287-92.

Stanner, S., K. Bulmer, C. Andres, O. Lantseva, V. Borodina, V. Poteen, and J. Yudkin. 1997. "Does malnutrition in utero determine diabetes and coronary heart disease in adulthood? Results from the Leningrad Siege Study, a cross sectional-study." *British Medical Journal*, 315: 1342-1348.

Stocker, C., J. Arch and M. Cawthorne. 2005. "Fetal origins of insulin resistance and obesity." *The Proceedings of the Nutrition Society*, 64(2): 143-151.

United States Department of Agriculture (USDA). 2009. "International Macroeconomic Data Set". *Economic Research Service, United States Department of Agriculture*. Available online at http://www.ers.usda.gov/Data/Macroeconomics/ [accessed 17 February 2011].

Valaoras, V. 1946. “Some effects of famine on the population of Greece.” *The Milbank Memorial Fund Quarterly*, 24(3): 215-234.

Valaoras, V. 1960. “A reconstruction of the demographic history of modern Greece.” *The Milbank Memorial Fund Quarterly*, 38(2): 115-139.

Victoria, C., L. Adair, C. Fall, P. Hallal, R. Martorell, L. Richter, and H. Singh Sachdev. 2008. “Maternal and child undernutrition: consequences for adult health and human capital.” *Lancet*, 371: 340-57.

Walker, S., T. Wachs, J. Meeks Gardner, B. Lozoff, G. Wasserman, E. Pollitt, and J. Carter. 2007. “Child development: risk factors for adverse outcomes in developing countries.” *Lancet*, 369: 145-157.

2 Long-Run Effects of Fetal Influenza Exposure: Evidence from Switzerland

2.1 Introduction[29]

Claiming at least 20 million lives worldwide, and more than 2.6 million in Europe alone, the 1918-19 influenza pandemic was the most lethal in recent human history (Patterson and Pyle 1991, Johnson and Mueller 2002, Ansart et al. 2009).

In comparison to prior and later influenza episodes which primarily affected the very young and elderly, the 1918-19 influenza showed high infection rates and mortality among people of reproductive age. Motivated by Barker's hypothesis (Barker 1998) that later-life outcomes are partly determined at the fetal stage, this unusual pattern has prompted research on the long-run effects of fetal exposure to the pandemic.

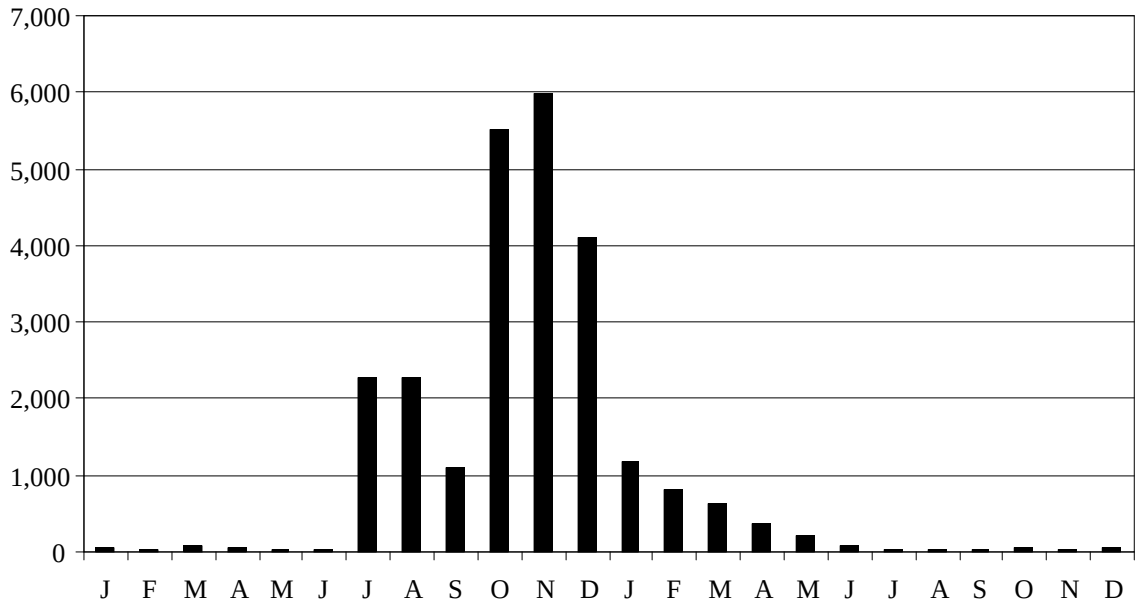
Using data from Switzerland, this work is the first to provide evidence for long-run socioeconomic effects for a European country. Switzerland is an especially suitable European country to study because it did not participate in the world wars and, unlike in other European countries, its population did not suffer severe forms of malnutrition and subsequent fertility shocks during or between the wars. This limits the scope for systematic selection into survival in my sample (Sonderegger 1991). Meanwhile, the influenza pandemic struck with similar severity as in other Western countries, killing an estimated 6.3 of 1,000 Swiss as compared to 5.2 of 1,000 U.S. citizens (Schweizerisches Bundesgesundheitsamt 1919, Thalmann, 1968, Patterson and Pyle 1991). An advantage of the Swiss data is that they contain detailed birthplace information. This permits me to limit the scope for bias from birthplace selection in my estimates and to test how variations in influenza intensity at the location of birth affect long-run socioeconomic outcomes.

2.2 Related Previous Studies

Starting with Almond and Mazumder (2005), a number of recent studies have investigated the link between fetal exposure to the 1918-19 influenza pandemic and long-run health and socioeconomic outcomes. Using U.S. data, Almond and Mazumder (2005) find that depending on the gestational age at the time of exposure, the pandemic increased the probability of stroke, diabetes, poor self-reported health, and difficulties with hearing, talking, walking and lifting. In related work for the U.S., Mazumder et al. (2010) show that fetal exposure to the pandemic led to a more than 20 percent increase in the risk of cardiovascular disease between age 60 and 82 and to lower male adult height at World War II enlistment. Also for the U.S. and depending on gestational age, Garthwaite (2008) finds influenza effects on coronary heart disease, diabetes, kidney disorders, self-reported health and educational attainment.

[29] This chapter is joint work with Thomas Stratmann (Neelsen and Stratmann 2011)

Figure 2.1: Monthly influenza deaths in Switzerland 1918-19

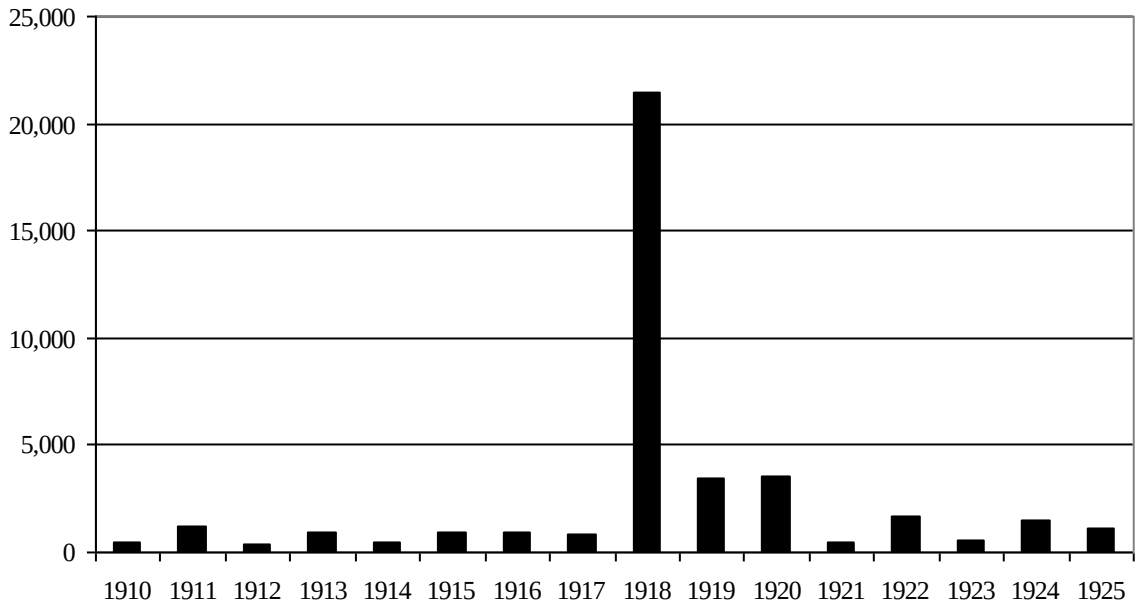


Source: Swiss Bureau of Statistics (1990)

Investigating socioeconomic influenza effects with U.S. census data, Almond (2006) shows that relative to the outcome trend for the 1912-22 cohorts, the 1919 cohort had a lower probability of high school completion, fewer years of education, lower income, lower socioeconomic status, and higher probabilities of welfare dependency, being poor and being disabled. Alternative specifications that exploit the variation in influenza severity across U.S. states confirm these results. With an empirical approach similar to Almond's, Nelson (2010) studies long-run socioeconomic effects of fetal influenza exposure for Brazil. His analysis shows that the 1919 cohort had fewer years of schooling, lower levels of literacy, a lower probability of college education and employment, and lower hourly wages than the surrounding cohorts. Again, the magnitude and statistical significance of the effects depend on the gestational age at the time of exposure.

To identify long-run influenza effects for Taiwan, Lin (2008) proxies the likelihood of fetal exposure with the province-level maternal mortality rate at the time the sampled individuals were in the fetal stage. His regression results indicate adverse effects on educational attainment that are larger for females for basic degrees and larger for males for advanced degrees. The evidence on long-run effects of fetal influenza exposure is not limited to the 1918-19 pandemic. Kelly (2009) examines the effects of fetal exposure to the 1957 Asian influenza pandemic for Britain. She identifies the effects by both over time and geographical variation in pandemic severity and finds that fetal exposure led to stunted fetal growth and lower school test scores at age 7 and 11.

Figure 2.2: Influenza deaths in Switzerland 1910-25



Source: Swiss Bureau of Statistics (1990)

2.3 The 1918-19 Influenza Pandemic in Switzerland

A first, mild wave of influenza arrived in Europe in the spring of 1918 (Sonderegger, 1991). It entered Switzerland from the northwest in mid-May, became epidemic in July and ended in mid-August. The second, extremely virulent fall wave began in late September 1918, peaked in October and November and trailed off in December. Influenza reemerged in some parts of Switzerland in February and March 1919 but with much lower intensity than in the prior outbreaks. While the influenza infection rate remained elevated until 1921, it did not reach epidemic levels after early 1919. Figures 2.1 and 2.2 show these monthly and yearly patterns of Swiss influenza deaths.

Initial estimates of infection rates suggested that the chance to contract influenza was independent of gender (Britten 1932). More recent sources, however, indicate that influenza mortality was somewhat higher among males. According to the Swiss Bureau of Statistics (1954), male mortality accounted for 58 percent of all influenza deaths.

With respect to the age-distribution of mortality, Switzerland showed the same unusual W-shaped pattern that was typical for the 1918-19 pandemic. As in other influenza waves, mortality was high among the very young and very old, but unlike in all other known influenza episodes, the 1918-19 virus claimed the majority of its victims among young and middle-aged adults that are commonly most resilient to infectious disease (Sonderegger, 1991).

There is a debate whether mortality in the 1918-19 pandemic varied by income. For Switzerland, Sonderegger (1991) analyzes influenza mortality data for different residential areas of the city of Bern. He finds a positive but not statistically significant relationship

between mortality and the number of people per living quarter, and statistically significant increases in mortality in rented, as opposed to owned homes. Furthermore, his analysis of cross-canton mortality differences suggests that influenza mortality increased with the size of the agricultural sector in a canton. The international evidence on a possible association of income and mortality is mixed. In a country-wise comparison of mortality levels, wealthier countries suffered lesser influenza deaths than poorer ones (Patterson and Pyle 1991). With respect to the relationship of income and mortality within countries, Frost (1920) and Sydenstricker (1931) for the U.S., and Witte (2006) for Germany, find that influenza mortality was elevated in groups of low social status. In contrast, Crosby's (1989) results suggest that influenza mortality was not significantly higher among the poor.

While there is conflicting evidence whether influenza mortality varied by income, various sources indicate that income did not affect the probability to contract influenza. Rotberg and Rabb (1985), for instance, suggest that mild undernutrition which might have occurred among the Swiss poor during the pandemic, does not significantly increase the likelihood of influenza infection. Consequently, Sonderegger (1991) finds no relationship between social background and the likelihood of influenza infection in Switzerland. Similar evidence exists for Germany (Witte 2006).

