Empirical Essays on the Effects of Early Life Conditions on Health Later in Life

Inauguraldissertation zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaften
der Universität Mannheim

Johannes Christian Schoch

vorgelegt im Mai 2014

Abteilungssprecher:	Prof. Dr. Eckhard Janeba
Referent:	Prof. Gerard van den Berg, Ph.D.
Korreferentin:	Prof. Dr. Andrea Weber

Verteidigung: 15. Juli 2014

Für meine Eltern

Acknowledgements

There are many people who helped me throughout the last years while I was writing this dissertation. First and foremost, I want to thank my parents, my brother, and my sister who supported me constantly and were never tired of listening to whatever problem I had or thought I had in the last years. The following phrase may be somewhat overused but nevertheless holds true: without you, I would not have made it.[1]

Second, I am indebted to my supervisor and co-author Gerard van den Berg for his patience, encouragement, and guidance. Furthermore, I want to thank my co-advisor Andrea Weber for many constructive comments on my research projects. I am also indebted to my co-authors and colleagues Anna Hammerschmid, Pia Pinger, and Katharina Walliczek for many useful discussions and their effort. I also want to thank the colleagues and seminar participants who helped me to improve my work, in partciular my colleagues from the Alexander von Humboldt Chair in Econometrics and Empirical Economics and the econometrics group at our department.

Last, there were many people who played a very important role at the intersection of university and private life or in my private life alone. I have to thank: my friends and 'partners in crime', Philipp Leibold and André Stenzel for advise and welcome diversion; my colleagues and friends Hannes Kammerer, Vera Molitor, and Michèle Weynandt for support, discussions about important and less important topics, lunch and coffee breaks, and everything else that brightened up my daily routine; Christian Michel for many interesting discussions about the peculiarities of life as well as cultural and philosophical matters; my friends from back in the days, most importantly Jonathan Fabarius, Simon Liegat, and Peter Popella, for backing me up and making me laugh tears at times; and every person accompanying me during the last years.

[1] Eine sinngemäße Übersetzung für meine Eltern: Ich möchte an erster Stelle meinen Eltern, meinem Bruder und meiner Schwester dafür danken, dass sie mich stets unterstützt haben und immer ein offenes Ohr hatten, wenn ich ein Problem hatte oder mir ein Problem einredete. Der folgende Satz mag ein wenig abgedroschen klingen, er drückt dennoch meine tiefste Überzeugung aus: Ohne Euch hätte ich es nicht geschafft.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This dissertation analyses a topic that is nowadays widely studied in economics and other social sciences, but originated from the medical literature: The long-run health effects of early life conditions. Broadly speaking, the underlying theoretical foundation is a model of 'fetal programming' which suggests that the fetal metabolism and organ development may react to outside influences via adaptation and epigenetic modifications (for instance, to guarantee survival up to reproductive age). For newborns, a similar mechanism has been suggested. Eventually, these adaptations may make an individual who has been exposed to adverse early life conditions more likely to suffer from, for instance, cognitive deficits, chronic diseases or reduced longevity. Empirical studies so far have found evidence for interrelations between a wide range of early life circumstances and later life health outcomes but have oftentimes failed to establish a robust causal link.

As economists, we are interested in the connection between conditions early in life and later life health mostly because of the potentially severe economic consequences: poor adult health outcomes are determinants of poor economic outcomes and productivity. Hence, early life conditions may be influential drivers of economic inequality and are thus of individual and societal interest. Moreover, it may be more efficient to compensate adverse early life conditions by adequate policies early on than to medicate later life health conditions or even out deficits in productivity that are related to those bad health outcomes. Knowledge about the exact risk factors and their impact is thus crucial for advising policy makers. As econometricians we take an interest in the methods employed to measure a statistical association or empirical effect since they determine the very nature of such an effect and thus crucially affect the relevance and external validity of an empirical result. This dissertation contributes both new findings and econometric insights to the existing literature.

Chapter 2, a joint project with Gerard van den Berg, starts out by using business cycle fluctuations around the time of birth of individuals as exogenous variation (i.e. variation beyond the control of an individual or not determined by individual level unobservables that also impact on the outcome of interest) in early life conditions. Specifically, we ask the question how such shocks influence the quantitative impact of later life events such as bereavement or the onset of diseases on the trajectory of cognitive and mental health outcomes. By employing panel data models, we measure causal effects of such events under relatively mild assumptions. We then differentiate the effects of these events along the state of the business cycle around birth. Our main finding is that the adverse effect of stroke on cognitive skills of women is cushioned if an individual was born under good economic conditions.

Business cycles are a recurring macro-economic phenomenon, unlike other exceptional and more extreme sources of exogenous variation in early life conditions used in the literature such as famines or other disastrous events. Moreover, since we use data from various European countries, we estimate average effects that are less context dependent and thus of more general relevance. An important question, however, remains unanswered: what signifies business cycle downturns on an individual level and what is it that drives the effects found in this chapter and in studies following a similar approach? Since we lack individual-level variables that are determined by business cycle shocks but only perform 'reduced form' analyses, we cannot isolate the exact drivers of our results.[1] Hence, it is not possible to give policy advice as to what particular compensatory action may be adequate to make up for adverse conditions that were induced by a bad economic environment around birth. Chapters 3 and 4 advance on the problem of limited informational content of reduced form effects in a different context and in two different dimensions.

The third chapter presents a study co-authored with Gerard van den Berg, Anna Hammerschmid, and Katharina Walliczek. It is concerned with the later life health effects of experiencing a famine in utero or shortly after birth. We use a unique combination of data on the German context of World War II and its aftermath – specifically on air raids on German cities during the war and food rations distributed during and after the war – to measure two reduced form effects: the effects of air raid attacks and the effects of famine early in life on

[1] I use the term 'reduced form' to describe analyses in which exogenous variables (possibly measured on a level higher than the individual level) that supposedly shift individual factors of interest (which may be endogeneous or unobserved) are used to proxy for the individual-level variation of interest. This approach precludes the estimation of an individual level causal effect or – in the case of my research questions – a parameter of a health production function, i.e. a parameter that describes the structural relationship between an individual level input and a health outcome.

various health outcomes. The results from this empirical exercise can be used to gain new insights on the nature of famine effects as they have been analysed in the medical literature for decades: comparing a set of individuals who experienced a famine within a critical age window to a control group that was not affected by this adverse condition at the same age provides a reduced form effect of early life famine and arguably an empirical test for theories of fetal programming triggered by nutritional shortages. However, as noted before, reduced form effects do not give an insight into what components of the aggregate adverse state are the drivers of the observed differences in later life health outcomes. A famine is defined as a state of mass starvation but psychological stress – either evoked directly by malnutrition but also by contextual factors paralleling the famine – may also be a relevant channel. Hence, the common interpretation of findings from famine studies were to be challenged if components different from individual malnutrition were quantitatively relevant determinants of later life health.

The analysis conducted in chapter 3 sheds light on this issue. Under the assumption that the experience of air raids evokes psychological stress but famine may evoke both stress and malnutrition on the individual level, we can show which outcomes are malleable to psychological stress and whether our famine effect contains a stress component of a relevant magnitude. The idea of our study is that if we find an outcome to be affected by air raids (i.e. stress) at a critical age window but not by famine, we have evidence for the stress component induced during exposure to the famine to be quantitatively small and a poor nutrition not to be a decisive factor for this outcome. On the other hand, if we find an outcome to be affected by the famine but not by air raids, this result hints at the fact that malnutrition may be the sole driving force. Our main findings imply that postnatal stress is a relevant determinant of later life hypertension among men. For male adult height, stress is not found to be relevant while malnutrition in utero is. Moreover, we find that female adult height shows no significant connection to stress but is reduced when a woman has been exposed to the famine in the first months after birth. Since adult height has been found to proxy overall health and also cognitive skills, we have an indication that the economic effects of malnutrition might be sizable. In general, the stress component inherent in famine exposure seems to be small in the context of our study.

In chapter 4, my co-authors Gerard van den Berg and Pia Pinger and I tackle another shortcoming of famine studies and start out by noting that a reduced form famine effect is in general not a quantitatively relevant measure of how a severe lack of nutrition in infancy or childhood affects later life health. More explicitly, during a famine, a fraction

of people may not be exposed to an individual level food shortage while for others hunger might be a problem also in non-famine times. That means that individual 'compliance' to famine, i.e. the shift from non-starvation to starvation induced by a famine, is unlikely to be perfect. Consequently, the effect of famine exposure on health outcomes, i.e. a reduced form famine effect, does not measure a causal effect of hunger or malnutrition since it is subject to attenuation bias.