2.4 Data and Empirical Approach

The data available on influenza survivors do not indicate whether or not an individual's mother contracted influenza during pregnancy. Therefore, I cannot directly estimate long-run effects of fetal influenza exposure at the individual level. However, with the drastic increase in influenza infections in October and November 1918, individuals born between October 1918 and August 1919 had a heavily increased likelihood of fetal influenza exposure compared to individuals born before and after. I exploit this exogenous shock to identify long-run influenza effects on the cohort-level.

My individual level data come from the Swiss 1970 census that forms a five percent representative sample of the Swiss population (Swiss Bureau of Statistics 2009). Besides detailed demographic information, the census contains data on educational attainment, occupation, and marital status that I use to construct five socioeconomic outcome measures. For an extension of my first empirical model I also combine the individual level data with a canton-level measure of influenza mortality that I obtain from the Swiss Statistical Yearbooks 1917-22 (Swiss Bureau of Statistics, 2010).

I investigate the long-run effects of fetal influenza exposure with two models. my first model analyzes the effect of being a member of the 1919 birth cohort – a large part of which experienced the pandemic as fetuses – compared to the surrounding cohorts that had a much lower probability of fetal influenza exposure. To assure comparability of my results with those of earlier work by Almond (2006), I also use the 1912-22 birth cohorts to estimate model (1)

$$y_i = cons + yob_i + yob_i^2 + 1919 + sex_i + bcan_{i,1} + ... + bcan_{i,24} + \varepsilon_i \quad (1)$$

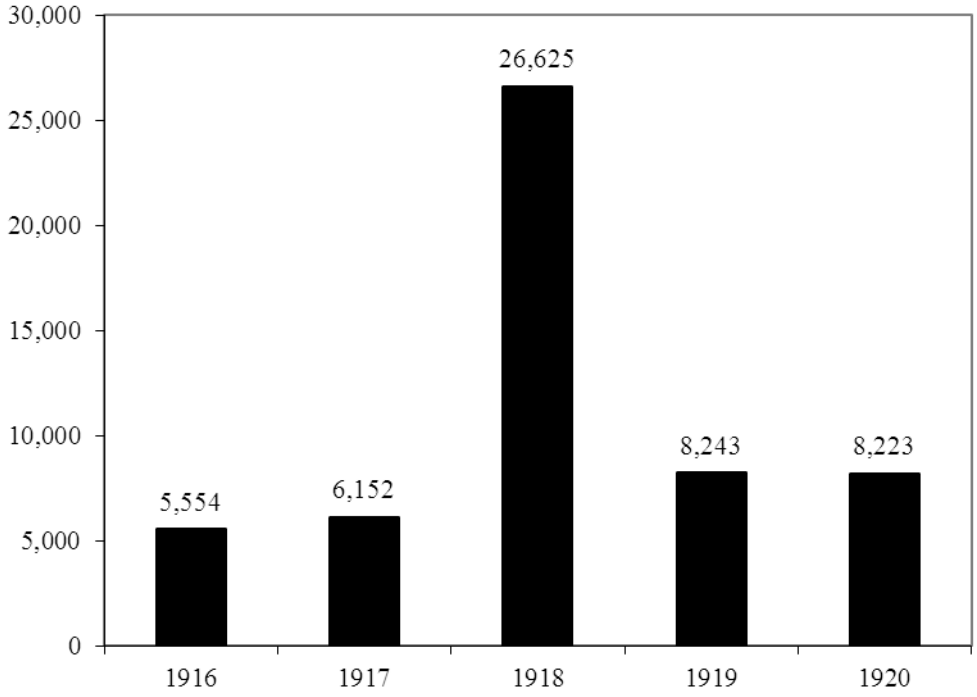
by ordinary least squares (OLS). The independent variable y_i represents the later-life outcomes for individual *i*. As outcomes, I use an indicator that equals one if an individual has not obtained any school degree, an indicator that equals one if the individual has obtained a vocational degree or any higher educational degree, an indicator that equals one if an individual has obtained a higher secondary school degree or any higher educational degree, and an indicator equaling one if an individual has never been married. For those individuals active in the labor market, y_i also represents a measure of socioeconomic status, specifically the individual's score on Ganzeboom et al.'s (1992) international socioeconomic index of occupational status (ISEI) that I compute from the ISCO88-coded occupation categories in the Swiss census. The index ranges from 16 to 90 and higher scores indicate occupations with higher socioeconomic status.

On the right hand side of model (1), *cons* represents a constant, and yob_i the year of individual *i*'s birth. yob_i thus captures a secular outcome trend. By also including the squared birth-year, yob_i^2, I allow for non-linearities in this trend. The independent variable of interest is *1919*, an indicator that equals one if an individual was born in 1919 and zero if the individual was born in one of the surrounding years in the sample. With the influenza pandemic peaking in late 1918, *1919* equals one for individuals with a highly increased likelihood of being exposed to influenza in the fetal stage. The coefficient on *1919* therefore measures the impact of fetal influenza exposure on the mean cohort outcomes. Because I use data from the 1970 census in my analysis, the 1919 cohort is 50 or 51 years of age at the time I observe the outcomes y_i.

One caveat in my study is a lack of detail in birth date information. Unlike other datasets that scholars have used to investigate Barker's hypothesis, the Swiss census only provides an individual's "age year" (census year minus birth-year). Because enumeration for the 1970 census took place from November 25 to 28, I can calculate the accurate birth-year for the majority of observations by subtracting the age year from 1970. The limited birth date information, however, does not permit me to tie long-run influenza effects to exposure at specific gestational ages.

$bcan_{i,1} + ... + bcan_{i,24}$ are 24 binary variables indicating in which of the then 25 Swiss cantons an individual was born. By including the birth-canton I rule out biases that result from an uneven mix of birth-cantons across cohorts. For instance, influenza mortality may have been particularly high in cantons with poor educational and labor market prospects. If, in addition, such birthplace selection through influenza mortality was stronger for fetuses than for young children, the share of individuals born in cantons with poor socioeconomic prospects would be lower in the 1919 cohort than in the surrounding cohorts. In this case, specifications without birthplace controls would understate the true influenza effects.

Figure 2.3: Deaths from ARD in Switzerland 1916-20



Source: Swiss Bureau of Statistics (2010), own calculations

Further, I include the indicator sex_i that equals one if an individual is female. Here, I control for the fact that both socioeconomic outcomes and the chance to succumb to influenza may vary systematically between males and females. The census data suggest that in early 20th century Switzerland, being female was associated with worse educational outcomes (Swiss Bureau of Statistics 2009). At the same time, females could be overrepresented in the 1919 birth cohort if female fetuses were less likely to die from maternal influenza infection than male fetuses. A model without gender control would hence overstate the true impact of fetal influenza exposure. Furthermore, in addition to controlling for gender in my basic specification, I allow for gender-specific 1919 birth cohort effects by estimating separate models for the female and male subsamples.

Model (1) permits me to investigate the aggregated, cohort-level effects of fetal influenza exposure. I can use these cohort-level effects to approximate the effects on the individual level, i.e., for persons whose mothers actually contracted influenza during pregnancy. Because influenza infection was likely random in terms of social background (Rotberg and Rabb 1985, Sonderegger 1991, Witte 2006), I obtain these approximations by dividing the 1919 cohort effect by the national infection rate. With an estimated 50 percent of the Swiss population contracting influenza during the 1918-19 pandemic, and assuming that infection rates among pregnant women were equal to those in the rest of the population, the individual effect is twice the size of the effect I estimate at the cohort-level.

In my second model, identification of the influenza effects builds not only on the variation of influenza severity over time but also on its variation across the Swiss cantons. I combine individual level data on later-life outcomes, birth-year and birthplace with canton-level measures of fetal influenza exposure to estimate model (2)

$$y_i = cons + ARDdr_i + unknowndr_i + undocdr_i + yob_i + sex_i + bcan_{i,1} + ... + bcan_{i,24} + \varepsilon_i \quad (2)$$

by OLS. The vector of independent variables y_i contains the same socioeconomic outcomes as in model (1). My key explanatory variable is $ARDdr_i$, the percentage of the canton population dying from influenza or other forms of acute respiratory disease (ARD) (pneumonia catarrhalis acuta, pneumonia crouposa, pleurisy, coryza, acute bronchitis, and laryngitis catarrhalis) in the year before individual *i*'s birth year (the *lagged ARD mortality rate*, hereafter). The coefficient estimate on $ARDdr_i$ thus measures the change in long-run socioeconomic outcomes due to a one percentage point increase in the lagged ARD mortality rate. Consistent canton-level ARD mortality data are available from 1916-20. Therefore, because model (2) uses the lagged ARD mortality rate to explain long-run outcomes, my sample for model (2) consists of the 1917-21 birth cohorts.

Figure 2.3 shows the total number of ARD-deaths in Switzerland between 1916 and 1920. The ARD-deaths time series reflects the timing of the influenza pandemic as shown in Figure 2.2.

The accuracy of reporting of ARD-deaths may vary between cantons and over time. Such variation could correlate with other determinants of later-life socioeconomic outcomes, as, for instance, access to health facilities and canton physician density. I therefore also include in model (2) the percentage of the canton population that died from unknown causes, $unknowndr_i$, or from undocumented causes, $undocdr_i$, in the year prior to the birth-year of individual *i*.

Further, as in model (1), yob_i is the birth-year which captures the linear outcome trend, and $bcan_{i,1} + ... + bcan_{i,24}$ are birth-canton indicators that, for instance, absorb time-invariant differences in the canton age and gender distributions. As a robustness test, I also present results for an alternative specification that includes canton-specific outcome trends measured by the interaction of each birth-canton indicator with the birth-year.

Finally, sex_i is defined as in model (1). As for model (1), I also estimate separate models for males and females for model (2).

In my data, the ARD mortality rate increases by more than a factor of four during the influenza pandemic, namely by 0.53 percentage points, from 0.16 percent in 1917 to 0.69 percent in 1918. Because the coefficients I estimate by model (2) measure the effects of a one percentage point $ARDdr_i$ increase, I can therefore approximate the pandemic's effects on the 1919 cohort by multiplying the $ARDdr_i$ coefficients by 0.53.

For models (1) and (2) to generate unbiased estimates of long-run influenza effects, the identifying assumption is that the error term ε_i is random with respect to the explanatory variables in the models. For this assumption to hold, membership in the 1919 birth cohort

Table 2.1: Summary statistics

	Model (1): 1912-22 cohorts			Model (2): 1917-21 cohorts		
	Full	**Male**	**Female**	**Full**	**Male**	**Female**
Dependent variables						
= 1 if higher secondary school degree or higher	.175 [.380]	.200 [.400]	.150 [.357]	.185 [.389]	.213 [.409]	.157 [.364]
= 1 if voc. degree or higher	.530 [.499]	.671 [.470]	.389 [.488]	.538 [.499]	.672 [.470]	.404 [.491]
= 1 if no educational degree	.002 [.047]	.002 [.044]	.002 [.049]	.002 [.049]	.002 [.046]	.003 [.051]
= 1 if never married	.114 [.318]	.100 [.299]	.129 [.335]	.111 [.314]	.098 [.297]	.124 [.330]
ISEI-score	40.421 [14.600]	40.568 [15.362]	40.078 [12.624]	40.796 [14.690]	40.995 [15.488]	40.336 [12.632]
Independent variables						
= 1 if born 1919	.087 [.281]	.087 [.282]	.086 [.281]	-	-	-
= 1 if female	.499 [.500]	-	-	.497 [.500]	-	-
% of canton population dying from ARD in year before birth	-	-	-	.282 [.209]	.283 [.208]	.282 [.208]
% of canton population dying from any respiratory disease in year before birth	-	-	-	. 447 [.217]	.447 [.217]	.447 [.217]
% of canton population dying from uncertain cause in year before birth	-	-	-	.024 [.012]	.025 [.012]	.025 [.012]
% of canton population dying from undocumented cause in year before birth	-	-	-	. 041 [.131]	.041 [.131]	.041 [.132]

and the lagged ARD mortality rate should be uncorrelated with unobserved long-run outcome determinants like parental socioeconomic status or genetic endowment.
Conversely, estimates are biased if the distributions of such omitted factors follow the same abrupt and idiosyncratic patterns as the pandemic. For example, if the 1919 cohort had superior (inferior) unobserved characteristics than the surrounding cohorts, for instance due to selective influenza mortality, the point estimate on the 1919 cohort indicator in model (1) would understate (overstate) the true influenza effects.

As I discuss above, social selection into survival during the pandemic may have been positive, i.e., the well-off may have had a higher chance of survival. If, in addition, such selection was stronger for fetuses than for young children, the share of individuals born to parents with low social status would be lower in the 1919 cohort than in the surrounding

cohorts. In this case, my estimates from models (1) and (2) would form lower bounds of the true influenza effects. The same logic applies for selection with respect to genetic endowments.

Additional bias can stem from events other than the 1918-19 influenza pandemic that exclusively affected the 1919 cohort or events that varied by the same geographical and time pattern as ARD mortality. For example, if the 1919 cohort had not only had a higher likelihood of fetal influenza exposure but also of fetal malnutrition, it would be impossible to distinguish the long-run effects of influenza from those of malnutrition. However, although Switzerland experienced a degree of food shortage at the time of the pandemic, it was not severe enough to impair fetal development (Greyerz 1977, Jost 1983, Abu-Saad and Fraser 2010). Finally, I am not aware of any later-life events that may have affected the 1919 cohort in particular compared to the surrounding cohorts, like, for instance, educational reform.

2.5 Results

Table 2.1 summarizes the full sample and gender-wise sample averages for the dependent and main independent variables in models (1) and (2). In both the 1912-22 and the 1917-21 samples fewer females than males have at least a higher secondary school degree or at least a vocational degree. Further, females are slightly more likely than males to never marry and have slightly lower ISEI-scores. There are no gender differences in the likelihood of having no school degree at all.

Table 2.2 presents outcome departures for the fetally-exposed 1919 cohort from the 1912-22 trend as estimated by model (1). The results I present in the following are robust to the use of different sample periods. The first two columns show the full sample results for specifications without (column 1) and with (column 2) indicators for the birth-canton. Columns 3 and 4 show results for the male and females subsamples, controlling for birth-canton effects. In Table 2.2 and all following tables, I present standard errors clustered at the birth-canton level in brackets.

The full sample results in column 1 show negative cohort effects for individuals born in 1919 – the cohort in which the likelihood of fetal exposure to influenza was drastically increased. Namely, being born in 1919 reduces the likelihood of a vocational degree by 0.5 percentage points which equals an 0.9 percent reduction compared to the sample average. The 1919 cohort's likelihood to never marry also departs from the 1912-22 trend with an increase of 0.3 percentage points (2.5 percent). Moreover, the 1919 cohort has a 0.17 point (0.42 percent) lower ISEI-score. The 1919 cohort effects on having at least a higher secondary school degree and on not having any educational degree also indicate negative influenza effects but are not statistically significant.

My results in column 2 support the hypothesis that mortality was higher in cantons with poor socioeconomic prospects in that the inclusion of birth-canton controls increases the negative 1919 cohort effect from column 1 in absolute value: in column 2, being born in 1919 reduces the likelihood of a vocational degree by 0.6 percentage points (1.1

Table 2.2: Model 1 - Long-run effects of fetal influenza exposure; 1919 birth cohort effects

	(1)	(2)	(3)	(4)
	Full sample	**Full sample**	**Males only**	**Females only**
Dependent variables				
Higher secondary school degree or higher	-.00194 [.00173] N = 598,500	-.00246 [.00175] N = 598,500	-.00220 [.00232] N = 299,562	-.00268 [.00225] N = 298,938
Vocational degree or higher	-.00544*** [.00113] N = 598,500	-.00621*** [.00111] N = 598,500	-.00751*** [.00219] N = 299,562	-.00479** [.00217] N = 298,938
No educational degree	.00033 [.00024] N = 598,500	.00032 [.00024] N = 598,500	.00064* [.00034] N = 299,562	-.00008 [.00032] N = 298,938
Never married	.00286** [.00113] N = 598,500	.00307** [.00111] N = 598,500	.00371** [.00162] N = 299,562	.00246 [.00155] N = 298,938
ISEI-score	-.16809 [.10094] N = 413,586	-.18270* [.10628] N = 413,586	-.19892 [.12121] N = 290,160	-.14297 [.18840] N = 123,426
Independent variables				
Birth-year	Yes	Yes	Yes	Yes
Sex	Yes	Yes	-	-
Birth-canton	No	Yes	Yes	Yes

Cluster-robust standard errors with clustering at the birth-canton level in brackets: * = p<0.10, ** = p<0.05, *** = p<0.01.

percent), increases the likelihood to never marry by 0.3 percentage points (2.5 percent), and decreases the ISEI-score by 0.18 points (0.45 percent).

In columns 3 and 4 of Table 2.2, I show 1919 cohort effect estimates for the male and female subsamples. For the male subsample in column 3, the statistically significant 1919 cohort effects are a reduction in the likelihood of having a vocational or higher degree of 0.8 percentage points (1.1 percent), an increase in the likelihood of no school degree of 0.06 percentage points (33 percent), and an increase in the likelihood of never getting married of 0.3 percentage points (3.8 percent). For females, the only statistically significant 1919 cohort effect is a 0.5 percentage point reduction (1.2 percent) in the likelihood of a vocational or higher degree. Although the signs are of the expected direction in both the female and male subsample regressions, for neither sex do I find a statistically significant 1919 cohort effect for the likelihood of having at least a higher secondary school degree, or on the ISEI socioeconomic status score.

Table 2.3: Model 1 - Long-run effects of fetal influenza exposure; 1918-20 birth cohort effects in male subsample

	Separate models			Single model with 3 cohort dummies		
	(1)	(2)	(3)	(4)	(5)	(6)
	1918 dummy	1919 dummy	1920 dummy	1918 dummy	1919 dummy	1920 dummy
Dependent variables						
Higher secondary school degree or higher	-.00407	-.00220	.00175	-.00461	-.00302	.00051
N = 299,562	[.00320]	[.00232]	[.00236]	[.00338]	[.00249]	[.00225]
Vocational degree or higher	-.00216	-.00751***	-.00257	-.00506	-.00953***	-.00502*
N = 299,562	[.00275]	[.00219]	[.00270]	[.00304]	[.00248]	[.00291]
No educational degree	-.00038	.00064*	.00031	-.00016	.00069*	.00041
N = 299,562	[.00028]	[.00034]	[.00024]	[.00028]	[.00034]	[.00024]
Never married	.00003	.00371**	.00032	.00112	.00418**	.00123
N = 299,562	[.00176]	[.00162]	[.00183]	[.00191]	[.00161]	[.00202]
ISEI-score	-.16487	-.19892	.03340	-.22679	-.25383**	-.04613
N = 290,160	[.14365]	[.12121]	[.09936]	[.13987]	[.12192]	[.07806]
Independent variables						
Birth-year	Yes	Yes	Yes	Yes	Yes	Yes
Birth-canton	Yes	Yes	Yes	Yes	Yes	Yes

Cluster-robust standard errors with clustering at the birth-canton level in brackets: * = p<0.10, ** = p<0.05, *** = p<0.01. Columns (1)-(3) show results for three different models where column 1 includes a dummy for the 1918 cohort, column 2 includes a dummy for the 1919 cohort, and column 3 includes a dummy for the 1920 cohort. Columns (4)-(6) refer to the coefficients estimated from a single model that includes one dummy for the 1918 cohort, one dummy for the 1919 cohort and one dummy for the 1920 cohort.

The lack of statistically significant effects in the female subsample indicates that the 1919 cohort effects that I find in columns 1 and 2 of Table 2.2 are primarily driven by the male subsample.

For the male subsample, Table 2.3 presents two different robustness tests for the results I show in Table 2.2. For reference, Table 2.3, column 2 replicates the results for the male 1919 birth cohort that I also show in Table 2.2, column 3. The first (third) column shows results for a specification that is equal to model (1) except that it includes a

1918 (1920) birth-year dummy instead of a 1919 birth-year dummy. The first (third) column thus shows the 1918 (1920) male birth cohort's departure from the 1912-22 trend.

Because the Swiss 1918 influenza pandemic had a (milder) summer wave and peaked in October and November, the likelihood of fetal influenza exposure was already increased for late-born members of the 1918 birth cohort (see Figure 2.1). The majority of the 1918 cohort, however, experienced the pandemic in its first year of life. Because among children, influenza vulnerability decreases in age, I expect any long-run influenza effects for the 1918 cohort to be smaller than in the 1919, mostly fetally-exposed cohort.

Further, while influenza prevalence remained elevated in 1919 and 1920 it had decreased to non-epidemic levels by 1919 (see Figures 2.2 and 2.3 for timelines of annual influenza and ARD-deaths). I therefore do not predict large negative birth cohort effects for those born in 1920.

The results in columns 1-3 of Table 2.3 are consistent with these predictions. They show that while the cohort effects for those born in 1918 and 1920 are of the expected sign for the majority of outcomes, there are no statistically significant departures from the 1912-22 trend. Further, with the exception of the not statistically significant effect on having a higher secondary school or higher degree, all 1919 birth cohort effects are of larger absolute magnitude than the effects I measure for the surrounding cohorts.

Columns 4-6 of Table 2.3 show 1918, 1919, and 1920 cohort effect estimates for an alternative specification that includes the three cohort dummy variables jointly. Because this specification changes the underlying outcome trend against which the cohort effects are measured, the sizes of the effects somewhat change compared to those in columns 1-3 of Table 2.3. The results I present in columns 1-3 and 4-6 of Table 2.3 are, however, qualitatively similar for the 1919 birth cohort that had a strongly increased likelihood of fetal influenza exposure. In contrast to the models with a single 1919 cohort dummy in Table 2.2, the specification with three cohort dummies yields a statistically significant reduction in the ISEI socioeconomic status score for the 1919 cohort. Further, in the model with three cohort dummies there is a statistically significant reduction in the probability of a vocational degree for the 1920 cohort that likely reflects elevated influenza infections in 1919. The effect is, however, only half as large as the effect I find for the 1919 cohort. In fact, in line with my predictions, for all outcomes for which I find statistically significant 1919 birth cohort effects, the magnitude of the 1919 cohort effect is larger than that for the 1918 and 1920 cohorts.

In summary, the results I present in Table 2.3 support the hypothesis that fetal influenza exposure harms long-run outcomes. They also indicate that the effects are visible at the cohort-level only if influenza infections reach epidemic levels.

Table 2.4 shows full sample results for model (2). Here I use regional and over time variations in the lagged ARD mortality rate to measure long-run effects of fetal influenza exposure. Table 2.4, column 1 shows results from regressing socioeconomic outcomes on the canton-level lagged ARD mortality rate, a linear outcome trend a gender dummy, and birth-canton indicators.

Table 2.4: Model 2 - Long-run effects of fetal influenza exposure in the full sample; using deaths from respiratory disease as a measure of exposure

	(1)	(2)	(3)	(4)
Dependent Variables				
Higher secondary school degree or higher N = 272,482	-.00371 [.00293]	-.00437 [.00367]	-.00425 [.00395]	-.00427 [.00391]
Vocational degree or higher N = 272,482	-.00886*** [.00309]	-.00974*** [.00270]	-.00844*** [.00297]	-.00807** [.00305]
No educational degree N = 272,482	.00067 [.00048]	.00064 [.00051]	.00064 [.00053]	.00063 [.00052]
Never married N = 272,482	.00726*** [.00255]	.00806*** [.00254]	.00878*** [.00284]	.00888*** [.00272]
ISEI-score N = 191,169	-.43875** [.2032]	-.45928* [.23538]	-.44812* [.23828]	-.43691* [.23042]
Independent Variables				
% of canton pop. dying from ARD in year before birth	Yes	Yes	Yes	No
% of canton pop. dying from any resp. disease in year before birth	No	No	No	Yes
Sex	Yes	Yes	Yes	Yes
% of canton pop. dying from undocumented or uncertain causes in year before birth	No	Yes	Yes	Yes
Birth-canton	Yes	Yes	Yes	Yes
Canton-specific trend	No	No	Yes	Yes

Cluster-robust standard errors with clustering at the birth-canton level in brackets: * = p<0.10, ** = p<0.05, *** = p<0.01.

Table 2.4, column 2 adds the canton share of deaths with unknown and undocumented cause, and column 3 shows results for a specification that in addition controls for canton-specific time trends. The results in column 4 are for a specification that is equal to that used in column 3 except that it explains long-run socioeconomic outcomes with a

mortality rate that besides ARD-deaths includes deaths from diphtheria, whooping cough, and pulmonary tuberculosis. The number of reported deaths from these diseases increased during the influenza pandemic. These increases may in part have been due to a misclassification of deaths that actually occurred from influenza (Witte 2006, Lin, 2008). The broader measure of influenza severity in column 4 therefore avoids a bias that may stem from variations in the degree of such misclassifications between cantons or over time.

The results in Table 2.4 provide further support for the hypothesis that an increased likelihood of fetal exposure led to reductions in educational attainment and to disadvantages on the marriage and labor markets. The results in column 2 show that a one percentage point increase in the lagged ARD mortality rate caused a one percentage point (1.8 percent) reduction in the likelihood of a vocational or higher degree. As discussed above, the pandemic's long-run effects on the 1919 cohort are about half the size of the coefficients I estimate in model (2). The 0.9 percent reduction in the likelihood of a vocational degree that I obtain from this approximation is close to the effect I measure by model (1), providing additional support for the plausibility of my findings in Tables 2.2 and 2.3.