For the analysis conducted in chapter 4 we thus combine survey data that contain retrospective information on early life circumstances and hunger suffered in childhood and aggregate level information on famines in three European countries in the 1940s. We use instrumental variable methods to estimate individual level causal effects of famine induced hunger in infancy and later childhood on adult height. We use adult height as the outcome since it is a compound measure of overall childhood health and the dietary history of an individual. Moreover, adult height has been shown to be significantly related to socio-economic outcomes, which implies that it should measure traits such as overall physical constitution and cognitive skills. For episodes of hunger very early in life that may be prone to recall bias, i.e. systematically underreported, we use two-sample instrumental variables methods to measure a local average treatment effect. More specifically, we estimate the marginal effect of a famine on individual level hunger within a subsample of respondents who were older when the respective famine took place and are thus assumed to actively recall the famine. We then use this estimate to rescale the average reduced form effect of famine in infancy on the outcome, as one would do it using a Wald estimator. The results imply that a commonly estimated reduced form effect of famine underestimates the individual level causal effect of early life hunger by at least a factor three. Specifically, we find that males who suffered from hunger between pregnancy up until age four are more than three centimeters shorter.

Chapter 2

Economic Conditions at Birth, Adverse Life Events, and the Decline of Mental Health and Cognitive Functioning Later in Life – Evidence from Eleven European Countries[1]

2.1 Introduction

Apart from the natural impact of progressing age, the mental and physical functioning of elderly individuals is affected by all sorts of adverse life events (e.g. Lin and Ensel (1989), Lindeboom et al. (2002), Fiske et al. (2003), Rosnick et al. (2007)). However, the impact of health shocks differs across individuals such that some may suffer more severe or long-lasting impairments than others. A natural starting point in detecting the source of such effect heterogeneity may be the current environment (e.g. the social network) and life style

[1]This chapter is joint work with Gerard J. van den Berg. An earlier version has previously been handed in as my Master Thesis at the University of Mannheim in the summer of 2011. Substantial changes have been made both concerning the empirical work and the verbal exposition where both more data have been exploited (SHARE survey wave 4) and the background on explanatory frameworks has been extended.

Data Acknowledgement: This chapter uses data from SHARE wave 4 release 1, as of November 30th 2012 or SHARE wave 1 and 2 release 2.5.0, as of May 24th 2011 or SHARELIFE release 1, as of November 24th 2010. The SHARE data collection has been primarily funded by the European Commission through the 5th Framework Programme (project QLK6-CT-2001-00360 in the thematic programme Quality of Life), through the 6th Framework Programme (projects SHARE-I3, RII-CT-2006-062193, COMPARE, CIT5- CT-2005-028857, and SHARELIFE, CIT4-CT-2006-028812) and through the 7th Framework Programme (SHARE-PREP, No 211909, SHARE-LEAP, No 227822 and SHARE M4, No 261982). Additional funding from the U.S. National Institute on Aging (U01 AG09740-13S2, P01 AG005842, P01 AG08291, P30 AG12815, R21 AG025169, Y1-AG-4553-01, IAG BSR06-11 and OGHA 04-064) and the German Ministry of Education and Research as well as from various national sources is gratefully acknowledged (see www.share-project.org for a full list of funding institutions).

of an individual. In contrast to possible immediate sources of heterogeneity, the present chapter is concerned with more remote factors, namely early childhood conditions.