Further, a one percentage point increase in the lagged ARD mortality rate increases the likelihood of never marrying by 0.8 percentage points (7.3 percent). Multiplication by 0.53 to approximate the effect in the 1919 fetally-exposed birth cohort gives a 3.7 percent increase which is again close to the effect estimated in model (1). Finally, I find a weakly statistically significant reduction in the ISEI socioeconomic status score of 0.46 points (1.1 percent) among individuals with an increased risk of fetal influenza exposure. Again, the calculation of the pandemic's effect on the 1919 cohort gives results that are similar to those obtained by model (1).

Dropping the undocumented and uncertain cause of death variable in column 1 and adding birth-canton-specific time-trends in column 3 yields similar results as I obtain in my main specification in column 2. In addition, there is no substantial change in the results of column 2 when I use the broader definition of deaths from respiratory disease in column 4.

Table 2.5 shows estimation results for the same specifications as in Table 2.4 for the male subsample. In my main specification in column 2, I do not find a statistically significant effect on the likelihood of having a higher secondary school degree, although the estimated coefficient is of the expected sign. In contrast, a one percentage point increase in the lagged ARD mortality rate decreases the likelihood of having at least a vocational degree by 1.4 percentage points (2 percent). Moreover, I find a statistically significant increase in the likelihood of having no school degree at all of 0.2 percent. While this increase is small in absolute size, it corresponds to a 76 percent increase in the likelihood of this outcome. Multiplying the point estimate for a one percentage point increase in ARD mortality rate by 0.53 to approximate the 1919 birth cohort effect gives a 38 percent reduction that comes close to the 33 percent reduction I estimate in model (1).

Table 2.5: Model 2 - Long-run effect of fetal influenza exposure in the male subsample; using deaths from respiratory disease as a measure of exposure

	(1)	(2)	(3)	(4)
Dependent variables				
Higher secondary school degree or higher N = 136,925	-.00478 [.00446]	-.00454 [.00531]	-.00411 [.00567]	-.00357 [.00558]
Vocational degree or higher N = 136,925	-.01223** [.00445]	-.01377*** [.00481]	-.01141** [.00553]	-.01100* [.00554]
No educational degree N = 136,925	.00141* [.00069]	.00151** [.00072]	.00150* [.00074]	.00144* [.00073]
Never married N = 136,925	.00745** [.00318]	.00640 [.00381]	.00646 [.00396]	.00621 [.00386]
ISEI-score N = 133,656	-.54362** [.22652]	-.58662** [.27072]	-.54951* [.28204]	-.54048* [.27468]
Independent variables				
% of canton population dying from ARD in year before birth	Yes	Yes	Yes	No
% of canton population dying from any respiratory disease in year before birth	No	No	No	Yes
% of canton population dying from undocumented or uncertain causes in year before birth	No	Yes	Yes	Yes
Birth-canton	Yes	Yes	Yes	Yes
Canton-specific trend	No	No	Yes	Yes

Cluster-robust standard errors with clustering at the birth-canton level in brackets: * = p<0.10, ** = p<0.05, *** = p<0.01.

Contrary to the results obtained from model (1), I do not find a statistically significant increase in the likelihood to never marry in the male subsample for model (2), even though the estimated coefficient has the predicted positive sign. Finally, model (2) yields a statistically significant reduction in the ISEI socioeconomic status score of 0.59 points (1.4 percent).

Table 2.6: Model 2 - Long-run effect of fetal influenza exposure in the female subsample; using deaths from respiratory disease as a measure of exposure

	(1)	(2)	(3)	(4)
Dependent Variables				
Higher secondary school degree or higher N = 135,557	-.00272 [.00347]	-.00423 [.00386]	-.00432 [.00388]	-.00497 [.00379]
Vocational degree or higher N = 135,557	-.00553 [.00420]	-.00550 [.00470]	-.00509 [.00487]	-.00486 [.00486]
No educational degree N = 135,557	-.00008 [.00056]	-.00026 [.00061]	-.00024 [.00061]	-.00020 [.00060]
Never married N = 135,557	.00713* [.00377]	.00990*** [.00287]	01132*** [.00294]	.01173*** [.00276]
ISEI-score N = 57,513	-.18890 [.35580]	-.16216 [.36313]	-.20646 [.37634]	-.18981 [.36784]
Independent Variables				
% of canton pop. dying from ARD in year before birth	Yes	Yes	Yes	No
% of canton pop. dying from any resp. disease in year before birth	No	No	No	Yes
% of canton pop. dying from undocumented or uncertain causes in year before birth	No	Yes	Yes	Yes
Birth-canton	Yes	Yes	Yes	Yes
Canton-specific trend	No	No	Yes	Yes

Cluster-robust standard errors with clustering at the birth-canton level in brackets: * = p<0.10, ** = p<0.05, *** = p<0.01.

While the results for the male subsample in Table 2.5, column 2 somewhat lose in statistical significance through the omission of controls for the share of undocumented and uncertain causes of death (column 1) and the inclusion of canton-specific time-trends (column 3), they remain similar in magnitude. They are also robust to using the broader measure of fetal exposure to respiratory disease (column 4).

Finally, Table 2.6 shows estimation results for the female subsample for the same specifications I used in Tables 2.4 and 2.5. The direction of the coefficients indicate that

an increase in the lagged ARD mortality rate decreased the likelihood of having a vocational degree and the likelihood of having a higher secondary school degree, increased the likelihood to never marry, and decreased the ISEI socioeconomic status score. With the exception of the likelihood to never marry (a 0.9 percentage point, or 8 percent increase), I, however, estimate none of the effects precisely. Again, the estimation of statistically significant point estimates in the full sample appears to be mainly driven by the variation in outcomes in the male subsample.

2.6 Discussion

I provide further evidence that the 1918 influenza pandemic's adverse effects went beyond its immediate death toll because infections of pregnant women impaired socioeconomic outcomes among their offspring. The depth of my analysis is, however, limited by data constraints. Most importantly, the lack of detailed birth date information in the Swiss census prevents me from testing to what degree long-run socioeconomic effects depend on the gestational age at the time of exposure. Furthermore, I can examine long-run effects only for a limited number of outcomes. For instance, I do not have data on health status or income.

I suspect that at least two mechanisms drive the transmission of the fetal health shock into long-run outcomes. First, maternal influenza infection has been shown to cause fetal growth retardation, pre-term births, birth defects and defects of the central nervous system (see literature review in Nelson, 2010). These innate conditions affect school performance either directly through impaired cognitive ability or indirectly through reduced school attendance (Case and Paxson, 2006). Lesser education subsequently impairs labor and marriage market outcomes. Second, Almond (2006) and Nelson (2010) present evidence that the negative income effects of fetal influenza exposure are not merely due to reductions in educational attainment. Instead, an additional worsening of health conditions during adulthood appears to further reduce the income opportunities of individuals with fetal influenza exposure. If this result from U.S. and Brazilian data can be transferred to the Swiss case, the ISEI-score reductions in the Swiss 1919 cohort may in part be the result of a worsening of health states during adulthood.

Unlike for males, I find little evidence for adverse education and labor market effects for females. One possible explanation for this is that in early 20th century Switzerland, women's education and labor market opportunities were largely determined by traditional gender roles rather than by cognitive ability and physical health (Joris and Witzig, 2001). Any influenza effects on educational and labor market potential would hence be muted in the female population. In contrast, it is plausible that female health is a signal for fertility and the health of future children. This would explain the marriage market effect in the female sample in model (2). In addition, the gender differences could partly be due to females being more resilient to fetal influenza exposure. This hypothesis is supported by Mazumder et al. (2010) and Garthwaite (2008) who find that later-life health impairments are more severe for males than for females.

The comparison of my results to those of prior studies is complicated by differences in outcome variables and in the education and labor market institutions of the countries under study. Almond (2006) finds 13-15 percent reductions in high school completion for individuals with fetal influenza exposure. This effect is several times larger than the 2.2 percent individual level effect I estimate for a similar measure, completion of a vocational-degree, in Switzerland. Nelson's (2010) results for Brazil indicate even larger reductions in male educational attainment than in the U.S. The author explains these differences with lower parental and public health investments in a developing country. Richer societies with more equal access to healthcare, he argues, would help the catch-up of individuals with fetal health shocks. Following this line of argument, the relatively low access barriers to healthcare and relatively equal distribution of wealth in Switzerland would provide an explanation for the smaller Swiss influenza effects. Similarly, if economic well-being in fact improved catch-up, and if children of well-off families are overrepresented in higher education tracks, this may explain why I find influenza effects primarily for lower educational degrees.

References

Abu-Saad, K. and D. Fraser. 2010. "Maternal nutrition and birth outcomes". *Epidemiological Review*, advance access published March 17, 2010.

Almond, D. 2006. "Is the 1918 influenza pandemic over? Long-run effects of in utero influenza-exposure in the post-1940 U.S. population". *Journal of Political Economy*, 114(4): 672-711.

Almond, D. and B. Mazumder. 2005. "The 1918 influenza pandemic and subsequent health outcomes: An analysis of SIPP data". *American Economic Review*, 95: 258-262.

Ansart, S., C. Pelat, P. Boelle, F. Carrat, A. Flahault, and A. Valleron. 2009. "Mortality burden of the 1918–1919 influenza pandemic in Europe". *Influenza and Other Respiratory Viruses*, 3(3): 99-106.

Barker, D. 1998. "In utero programming of chronic disease". *Clinical Science*, 95, 111-128.

Britten, R. 1932. "The incidence of epidemic influenza 1918-19". *Public Health Reports*, 47(6): 303-339.

Case, A., and C. Paxson. 2006. "Children's health and social mobility". *Opportunity in America*, 16(2): 151-173.

Crosby, A. 1989. America's forgotten pandemic: The influenza of 1918. Cambridge: Cambridge University Press.

Frost, W. 1920. "Statistics of influenza morbidity: With special reference to certain factors in case incidence and case fatality". *Public Health Reports*, 35(11): 584-98.

Ganzeboom, H., P. De Graaf, and D. Treiman. 1992. "A standard international socio-economic index of occupational status". *Social Science Research*, 21: 1-56.

Garthwaite, C. 2008. "The effect of in-utero conditions on long term health: Evidence from the 1918 Spanish flu pandemic". http://www.kellogg.northwestern.edu/faculty/garthwaite/htm/fetal_stress_garthwaite_053008.pdf [last accessed online November 26, 2010]

Greyerz, H. 1977. "Der schweizerische Bundesstaat seit 1848". In Handbuch der Schweizer Geschichte, Vol. 2. Zürich.

Johnson, N., and J. Mueller. 2002. "Updating the accounts: Global mortality of the 1918–1920 "Spanish" influenza pandemic". *Bulletin of Historical Medicine*, 76: 105-115.

Joris, E., and H. Witzig. 2001. Frauengeschichte(n). Dokumente aus zwei Jahrhunderten zur Situation der Frauen in der Schweiz. Zürich: Limmat-Verlag.

Jost, H. 1983. "Bedrohung und Enge". In: Geschichte der Schweiz und der Schweizer, Vol. 3, 101-90. Basel: Helbing and Lichtenhahn.

Kelly, E. 2009. "The scourge of Asian flu: in utero exposure to influenza pandemic and the development of a cohort of British children". *IFS Working Paper*, W09/17.

Lin, M. 2008. "Does "in utero" pandemic exposure matter? Evidence from the long term effect of the 1918 influenza on educational attainments in Taiwan". *National Taiwan University Working Paper*.

Mazumder, B., D. Almond, K. Park, E. Crimmins, C. Finch. 2010. "Lingering prenatal effects of the 1918 influenza pandemic on cardiovascular disease". *Journal of Developmental Origins of Health and Disease*, 1(1): 26-34.

Neelsen, S. and T. Stratmann. 2011. "Long-run effects of fetal influenza exposure: Evidence from Switzerland". *Social Science & Medicine*, available online 11 November 2011.

Nelson, R. 2010. "Testing the fetal origins hypothesis in a developing country: Evidence from 1918 Influenza pandemic". *Health Economics*, 19(10): 1181-1192.

Nussbaum, W. 1981. Die Grippe-Epidemie 1918-1919 in der schweizerischen Armee. Bern.

Patterson, D. and G. Pyle. 1991. "The geography and mortality of the 1918 influenza pandemic". *Bulletin of the History of Medicine*, 65(1): 4-21.

Rotberg, R. and T. Rabb. 1985. Hunger and history: The impact of changing food production and consumption patterns on society. Cambridge: Cambridge University Press.

Schweizerisches Bundesgesundheitsamt. 1919. *Bulletin des Schweizer Gesundheitsamtes 1918-19*. Bern: Schweizerisches Bundesgesundheitsamt.

Sonderegger, C. 1991. Die Grippeepidemie 1918-19 in der Schweiz. Bern: Lizentiatsarbeit, Historisches Institut, Universität Bern.

Swiss Bureau of Statistics. 1954. Gestorben infolge Grippe nach Geschlecht, Sterbemonaten und Kantonen 1918-20. Bern: Swiss Bureau of Statistics.

Swiss Bureau of Statistics. 1990. Grippetote nach Geschlecht und Alter 1877-1989. Bern: Swiss Bureau of Statistics.

Swiss Bureau of Statistics. 2009. Schweizerische Volkszählung 1970. Mikrodaten. Bern: Obtained from Swiss Bureau of Statistics.

Swiss Bureau of Statistics. 2010. *Das Jahrbuch Archiv*. http://www.bfs.admin.ch/bfs/portal/de/index/dienstleistungen/publikationen_statistik/statistische_jahrbuecher/stat__jahrbuch_der/jahrbuch-archiv.html [last accessed November 26, 2010]

Sydenstricker, E. 1931. "The incidence of Influenza among persons of different economic status during the epidemic of 1918." *Public Health Reports*, 46(4): 154.

Thalmann, H. 1968. Die Grippepandemie in Zürich 1918-19. Dissertation, Zürich.

Witte, W. 2006. Erklärungsnotstand - Die Grippe-Epidemie 1918-1920 in Deutschland unter besonderer Berücksichtigung Badens. Centaurus Verlag, Herbolzheim.

3 Infant Exposure to an Indonesian Midwife Program and Adolescent Height and Cognitive Skill

3.1 Introduction

More than 200 million of today's children will fail to reach their full developmental potential because of poor living conditions in their formative years (Grantham-McGregor et al. 2007). These conditions include early-life exposure to malnutrition, infectious disease, inadequate cognitive stimulation, and violence (Walker et al. 2007). By impairing health and the development of cognitive abilities, these early-life insults not only harm individual well-being and socioeconomic prospects. They also bear substantial societal costs that arise in the form of human capital losses, forgone productivity and income, higher healthcare costs, and rising crime rates (Naudeau et al. 2011a). Conversely, improving early-life conditions has high societal payoffs (Heckman and Masterov 2007). To improve child health in the short run and achieve long-run improvements in individual and societal development, national governments and donors are thus reprioritizing their policies towards the youngest (Naudeau et al. 2011b). To be efficient, this process requires comprehensive appraisals of the costs and benefits of such policies – including those that accrue in the long-run.

However, while a large number of studies have investigated long-run effects of early-life interventions in developed countries (Almond and Currie 2011) similar work for developing countries remains scarce and results are incoherent (Walker et al. 2007). This study aims to contribute to the closing of this knowledge gap by examining impacts of infant exposure to a large-scale Indonesian midwife program on adolescent height-for-age and cognitive skill. Using data from the longitudinal Indonesian Family Life Survey (IFLS) the study is the first to investigate the program's long-run effects and one of the first on later-life impacts of early-life community-level interventions in a developing country (see Walker et al. 2005 and Hoddinott et al. 2008 for exceptions).

The program dispatched trained midwives to over 50,000 Indonesian communities in the early- to mid-1990s. The midwives' responsibilities went beyond birth attendance and included basic curative and preventive care for pregnant women and young children, the distribution of basic medicines and nutritional supplements, vaccinations, the education of mothers on topics like oral rehydration, breastfeeding and child nutrition, and the raising of awareness for public health issues like sanitation and personal hygiene in the community (Frankenberg et al. 2005). With this broad set of services, the program bore potential for not only achieving its primary goal of reducing maternal mortality but also for improving early-life living conditions, and, thereby, the long-run physical and cognitive development of individuals exposed to the program at young age.

The intuitive starting point for examining the program's long-run effects is the comparison of users and non-users of midwife services. This approach, however, suffers from several weaknesses. Most importantly, use of interventions like the village midwife program is typically not random in terms of other determinants of child development. For

instance, pregnant women who consult a program midwife may, on average, have more complicated pregnancies than pregnant women who do not. If pregnancy complications affect long-run outcomes, failure to control for this confounder causes omitted variables bias (OVB) in the measurement of program effects. The same holds for variables like parental wealth and education. Moreover, even when controlling for all joint determinants of midwife use and child development, program spill-overs can blur the distinction of treatment and control groups, and thus confound program effect estimation. Such spillovers occur when program benefits also accrue to mothers and children that do not consult a midwife directly – an aspect that gains particular relevance through the midwives' role as proponents of public health initiatives in the community (Frankenberg et al. 2005).[30]

To avoid such biases, scholars, including those of the midwife program, have used community- instead of individual-level variation in program exposure to identify causal effects (Frankenberg et al. 2005). The basic idea is to compare the long-run outcomes of individuals from communities where the intervention took place to those of individuals from communities with no intervention. However, if the placement of the intervention is not random across communities but correlates with unobserved community characteristics that determine long-run outcomes, OVB arises just like in the case of individual-level unobserved heterogeneity (Angeles et al. 1998).

At least in its initial implementation phase, the midwife program gave priority to placing midwives in poor, underserved communities (Frankenberg and Thomas 2001). Midwife placement therefore certainly correlated with variables that negatively affect child development. Hence, if the placement mechanism is not controlled for in the empirical model, OVB prohibits inference of causal midwife effects.

I address this issue with an identification strategy that exploits not only spatial but also over-time variation in early-life midwife exposure. In less formal terms, my control group not only consists of individuals from communities that did not participate in the program (spatial), but also of individuals from participating communities who were exposed at ages at which long-run effects are less likely (over-time). The latter, intra-community variation in midwife exposure permits me to include community fixed effects in my empirical model – and thus rule out all time-invariant determinants of midwife placement as sources of OVB.

My empirical analysis shows that presence of a midwife in the year after the birth year is associated with statistically and economically significant improvements in height-for-age and cognitive skill at ages 10-14. The effects are larger for adolescent females than for males and restricted to children of less-well educated mothers. The latter finding indicates that the program contributes to the reduction of socioeconomic inequities in child development.

[30] Spillovers can also occur through maternal learning, i.e. for children that did not receive midwife care directly but whose mothers consulted midwives for prior births. However, because the treated cohorts in this study come from a rather short time-interval, the number of siblings for which such spillovers may occur is limited.

The rest of this chapter is organized as follows. Section 3.2 provides details on the midwife program and Section 3.3 reviews prior evaluation studies. Section 3.4 introduces the dataset and empirical approach and Section 3.5 presents estimation results. Section 3.6 concludes.

3.2 The Indonesian Midwife Program[31]

Aiming to reduce maternal mortality and make basic maternal and child care more accessible to underserved populations, the Indonesian government initiated a large scale community midwife program (*bidan di desa*) in 1989. With 13,000 midwives at the time of program initiation, the goal was to place a midwife in all 54,000 communities in need by the mid-1990s (World Bank 1991). The program recruited nursing academy graduates to one-year midwifery trainings to subsequently place them in selected communities. The placement process initially prioritized poor and remote areas with little access to healthcare services. To guarantee midwives a steady income despite offering their services at subsidized rates or free of charge, they received a government salary in the first three to six years of their service. To top up this income, they were allowed to practice privately after hours on a fee-for-service basis.

Besides their primary task of reducing maternal mortality through skilled birth attendance, the midwives provide a broad set of preventive and curative services to improve child health. These include the provision of pre- and neo-natal care, the administration of essential medicines and micronutrients such as iron and vitamin A, vaccinations, and the education of mothers and the community on topics like family planning, child nutrition, hygiene and sanitation.

By 1997, the program had placed over 52,000 midwives across the country, achieving coverage of 96 percent of targeted communities and raising midwife density from 0.2 per 10,000 inhabitants in 1986 to 2.6 in 1996 (Ministry of Health 1997, 2000).

3.3 Earlier Evaluation Studies

Earlier evaluation studies indicate high uptake of midwife services and short-run impacts on maternal and child health. The share of births attended by skilled midwives doubled to 55 percent between 1990 and 2003 (Shankar et al. 2008) while the socioeconomic divide in skilled birth attendance dropped sharply (Hatt et al. 2007).[32,33] Frankenberg et al. (2009) show that presence of a midwife increased pregnant women's intake of iron tablets and their uses of antenatal care and skilled as opposed to traditional

[31] For detailed descriptions of the program and of Indonesia's healthcare infrastructure see Sweet et al. (1995), Kosen and Gunawan (1996) and Frankenberg and Thomas (2001).

[32] In contrast to making access to skilled birth attendance more equitable, Hatt et al. (2007) show that the program coincided with a widening of the wealth and education gap between users and non-users of cesarean sections.

[33] The results by Achadi et al. (2007), Makowiecka et al. (2008), and Ensor et al. (2009) indicate that after the phasing out of government placement of midwives in the late 1990s, underservicing of remote communities and the socioeconomics gap in access to skilled birth attendance again increased during the 2000s.

birth attendance. Their results also indicate that these effects extended across the entire distribution of women's education. Further, Frankenberg et al. (2005) show that midwife presence greatly improved access to maternal and child care both through village midwives being the first healthcare provider to practice in a community, and, in communities where other providers were already present, by offering care at cheaper prices.

Against this background, various studies have examined midwife presence effects on the health of mothers and young children. Frankenberg and Thomas (2001) use data from the 1993 and 1997 ILFS waves to investigate effects on the Body-Mass-Index (BMI) of women of reproductive age as a proxy for maternal health. Their identification strategy compares the BMI-trend differences between women of reproductive age and men and older women in communities with midwives to the BMI-trend differences in communities without. The results indicate that midwife presence leads to statistically significant BMI-increases for women of reproductive age. The same study provides estimates of midwife effects on birth-weight. Controlling for community-level, time-invariant birth-weight determinants and a set of maternal characteristics, the authors find positive birth-weight effects. They, however, concede that their estimates may be upward biased because the subsample of Indonesian babies that is weighted at birth likely constitutes a positively selected group.

With data from four waves of the Indonesian Demographic and Health Survey (IDHS) between 1991 and 2002, Hatt et al. (2009) model time-trends in first-day and neonatal mortality, controlling for individual characteristics, including birth context variables like the type of birth attendance and the location of birth (e.g. at home or in a public health facility). They find no decrease in first-day mortality over the 1986-2002 period but a decrease in neonatal mortality by an annual average of 3.2 percent. This trend, however, appears to be independent of the midwife program's timing of implementation. Contrary to Hatt et al. (2009), Shrestha (2010) finds evidence for a midwife effect on child mortality. Exploiting variation in the timing of midwife placement across communities and controlling for district-level unobserved heterogeneity, she shows that high midwife density in a district is associated with lower neonatal but not lower post-neonatal mortality.

Finally, Frankenberg et al. (2005) use 1993 and 1997 IFLS data to analyze the midwife program's effect on young children's height-for-age. Their identification strategy bases on both spatial and over-time variation in community midwife presence and rules out OVB from time-invariant community-level characteristics through the inclusion of community fixed effects. Midwife presence in the first years of life, they find, is associated with increases in height-for-age that are particularly large among children of less-well educated mothers.

3.4 Data and Empirical Approach

My analysis uses data from the first four waves of the longitudinal IFLS (Frankenberg and Karoly 1995, Frankenberg and Thomas 2000, Strauss et al. 2004, Strauss et al. 2009). In the first wave of 1993 interviews were conducted with over 7,000 households from 312 enumeration areas in 13 of Indonesia's 27 provinces. Therewith, the survey was representative of 83 percent of the country's population. The subsequent waves (1997, 2000, and 2007) included interviews both with the original first wave households and with households founded by former first wave household members.

Because I examine infant midwife exposure effects on the physical and cognitive development of adolescents aged 10-14, the individual data in my empirical analysis come from the 1985-89 cohorts in the 2000 wave and the 1992-96 cohorts in the 2007 wave. My measure of long-run effects on physical development is the adolescents' height-for-age *z*-score (HAZ).[34] The HAZ compares the height of a sample adolescent to that of a well-nourished adolescent of the same age and sex from the U.S, as defined in the 2000 Center for Disease Control Growth Reference (Kuczmarski et al. 2000). It reports the adolescent's height deviation from the reference population's median divided by its standard deviation.[35] A HAZ two or more standard deviations below the reference median indicates *stunting*, i.e. severe growth faltering. Typical causes of stunting include mal- and undernutrition and exposure to infectious disease during early childhood and thus may be affected by midwife presence during infancy (Martorell and Habicht 1986, Frankenberg et al. 2005).