A wealth of empirical studies from different fields of science confirm that there is a relationship between early life conditions and mental health later in life. Since ethical concerns obviously forbid experimental studies on human beings and individual level information on early life conditions are often missing, many studies proceed by exploiting "natural experiments" in the form of aggregate shocks that evoke adverse life conditions for a fraction of a population. One then compares individuals exposed to this aggregate level shock within a certain age window to whose who were not affected. Results suggest that fetal or postneonatal exposure to famines and thereby implied potential malnutrition can lead to higher risk of, e.g., showing signs of antisocial personality disorder in the early years of adulthood (Neugebauer et al. (1999)), Schizophrenia (Susser et al. (1996)), or symptoms of depression (Stein et al. (2009)); for a recent collection of findings from famine studies, see Lumey et al. (2011). Almond (2006) and Almond and Mazumder (2011) using the 1918 US Influenza Pandemic and Ramadan fasting as sources of exogenous variation, respectively, find that cognitive skills and mental health outcomes are significantly worse for individuals experiencing adverse conditions *in utero*. Case and Paxson (2009) show that the prevalence of infectious diseases both at birth but more so during the very first years of infancy impairs cognitive development. All these studies show how long-lasting the bequests of early childhood conditions are.

Empirical evidence on the importance of early life conditions on the trajectory of mental health in later life and the ability to cope with stressful events is rather scarce. Our study heavily draws from two recent empirical papers. Closest to this study is the the work by van den Berg et al. (2010) and Scholte et al. (2014) where the authors demonstrate that the cognitive abilities and physical functioning of individuals born under bad economic conditions are affected more severely by adverse life events. For the mental health outcomes studied in van den Berg et al. (2010), effect heterogeneity is most pronounced in the case of stroke. In a related framework, Doblhammer et al. (2013a) use survey data from various European countries and show that being born during a period of extraordinarily good or bad economic conditions has a significant impact on the level of various cognitive skill measures later in life. By using business cycle fluctuations at birth as a proxy for an exogenous shocks to economic well-being on the household level, both studies solve an endogeneity problem that evolves from unobserved components that drive both economic well-being and health. Business cycle shocks are independent of such components but nevertheless are expected

to have an impact on the single individual. Possible channels may be that recessions, for instance, evoke a insecure economic environment for some households, thereby potentially implying food insecurity or malnutrition and psychological stress. For children born during such a period or being *in utero* at that time, it is conceivable that such an environment may have long-run health effects.

This chapter combines elements of the previously mentioned studies by using pan-European survey data to analyse how later life shocks to health or general well-being impact on cognitive performance and how these effects are heterogeneous along early life conditions, measured as business cycle fluctuations. We thus complement the study of van den Berg et al. (2010) by using a related approach but different data; we also contribute to the literature by analysing how adverse life events affect not only cognitive skills but also clinical depression. Our analysis will thus show whether the results found in studies of only national perspective are context-independent, i.e. whether heterogeneity in the effects of adverse events later in life along the dimension of early life conditions is present in societies regardless of the overall level of economic prosperity, cultural or institutional differences etc. Using data from various countries is also appealing since it offers an additional source of cross-sectional variation. Hence, one needs not to rely on comparing only individuals from one single country at different stages of the business cycle which makes it hard to disentangle cohort effects from business cycle conditions.

The empirical framework we employ provides causal evidence under relatively weak assumptions.Establishing a causal link is of societal interest since first, mental health status is an influential determinant of individual economic decision making (e.g. Christelis et al. (2010), Banks et al. (2010), Banks and Oldfield (2007)). Second, in an ageing society the number of persons at risk to be impaired by adverse life events is expected to grow rapidly; since health care is associated with costs both for the single individual and her family and the society, there may be scope for mitigating the burden by already laying a foundation for cognitive robustness very early in life at considerably lower cost.

To briefly preview the results, we find that early life conditions have a significant relevance for the impact of adverse life events on the trajectory of cognitive performance among women only; the results show that adverse effects of stroke and to a lesser extent diabetes are cushioned when a woman is born in a relatively favorable early life environment. For men, we find no convincing evidence for effect heterogeneity. Moreover, adverse life events' impacts on changes in clinical depression are not found to be different

among individuals born under good economic conditions compared to those born in a less favorable environment.

The remainder of this chapter is structured as follows. The next section will give an overview about theoretical explanatory frameworks. Section 2.3 presents the data, variables, and sample selection criteria. Section 2.4 presents the empirical model and discusses potential identification issues. Section 2.5 discusses the results and the robustness of the findings. Section 2.6 concludes.