As a second outcome variable, I investigate the association between infant midwife exposure and adolescent weight-for-age *z*-scores (WAZ). Other than using weight instead of height, the WAZ method of calculation is equal to that of the HAZ. A deviation of two or more standard deviations from the reference median indicates underweight. In contrast to the HAZ as a long-run measure of child development, the WAZ reflects short- to mid-run nutritional and health conditions (World Health Organization 1986). Therefore, as the individuals in my sample are between 10 to 14 years of age, I expect no systematic infant midwife exposure effect on the WAZ.

I measure midwife exposure effects on the development of cognitive abilities by the adolescents' performance in IFLS's cognitive skill test for children aged 7-14 (Strauss et al. 2009).[36] The test comprises of two parts. The first is a shape matching exercise with twelve tasks. In the second, the child is asked to solve five mathematical tasks of ascending difficulty. Figures 3.1 and 3.2 of the Appendix sample one task from each part of test. The outcome for my empirical analysis is the total number of correct answers in complete tests to which I refer as the cognitive skill score (CSS) hereafter.

IFLS complements the household interviews with portrayals of the living conditions in the enumerated communities. This information is gathered in interviews

[34] In IFLS, a trained health worker takes height and weight measurements of each household member.
[35] I calculate *z*-scores using Stata's zanthro command (Vidmar et al. 2004). Using zanthro's cutoff option, I drop extreme observations with z-scores with absolute value over five from the sample.
[36] The tasks in the cognitive skill test are identical in the 2000 and 2007 waves.

Table 3.1: Introduction of midwives in sample communities 1993-97

Arrival year	Frequency	Percent	Cumulative
1993 or earlier	76	31.54	31.54
1994	18	7.47	39.00
1995	18	7.47	46.47
1996	11	4.56	51.04
1997	8	3.32	54.36
None 1993-97	110	45.64	100.0
Total	241	100.0	-

with key informants like community headmen and women group leaders. Among other things, it indicates if the community had a midwife at the time of the respective survey wave and, for communities with a midwife, retrospective midwife arrival year data. I use this information from the women group leader interviews in the 1993 and 1997 IFLS waves to reconstruct for the 1986-90 and 1993-97 periods whether or not a community had a program midwife.[37]

Dispatching of the first midwives started in 1990 but the program reached substantial coverage only in subsequent years. I therefore code all sample communities as not having a program midwife for the 1986-90 period. For the 1993-97 period, I obtain the community midwife status as follows. For all communities that report a midwife in the 1993 and 1997 waves, I assume that a midwife was present throughout 1993-97. Because this period is rather short, and because the initial government contracts lasted up to six years, it is unlikely that a substantial part of communities lost and regained a midwife in between. By the same argument, I assume that communities without a midwife in both the 1993 and 1997 waves did not have a midwife throughout 1993-97. For communities that report no midwife in 1993 but one in 1997, I use the arrival year data from the 1997 wave to determine the midwife status for every year in the 1993-97 period.

Likely due to recollection error, the arrival year data for 15 IFLS communities contradicts the midwife status information in the 1993 and 1997 waves. Such contradictions arise when the reported arrival year in the 1997 wave is 1992 or earlier, whereas the key informant in the 1993 wave declared that the community currently did not have a midwife. Moreover, there are 35 communities that report a midwife in the 1993 but none in the 1997 wave. Because there is no midwife departure year information in the 1997 wave, it is impossible to reconstruct midwife presence for each year in the 1993-97

[37] The 2000 and 2007 waves also include retrospective midwife arrival and departure year information. Cross-checks of the retrospective information between the four IFLS waves, however, indicate that recollection error accumulates over time and that arrival and departure year responses in later waves often only refer to the last midwife to arrive in the community. This makes a reconstruction of midwife presence status over more than two waves impossible.

period for these communities.[38] As my analysis relies on precise midwife presence information, I omit observations from the 15 communities with contradictory arrival year data and from the 15 communities that lost a midwife in 1993-97 from my main sample. Table 3.1 shows the timing of the midwife program's implementation for the 241 sampled communities.

I use the community midwife presence information for the 1986-90 and 1993-97 periods to obtain whether or not a midwife was present in the year after an individual's birth year. My sample consists of the individuals for which this information is available, namely of the 1985-89 and 1992-96 ILFS cohorts.

My identification of midwife effects bases on both the spatial differences in the timing of midwife introduction and the assumption that the program primarily affected the development of children that were very young at the time of midwife exposure. The rationale for this assumption is twofold. The first and main rationale comes from the biological and medical literatures. It is now established that the role of nutrition and disease – as primary fields of midwife activity – in shaping long-run outcomes is particularly crucial in the period from conception to the 24th month after birth (for a review, see Victoria et al. 2008).[39] The second rationale bases on program design: because pregnant women and newborn children were its primary target group, I consider large program impacts on children of school age unlikely. Thus, I use midwife exposure in the year after an individual's birth year – a period that depending on the birth date extends from the 1st to the 24th month of life – as a cutoff point to separate treatment and control groups, and estimate

$$y_i = cons + \beta_{MW} \cdot MW_i + BY_{i,1985\text{-}89,1992\text{-}96} + Sex_i + MAT_{i,EDUC,\,HEIGHT} + CB_{i,1\text{-}241} + \varepsilon_i \,. \quad (1)$$

The vector of dependent variables y_i represents the three outcome variables for individual *i*. For the HAZ and WAZ, I estimate equation (1) with an Ordinary Least Squares (OLS) model and for the CSS as a count outcome using a Poisson model.[40]

On the right hand side of (1), *cons* is a constant and MW_i the explanatory variable of interest. MW_i is a binary variable that equals one if a midwife was present in an individual's birth community in the year after the individual birth year and zero if not. MW_i 's coefficient β_{MW} thus measures the association between midwife presence in the

[38] The 2000 wave is the first that includes midwife departure year data. The data is, however, only raised for communities that had a midwife during the 1997 but none during the 2000 wave. Therefore, these communities are not the same as the communities that lost a midwife between 1993 and 1997.

[39] Further support for this assumption comes from one of the first studies on long-run effects of an early-life nutritional intervention: Hoddinott et al. (2008) show that a Guatemalan child feeding program was effective in increasing hourly wages for males with program exposure between the first and 24th month of life, but not for males exposed at later ages.

[40] I find no indication for over-dispersion or excess zeros for the CSS variable that would motivate the use of negative binominal or zero-inflated models: testing the model form using Stata's estat gof command, the goodness-of-fit chi-squared test is not statistically significant for any of the models for which I present results in Tables 3.3-3.6. In a robustness test, I obtain OLS estimates of the CSS effects. The estimates are almost identical in magnitude and statistical significance for all specifications in Tables 3.3-3.6. Table 3.6, column (4) shows OLS estimates for the CSS, the shape matching score, and the mathematical score for the full sample. Further OLS results can be provided upon request.

year after the birth year and the outcomes y, holding the control variables in (1) constant. Hereafter, I refer to midwife exposure in the year after the birth year as *infant midwife exposure*.

Equation (1) includes the following control variables. For each of the 1985-89 and 1992-96 birth years except a reference year, $BY_{i,1985\text{-}89,1992\text{-}96}$ represents an indicator that equals one if an individual was born in the respective year, and zero otherwise. These birth year fixed effects control for all determinants of the outcomes y that are constant across entire birth cohorts, for instance, the age at which a sample cohort experienced the 1997 Asian Financial Crisis and its aftermath. The birth year fixed effects also control for the age at which I observe the outcomes. This is important because unlike the HAZ and WAZ, the CSS is not age-standardized.

The indicator Sex_i equals one if an individual is male and controls for sex differences in outcome levels. $MAT_{i,EDUC,\ HEIGHT}$ represents two control variables: an indicator that equals one if an individual's mother has higher than primary education, and a variable indicating the mother's height in centimeters. By controlling for maternal characteristics, I address OVB that arises if mothers of children with infant midwife exposure differ from mothers of children without. As discussed in Section 3.3, there is evidence that the midwife program reduced child mortality (Shresta 2010) and lead to larger health improvements for children of poorly educated mothers (Frankenberg et al. 2005). If this is true for my sample, there are more children with poorly educated mothers in the cohorts with infant midwife exposure than in the cohorts without. If, as shown in earlier work, poor maternal education correlates with lower levels of child development (e.g. Wachs 2005), not controlling for maternal education would cause downward bias in my estimates of the program's long-run effects. The inclusion of maternal height and education as predictors of child development aims to mitigate such concerns.

The birth community fixed effects $CB_{i,1\text{-}241}$ form the final set of controls in (1). For each of the 241 sampled communities except a reference community, $CB_{i,1\text{-}241}$ represents an indicator that equals one if an individual was born in the respective community, and zero otherwise. These community fixed effects control for birth-community-specific determinants of the outcomes y that are constant over the 1985-97 period. With the inclusion of these controls I aim to capture the non-random midwife placement mechanism I discuss in Section 3.2. Because this midwife placement mechanism served poor communities first, and because poverty negatively correlates with child development, not controlling for the determinants of midwife placement would cause OVB, i.e. a correlation of MW_i and the error term ε_i, and thus an understatement of the true midwife effects.[41]

[41] See Angeles et al. (1998) for a more general discussion of biases that can arise in the evaluation of non-randomly placed programs.

Table 3.2: Descriptive statistics – full sample and subsample of communities gaining a midwife in the 1986-97 period

	(1) Full sample		(2) Communities gaining midwife	
	Mean	S.D.	Mean	S.D.
Dependent variables				
Height-for-age *z*-score	-1.67	1.06	-1.76	1.03
Weight-for-age *z*-score	-1.53	1.29	-1.62	1.26
Cognitive skill score	12.38	3.11	12.04	3.18
Independent variables				
= 1 if midwife during infancy	.25	.43	.39	.49
= 1 if male	.48	50	.47	.50
Birth year	1990.60	3.77	1990.49	3.75
Maternal height (cm)	150.33	6.86	150.27	6.53
= 1 if mother has higher than primary ed.	.34	.47	.27	.44
# of observations	3,306		2,075	

3.5 Results

Table 3.2 shows descriptive statistics for the three outcomes and the independent variables in equation (1). The main sample consists of 3,306 individuals for whom no data for the variables in equation (1) are missing. Column (1) reports means and standard deviations for the full sample, and column (2) for the subsample of 2,075 individuals that come from the 131 communities that gained a midwife in the 1986-97 period. In the full sample, 25 percent of individuals had infant midwife exposure – they form the treatment group. The control group consists of the 75 percent without.

For all three outcomes, the full sample means in column (1) are larger than those in the subsample of midwife-gaining communities in column (2). This, together with the observation that mothers from midwife-gaining communities are on average shorter and less-well educated reflects the non-random midwife placement mechanism I discuss in Section 3.2 and underscores the necessity of accounting for this mechanism in the empirical model.

Tables 3.3-3.6 show my estimates of midwife effects for different samples and model specifications. To account for spatial autocorrelation in the model residuals, I cluster standard errors at the birth-community-level in all estimations (Bertrand et al. 2004).

Table 3.3, column (1) shows full sample estimates of the midwife effect for a specification that controls for birth year and sex, column (2) for a specification that in

Table 3.3: Midwife presence during infancy and adolescent outcomes – full sample

		(1)	(2)	(3)
Height-for-age *z*-score				
= 1 if midwife during infancy	β_{MW}	-.13	.14**	.14**
	S.E.	[.08]	[.07]	[.06]
Birth community		No	Yes	Yes
Maternal height and education		No	No	Yes
Weight-for-age *z*-score				
= 1 if midwife during infancy	β_{MW}	-.21*	.00	.00
	S.E.	[.09]	[.08]	[.08]
Birth community		No	Yes	Yes
Maternal height and education		No	No	Yes
Cognitive skill score				
= 1 if midwife during infancy	β_{MW}	-.040***	.046**	.047***
	S.E.	[.014]	[.018]	[.018]
Birth community		No	Yes	Yes
Maternal height and education		No	No	Yes

Notes: all models control for sex and birth year; height-for-age and weight-for-age *z*-score models estimated using OLS, and cognitive skill score model using the Poisson estimator; standard errors with clustering at the birth community level reported in brackets; N = 3,306; ***, **, * signify statistical significance on the 1, 5, and 10 percent levels, respectively.

addition controls for birth community fixed effects, and column (3) for the full specification of equation (1) that moreover controls for maternal height and education.