2.2 Explanatory frameworks

Numerous theoretical considerations point at the important role of early life conditions for adult mental health outcomes such as depression or cognitive abilities in general. For instance, Welberg and Seckl (2001) survey a strand of the neuroendocrinological literature that builds on the well-known "fetal programming" hypothesis of Barker (1992). They suggest that prenatal and perinatal exposure to stress, as it might be triggered by an economic downturn in the context of our study, may have a profound effect on parts of the central nervous system and thus may impair cognitive abilities or may have detrimental effects on individual's ability to cope with stress later in life.[2] They also cite a range of experimental animal studies that are in line with the proposed underlying channels. In general, the effect of early life stress on an individual's capabilities to cope with stressful events later in life is not unambiguous. It is conceivable that early life stress impacts on brain regions and eventually the physiological response of an organism and makes it more vulnerable to stress later in life (Pechtel and Pizzagalli (2011)). However, an adaptive response to early life stress could also imply that early life adversity programs an individual in a way that makes him or her more fit for later life adversity. Hence, early life adversity and stress could enhance an individual's stress coping skills and promote resilience.[3]

Another line of reasoning why early childhood conditions should affect the quantitative impacts of adverse events later in life on cognitive functioning can be centered around relatively simple neuropsychological arguments. (Further explanatory frameworks for the relationship of early life to mental deterioration can be found in van den Berg et al. (2010).) First, there is a vast interdisciplinary body of evidence that the *absolute level* of cognitive skills is determined to a great part very early in the life of an individual (for a synthesis

[2] See Almond and Currie (2011b) or Doblhammer et al. (2013a) for a collection of further possible biological mechanisms.

[3] Experimental animal studies have found early life stress to promote resilience; see, e.g., Lyons et al. (2010).

of findings from several strains of science, see Knudsen et al. (2006)). In this context, early means the first year after birth as well as prenatal. Since cognitive skills are determined by a learning process in which different skills and abilities beget each other in a hierarchical order (see again Knudsen et al. (2006) and references therein), a cognitive impairment very early in life (caused by stress, malnutrition, or other deprivation) can affect the overall accumulation of cognitive traits.

Second, it is often found that cognitive skills are highly correlated with physical indicators such as head circumference (e.g. Ivanovic et al. (2004)). High ability also manifests itself in better scholastic performance. Educational attainment but also physical indicators such as brain size and synapse count, are found to correlate with the ability of an individual to deal with diseases that cause cognitive impairment, adverse shocks, or the effects of ageing in general, i.e. to mitigate the *decline* in cognitive functioning. For a general overview, conceptual classifications, and evidence on this phenomenon (mostly on the impact of Alzheimer's disease but also with more general scope), see Stern (2002, 2009) and references therein.

According to Stern (2002, 2009), individuals differ in their level of "reserve" which can be regarded as a means of slowing down cognitive impairment due to ageing or adverse shocks such as mental diseases. He gives an overview of different concepts where so-called "passive models" (the brain's reserve capacity is run down until shocks yield clinical impairment) emphasize the role played by physical indicators (brain size) whereas "active models" (the brain actively tries to cope with adverse influences) stress the role of education and occupation. Both approaches to reserve suggest that not only cognitive skill accumulation but also the robustness of the same is decisively influenced by early childhood circumstances.

A last example of a potential pathway from early childhood to late-life cognitive decline could be established by findings from telomere biology. Telomeres are DNA-protein complexes which cap the ends of chromosomes and protect chromosomes from possible loss of essential information. They are often referred to as a biomarker for the biological age of a cell since the length of telomeres shortens with cell reproduction and by exposure to adverse environmental factors (such as smoking). Recent studies by Tyrka et al. (2010) and Kananen et al. (2010) found an association between childhood maltreatment and associated stressors and telomere length in adulthood. Adult stress has also been found to be associated with accelerated telomere attrition (e.g. Epel et al. (2004)). It has been hypothesised that telomere length could be an influential marker of cognitive function (Harris et al. (2006) state this

hypothesis but find a significant association between telomere length and verbal fluency only). Moreover, Martin-Ruiz et al. (2006) find that longer telomere length in patients who suffered a stroke has been associated with a lower risk of developing dementia and less adverse development in cognitive function. Even though most of the studies do not manage to establish causality due to a lack of longitudinal data, this evidence may be suggestive of a link between early life stress, thereby implied changes in cell metabolism, and robustness of brain function to neurodegenerative disorders, health shocks, or other adverse factors such as psychological stress. In addition, Cameron and Demerath (2002) cite evidence for a link from prenatal malnutrition in combination with subsequent catch-up to adverse health outcomes in adulthood and reduced longevity via accelerated telomere shortening. Because the present study considers inherently temporary shocks to economic well-being at times where the observational units have possibly been *in utero*, this may be another relevant connection from early life conditions to late life outcomes.