In column (1), the midwife coefficient estimates have a negative sign for all three outcomes and are statistically significant for the WAZ and the CSS, suggesting that infant midwife exposure impaired adolescent outcomes. Rather than providing evidence for perverted midwife effects, this result reflects the fact that the placement mechanism favored poor communities, causing an overrepresentation of individuals from communities with inferior developmental prospects in the treatment as compared to the control group. Not controlling for this in the specification in column (1) introduces downward OVB into the estimates of the midwife effects. In line with this argument, the midwife effect estimates increase for all outcomes when I include the birth community fixed effects in the specification in column (2).

Moreover, they change only slightly when I, in addition, include controls for maternal education and height in column (3). The latter finding indicates that the midwife program caused no or only small changes in the socioeconomic patterns of fertility and child mortality. Because of the small magnitude of differences between the effects in columns (2) and (3), I limit the following discussion of the midwife effects' sizes to the full specification results in column (3).

Table 3.4: Midwife presence during infancy and adolescent outcomes – by sex

		(1)	(2)
		Female	**Male**
Height-for-age *z*-score			
= 1 if midwife during infancy	β_{MW}	.18*	.09
	S.E.	[.11]	[.10]
Weight-for-age *z*-score			
= 1 if midwife during infancy	β_{MW}	.06	-.02
	S.E.	[.13]	[.12]
Cognitive skill score			
= 1 if midwife during infancy	β_{MW}	.066**	.038*
	S.E.	[.026]	[.022]
# of observations		1,589	1,717

Notes: all models control for birth year, birth community, maternal height and education; height-for-age and weight-for-age *z*-score models estimated using OLS, and cognitive skill score model using the Poisson estimator; standard errors with clustering at the birth community level reported in brackets; **, * signify statistical significance on the 5 and 10 percent levels, respectively.

In this specification, infant midwife exposure is associated with a statistically significant .14 point increase in the HAZ, corresponding to an 8.4 percent increase towards the sample mean. This result indicates long-run improvements in physical development through the midwife program. Presence of a midwife during infancy is also associated with a 4.8 percent higher CSS.[42] In contrast, I find no association of infant midwife exposure with the WAZ as a measure of short- and mid-run health conditions, suggesting that the HAZ and CSS effects are in fact driven by early- rather than later-life circumstances.

Table 3.4 shows infant midwife exposure effects for the female subsample in column (1) and for the male subsample in column (2). There is no statistically or economically significant effect on the WAZ for either sex. For both long-run outcomes, however, the results indicate substantially larger program effects on females. Female infant midwife exposure is associated with a .18 point increase in the HAZ, and a 6.8 percent increase in the CSS. The male HAZ effect, in contrast is just .09 points and not statistically significant, whereas the male CSS effect is 3.9 percent and statistically significant at the 10 percent level. This finding may reflect parental son preference in access to early-life nutrition and healthcare in Indonesia. The midwife program lowered the financial and time barriers for access to adequate early-life nutrition and care. If, other things equal, Indonesian parents are willing to invest more in male offspring, and if the

[42] For the CSS, I obtain the effects in percent by the IRR option in Stata's poisson estimation command. IRR exponentiates the Poisson model coefficients.

Table 3.5: Midwife presence during infancy and adolescent outcomes – by maternal education

		(1) Maternal education primary of lower	(2) Maternal education higher than primary
Height-for-age *z*-score			
= 1 if midwife during infancy	β_{MW}	.20**	.05
	S.E.	[.09]	[.12]
Weight-for-age *z*-score			
= 1 if midwife during infancy	β_{MW}	-.04	.13
	S.E.	[.11]	[.16]
Cognitive skill score			
= 1 if midwife during infancy	β_{MW}	.047*	.011
	S.E.	[.025]	[.024]
# of observations		2,188	1,118

Notes: all models control for sex, birth year, birth community, and maternal height; height-for-age and weight-for-age *z*-score models estimated using OLS, and cognitive skill model using the Poisson estimator; standard errors with clustering at the birth community level reported in brackets; **, * signify statistical significance on the 5 and 10 percent levels, respectively.

developmental benefits of early-life investments are downward sloping, this may explain why the lowering of access barriers benefited females more than males.

Table 3.5 splits the sample by levels of maternal education. Column (1) shows infant midwife exposure effects for individuals with mothers with no formal or primary education, and column (2) for individuals with mothers with higher than primary education. By splitting the sample this way, I examine if children from different societal strata benefited differently from infant exposure to the midwife program.

The comparison of results from columns (1) and (2) indicate that the midwife effects on child development are limited to individuals with poorly educated mothers. Among them, infant midwife exposure is associated with a .2 point higher HAZ and a 4.8 percent higher CSS. Both effects are about four times larger than the not statistically significant HAZ and CSS effects I find for individuals with better educated mothers in column (2). These findings are in line with those of Frankenberg et al. (2005) who find that early-life midwife presence leads to larger HAZ improvements in young children if their mothers have no formal of just primary education. They indicate that the midwife program contributed to the reduction of social inequities in access to adequate early-life nutrition and healthcare, and thus to greater equity in the developmental opportunities of children of different socioeconomic backgrounds.

Table 3.6 provides four robustness tests for the full sample results in Table 3.3. All results in the table come from models that use the full set of controls in equation (1).

Table 3.6: Midwife presence during infancy and adolescent outcomes – robustness tests

		(1) Imputed midwife arrival year data	(2) Midwife arrival year data from community headman	(3) Shape matching and math. scores separately	(4) OLS estimates of count outcomes
Height-for-age *z*-score					
= 1 if midwife during infancy	β_{MW}	.18***	.15**	-	-
	S.E.	[.56]	[.07]		
Weight-for-age *z*-score					
= 1 if midwife during infancy	β_{MW}	.10	.03	-	-
	S.E.	[.08]	[.08]		
Cognitive skill score					
= 1 if midwife during infancy	β_{MW}	.037**	.048***	-	.53**
	S.E.	[.015]	[.019]		[.023]
Shape matching score					
= 1 if midwife during infancy	β_{MW}	-	-	.039**	.31*
	S.E.			[.019]	[.17]
Mathematical score					
= 1 if midwife during infancy	β_{MW}	-	-	.070***	.22**
	S.E.			[.028]	[.09]
# of observations		3,931	3,485	3,306	3,306

Notes: all models control for sex, birth year, birth community, and maternal height and education; height-for-age and weight-for-age *z*-score models estimated using OLS, and cognitive skill, shape matching, and mathematical score models in columns (1)-(3) using the Poisson estimator; cognitive skill, shape matching, and mathematical score models in columns (4) estimated using OLS; standard errors with clustering at the birth community level reported in brackets; ***, **, * signify statistical significance on the 1, 5 and 10 percent levels, respectively.

Column (1) shows midwife effect estimates for a sample that includes 625 additional observations from the 15 communities for which midwife arrival year data in the 1993 and1997 IFLS waves is contradictory, and from the 35 communities that lost a midwife between the 1993 and 1997 waves. To include these observations in the sample, I impute the contradictory and missing midwife presence data as follows: for communities that gained midwife between the 1993 and 1997 waves but for which the arrival year in the 1997 wave is earlier than 1993, and thus contradictory, I assume that the midwife arrived in 1994. For communities that lost a midwife between the 1993 and 1997 waves, I assume

that the midwife left in 1995, i.e. that the 1992-94 cohorts had infant midwife exposure, while the 1995 and 1996 cohorts did not.

The results for the extended sample in column (1) of Table 3.6 are qualitatively similar to those for the main sample in column (3) of Table 3.3. The HAZ effect is somewhat larger, and the CSS effect somewhat smaller for the imputation sample in Table 3.6. There is no statistically significant effect on the WAZ for either sample. In summary, these findings indicate that the results in Table 3.3 are not driven by the omission of communities with contradictory or incomplete midwife presence data.

My analysis in Tables 3.3-3.5 uses midwife presence and arrival year data from interviews with women group leaders. IFLS also obtains this information from community headmen. My first rationale for using women group data for my main analysis is the assumption that because maternal and child health are main fields of activity for women groups, their leaders provide more precise arrival year data. The second rationale comes from the survey literature. Unlike the women group arrival year data, the headmen data show a heaping of arrivals in 1990 when the dispatching of midwives only started and had not reached the rates of later years. Against this background, such heaping is likely the result of recollection error in an environment where written records are rarely kept (Beckett et al. 2001). The correlation of arrival years from the two sources is nevertheless high (.76) and the majority of differences is in pre-1993 years. Therefore, because I use arrival year data only for 1993-97, no substantial differences should exist between models that use the women group and headmen data to construct the infant midwife exposure variable. Column (2) of Table 3.6 tests for this by showing results for the headmen arrival year data. In line with the prediction, the estimates indicate qualitatively similar effects for both data sources, with the headmen data effects being just marginally larger.

Column (3) of Table 3.6 shows infant midwife exposure effects I obtain for the outcomes of the two parts of the cognitive skill test separately. If the educational infrastructure improved more for individuals with infant midwife exposure than for individuals without, this, rather than midwife exposure may cause the increase I measure for the CSS. If the association between infant midwife exposure and CSS is in fact spurious, I should find little or no effect of midwife exposure on the outcome of the shape matching exercise that I assume to depend less on formal schooling. The results in column (3) do not support this hypothesis. While the effect on the mathematical skill score is larger at 7.2 percent, the shape matching exercise effect is 3.9 percent and statistically significant. Rather than providing evidence for a spurious correlation of infant midwife exposure and cognitive skill, the fact that the mathematical skill score is larger permits the tentative interpretation that the midwife effect on cognitive skill works through two channels. First, the effect on performance in the shape matching exercise indicates a direct midwife effect on cognitive skill formation through better brain development. Second, the larger mathematical skill score effect suggests that midwife presence not only increased school performance through better cognitive ability but that better child health had an

additive effect on performance through, for instance, higher school attendance (Case and Paxson 2006).

Finally, Table 3.6, column (4) shows full sample estimates of midwife effects on the cognitive outcomes that I obtain using OLS instead of Poisson estimation. While the distributions of the CSS and the shape matching and mathematical scores justify the use of a count model, OLS estimation provides a test for the robustness of the above findings towards a different model specification. The estimates in column (4) suggest that the cognitive skill effects that I obtain by Poisson models are not driven my choice of a count model. Instead, the .53, .31, and .22 point reductions in the CSS, shape matching, and mathematical scores I obtain by OLS are statistically significant and almost of identical size (4.4 percent, 3.9 percent, and 7.3 percent) as the corresponding Poisson model estimates.

3.6 Conclusions

This work examines effects of infant exposure to a large scale Indonesian midwife program on height and cognitive skill at age 10-14. Controlling for cohort and time-invariant birth community effects, maternal characteristics, and sex, I find economically and statistically significant improvements in these long-run outcomes for individuals with infant midwife exposure. Splitting the sample by sex and maternal education reveals that the effects are larger for girls and restricted to individuals with poorly educated mothers. This indicates that the program contributed to the reduction of both sex and social inequalities in child development.

Frankenberg et al. (2005) examine channels by which early-life midwife exposure may lead to improvements in the height-for-age *z*-scores of young children that may also apply for the formation of adolescent outcomes. They find that midwife presence leads to longer exclusive breastfeeding and higher uptake of prenatal care. Both have been shown to affect child development, including in the long-run (Frongillo and Habicht 1997, Onyango et al. 1998, Walker et al. 2007). Lacking data impedes the identification of other impacts of midwife presence on early-life health and nutrition conditions. However, with the broad set of midwife activities outlined in Section 3.2, other channels, for instance, through better hygiene that reduces the prevalence of early-life diarrhea, are likely.

A causal interpretation of the association between infant midwife exposure and adolescent outcomes is warranted if no unobserved phenomenon of equal spatial distribution and timing as the midwife program exists. For 1985-97, I am not aware of any government or donor interventions on early-life conditions that were of similar scope. It is, however, possible that the rapid economic expansion that Indonesia experienced between the late 1980s and the 1997 crisis caused particularly large improvements in early-life health conditions in the poorest communities that were also targeted by the midwife program. With no community- or sub-district-level data on economic conditions for 1985-97, I therefore cannot rule out that the associations I present in Section 3.5 are in part driven by community specific growth trends.

The inclusion of controls for maternal education and height addresses this issue insofar as improvements in parental backgrounds are proxies for lower child mortality and therefore reflect improvements in economic conditions. Another argument against a spurious correlation of infant midwife exposure and child development is that the program only started in the 1990s. The fact that the program prioritized communities that by this time were still underdeveloped makes it unlikely that these communities were the ones that experienced particularly large improvements in early-life living conditions in the 1985-97 period. Finally, if the midwife effect was in fact an economic growth effect, I should also measure significant improvements in the WAZ, as communities with faster economic growth would also show larger relative improvements in more short-run determinants of body size. While these counterarguments may alleviate some of the concerns about a spurious correlation, only explicit controls for local gross domestic product or better proxies thereof can provide a definite answer. Future investigations may address this.

Equally, it is of interest why the program apparently benefited girls more than boys. While son preference may offer an explanation, this interpretation is at odds with the finding that unlike for other East-Asian countries, there is little evidence for son preference for Indonesia (Kevane and Levine 2003).

If data from future IFLS waves permit the tracking of individuals with early-life midwife exposure into adulthood, future work may examine impacts on education, labor market performance, and adult health. Contributing to a full appraisal of program benefits, this forms the basis for comprehensive cost-benefit-analysis and subsequent assessments of the program's transferability into other country contexts. The Indonesian midwife program is one of many community-level interventions in developing countries to improve early-life health and nutrition (Claeson and Waldman 2000). Practitioners highlight that more studies on their mid- and long-run impacts are needed to single out the most efficient ones (Alderman 2011).

The need for cost-efficient, large-scale interventions to improve early-life living conditions is urgent in a time of global economic disruption. Despite tightening budgets, developing countries cannot afford to risk the productive potential of more generations. This work provides further evidence that the focusing of government and donor efforts on child development has high rates of return – not only for the individual but for the competitiveness and development of societies as a whole.

References

Achadi, E., S. Scott, E. Pambudi, K. Makowiecka, T. Marshall, A. Adisasmita, P. Deviany, and C. Ronsmans. 2007. "Midwifery provision and uptake of maternity care in Indonesia". *Tropical Medicine and International Health*, 12(12): 1490-1497.

Alderman, H. 2011. "Introduction". In: Alderman, Harold (ed.): No Small Matter - The Impact of Poverty, Shocks, and Human Capital Investments in Early Childhood Development, The World Bank, Washington D.C.: 9-50.

Almond, D. and J. Currie. 2011. "Human capital development before age 5". In: Orley Ashenfelter and David Card (eds.): Handbook of Labor Economics, vol. 4B, Elsevier, Amsterdam: 1315-1486.

Angeles, G., D. Guilkey, and T. Mroz. 1998. "Purposive Program Placement and the Estimation of Family Planning Program Effects in Tanzania". *Journal of the American Statistical Association*, 93(443): 884-899.

Beckett, M., J. Da Vanzo, N. Sastry, C. Panis, and C. Peterson. 2001. "The Quality of Retrospective Data: An Examination of Long-run Recall in a Developing Country". *The Journal of Human Resources*, 36(3): 593-625.

Bertrand, M., E. Duflo, and S. Mullainathan. 2004. "How much should I trust differences-in-differences estimates?". *The Quarterly Journal of Economics*: 119(1): 249-275.

Case, A. and C. Paxson. 2006. "Children's health and social mobility". *Opportunity in America*, 16(2): 151-173.

Claeson, M. and R. Waldman. 2000. "The evolution of child health programmes in developing countries: from targeting diseases to targeting people". *Bulletin of the World Health Organization*, 78(10): 1234-1245.

Ensor, T., Z. Quayyum, M. Nadjib, and P. Sucahya. 2009. "Level and determinants of incentives for midwives in Indonesia". *Health Policy and Planning*, 24: 26-35.

Frankenberg, E. and L. Karoly. 1995. "The 1993 Indonesian Family Life Survey: Overview and Field Report". RAND, Santa Monica, CA.

Frankenberg, E. and D. Thomas. 2000. "The Indonesia Family Life Survey (IFLS): Study Design and Results from Waves 1 and 2". RAND, Santa Monica, CA.

Frankenberg, E. and D. Thomas. 2001. "Women's Health and Pregnancy Outcomes: Do Services Make a Difference?". *Demography*, 38(2): 253-265.

Frankenberg, E., W. Suriastini and D. Thomas. 2005. "Can expanding access to basic healthcare improve children's health status? Lessons from Indonesia's 'midwife in the community' programme". *Population Studies*, 59(1): 5-19.

Frankenberg, E., A. Buttenheim, B. Sikoki, and W. Suriastini. 2009. "Do Women Increase Their Use of Reproductive Health Care When It Becomes More Available? Evidence from Indonesia". *Studies in Family Planning*, 40(1): 27-38.

Frongillo, E. and J. Habicht. 1997. "Investigating the weanling's dilemma: lessons from the Honduras". *Nutrition Reviews*, 55(11): 390-395.

Grantham-McGregor S., Y. Cheung, S. Cueto, P. Glewwe, L. Richter, B. Strupp, and International Child Development Steering Group. 2007. "Developmental potential in the first five years for children in developing countries." *Lancet*, 369: 60-70.

Hatt, L., C. Stanton, K. Makowiecka, A. Adisasmita, E. Achadi, and C. Ronsmans. 2007. "Did the strategy of skilled attendance at birth reach the poor in Indonesia?". *Bulletin of the World Health Organization*, 85: 774–782.

Hatt, L., C. Stanton, C. Ronsmans, K. Makowiecka, and A. Adisasmita. 2009. "Did professional attendance at home births improve early neonatal survival in Indonesia?". *Health Policy and Planning*; 24: 270-278.

Heckman, J. and D. Masterov. 2007. "The Productivity Argument for Investing in Young Children". *Review of Agricultural Economics*, 29(3): 446-493.

Hoddinott, J., J. Maluccio, J. Behrman, R. Flores, and R. Martorell. 2008. "Effect of a Nutrition Intervention during Early Childhood on Economic Productivity in Guatemalan Adults". *Lancet*, 371: 411-416.

Kevane, M. and D. Levine. 2003. "Changing Status of Daughters in Indonesia". Center for International and Development Economics Research, Working Paper Series 3171.

Kosen, S. and S. Gunawan. 1996. "Health services in Indonesia." *Medical Journal of Australia*, 165(11/12): 641-644.

Kuczmarski, R., C. Ogden, L. Grummer-Strawn, K. Flegal, S. Guo, R. Wei, Z. Mei, L. Curtin, A. Roche, and C. Johnson. 2000. "CDC growth charts: United States." *Advance Data*, 314: 1-27.

Makowiecka, K., E. Achadi, Y. Izati, and C. Ronsmans. 2008. "Midwifery provision in two districts in Indonesia: how well are rural areas served?". *Health Policy and Planning*, 23: 67-75.

Martorell, R. and J. Habicht. 1986. "Growth in early childhood in developing countries". In: Falkner, Frank and J. M. Tanner (eds.): Human Growth: A Comprehensive Treatise, Volume 3, New York, Plenum Press: 241-262.

Ministry of Health. 1997. "Indonesia: health profile". Centre for Health Data, Jakarta.

Ministry of Health. 2000. "Indicators for the third evaluation of HFA 2000". Government of Indonesia, Jakarta.

Naudeau, S., S. Martinez, P. Premand, and D. Filmer. 2011a. "Cognitive Development among Young Children in Low-Income Countries". In: Alderman, Harold (ed.): No Small Matter - The Impact of Poverty, Shocks, and Human Capital Investments in Early Childhood Development, The World Bank, Washington D.C.: 9-50.

Naudeau, S., N. Kataoka, A. Valerio, M. Neuman, and L. Elder. 2011b. Investing in young children: an early childhood development guide for policy dialogue and project preparation. The World Bank, Washington DC.

Onyango, A., K. Koski, and K. Tucker. 1998. "Food diversity versus breastfeeding choice in determining anthropometric status in rural Kenyan toddlers". *International Journal of Epidemiology*, 27(3): 484-489.

Shankar, A., S. Sebayang, L. Guarenti, B. Utomo, M. Islam, V. Fauveau, and F. Jalal. 2008. "The village-based midwife programme in Indonesia". *Lancet*, 371(12): 1226-1229.

Shresta, R.. 2010. "The Midwife Program and Infant Mortality in Indonesia". *Bulletin of Indonesian Economic Studies*, 46(2): 193-211.

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

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

Sweet, B., V. Tickner, and G. Maclean. 1995. "Midwifery in Indonesia: A professional snapshot." *Modern Midwife*, 5(6): 8-13.

Victoria, C., L. Adair, C. Fall, P. Hallal, R. Martorell, L. Richter and H. Singh Sachdev. 2008. "Maternal and child undernutrition: consequences for adult health and human capital." *Lancet*, 371: 340-357.

Vidmar, S., J. Carlin, and K. Hesketh. 2004. "Standardizing anthropometric measures in children and adolescents with new functions for egen". *The Stata Journal*, 4(1): 50-55.

Wachs, T. 2005. "Linking Nutrition and Education: a Cross-Generation Model". *Food and Nutrition Bulletin*, 26: 159-167.

Walker S., S. Chang, C. Powell, and S. Grantham-McGregor. 2005. "Effects of early childhood psychosocial stimulation and nutritional supplementation on cognition and education in growth-stunted Jamaican children: prospective cohort study". *Lancet*, 366: 1804-1807.

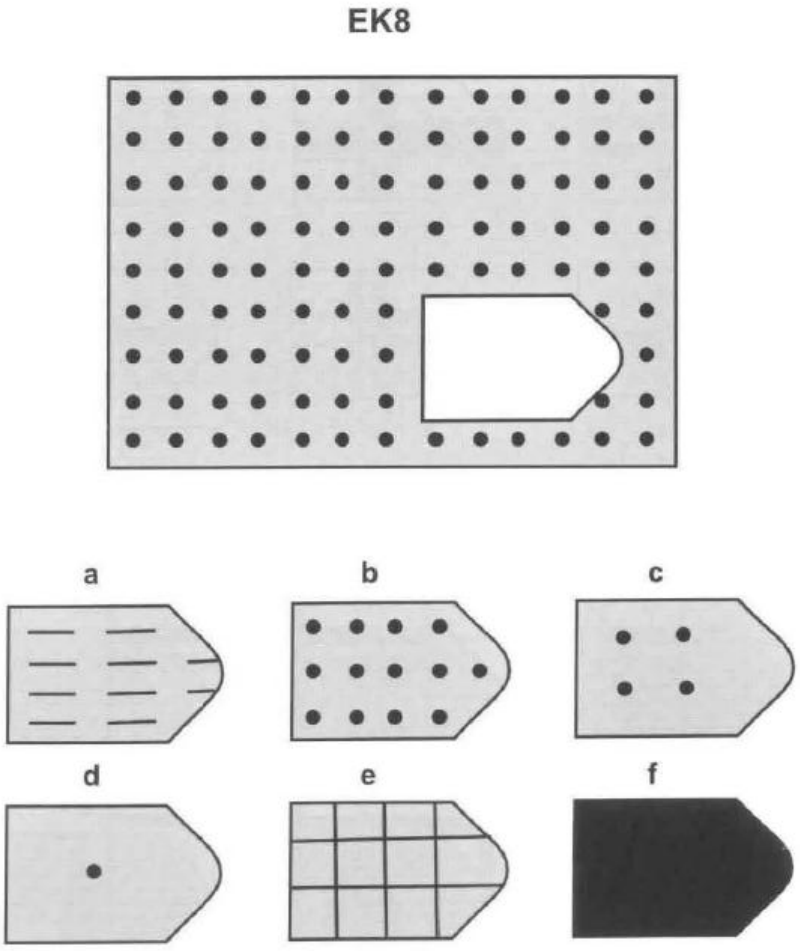
Walker S., T. Wachs, J. Meeks Gardner, B. Lozoff , G. Wasserman, E. Pollitt, J. Carter, and the International Child Development Steering Group. 2007. “Child development: risk factors for adverse outcomes in developing countries.” *Lancet*, 369: 145-157.

World Bank. 1991. “Staff Appraisal Report Indonesia: Fifth Population Project (Family Planning and Safe Motherhood)”. The World Bank, Washington D.C.

World Health Organization. 1986. “Use and Interpretation of Anthropometric Indicators of Nutritional Status”. *Bulletin of the World Health Organization*, 64(6): 929-941.

Appendix

Figure 3.1: Sample task in IFLS shape matching test



Source: Strauss et al. (2009)

Figure 3.2: Sample task in IFLS mathematical test

EK16. 56/84 = . . .

a. 4/7

b. 2/3

c. 3/4

d. 5/6

Source: Strauss et al. (2009)

Akademischer Lebenslauf

April 2009 – Mai 2012	Promotionsstudium an der Ludwig-Maximilians-Universität München und Tätigkeit als wissenschaftlicher Mitarbeiter am ifo-Institut für Wirtschaftsforschung München
April 2010 – Juni 2010	Forschungsaufenthalt an der George Mason University in Arlington (Virginia), Vereinigte Staaten von Amerika
Oktober 2001 – April 2008	Diplomstudium Volkswirtschaftslehre sozialwiss. Richtung (Schwerpunkt Politikwissenschaft) an der Universität zu Köln
May 2000	Abitur am Gymnasium am Neanderthal, Erkrath
26.11.1980	Geboren in Düsseldorf