2.3 Data Description

2.3.1 Individual survey data

Sample selection. The empirical analysis uses data drawn from the Survey of Health, Aging, and Retirement in Europe (SHARE) which is a pan-European longitudinal survey of individuals aged 50 and over. The SHARE data is conceptually closely related to the US-American Health and Retirement Study (HRS) and the English Longitudinal Study of Ageing (ELSA) both in survey design and coverage of topics. Up to now, there are four waves of data available of which only the first two and the fourth are used. The third wave is not used since it has been concerned with reconstructing the entire life histories of the participants while the crucial health indicators employed as dependent variables in our analysis are not included. For general information on the survey design and other methodological issues, see Börsch-Supan and Jürges (2005) and Börsch-Supan et al. (2008).

We select individuals born between 1920 and 1945 with non-missing information in the first two waves of the survey from countries that took part in the first two waves. With exception of Greece which has not been part of wave four, respondents from all of these countries were re-interviewed also in waves three and four. Because we will estimate a model in first differences in the empirical analysis, we will often refer to changes between two subsequent waves of the survey. Apart from deleting observations with missing

relevant information, a sample selection criterion is that immigrants have been dropped. Moreover, as a consequence of sample selection, the working sample does not contain any observations that were surveyed by means of a proxy interview since the dependent variables are taken from modules which were not covered in this case (see Börsch-Supan and Jürges (2005, ch. 7)). The size of the samples used in the regressions is also reduced by the fact that some respondents may have answered questions used to construct the dependent variables, but were surveyed by means of proxy interviews in other modules concerning, e.g., demographics or children.[4]

Outcomes. The first dependent variable is an indicator that comprises symptoms of clinical depression on a scale from zero to twelve (the so-called Euro-D scale, see Prince et al. (1999)). One symptom accounts for one point on the scale. In the regression analysis, we transform this score by subtracting each respondent's value from 12 such that higher scores refer to a better mental condition.

Further, we use four measures of cognitive capacity.[5] The first is a numeracy score that ranges from one to five with five indicating the best possible performance. This score is based on the number of correct answers to a sequence of four simple arithmetic computations. This variable has the drawback that for wave four respondents, it has been mainly asked for new entrants into the survey. This means that when estimating a model in first differences, the data used almost exclusively stem from the first two waves.

The second measure is based on the performance in a word recall exercise. After a list of ten words was read out to the respondent, she was asked to name all the words she could recall right away. The number of right answers is stored in a variable ranging from zero to ten. After other tests have been conducted, the respondent was asked again which words from the list she could recall. This delayed recall capacity is stored in another index. We sum both recall scores and thus get a general indicator for recall capacity ranging from 0 to 20.

Another outcome variable indicates verbal fluency skills. Respondents were asked to name as many animals as they could within one minute.[6] The number of animals the respondent was able to name is stored in the verbal fluency score and is often referred to as

[4]There are two types of proxy interviews, full proxy and proxy in combination with the respondent. The latter type has been kept in the sample in order not to reduce the working samples too much. Full proxy interviews have been dropped in order to reduce errors in variables constructed from statements regarding diagnosis of diseases or death of family members.

[5]Christelis et al. (2010) use the same indicators for cognitive skills and provide further information on the generation of the variables.

[6]Compared to the other outcomes, there is a larger but still small fraction of respondents not participating in this test; this will show up in the regression results in the form less observations.

a measure of executive function. Note that all scores are such that higher values correspond to better cognitive skills. That is, when taking first differences between the values of the interviews, an improvement in cognition will turn up as a positive number.

Explanatory variables. For the first differenced explanatory variables, we use the following information. For diabetes, our indicators are based on questions from all survey waves where respondents were asked whether they ever had been diagnosed with this condition. For instance, if an individual reported a diagnosis in the second wave but not in the first, the indicator for the onset of this condition between waves one and two is set to one. For all other respondents (those who reported a diagnosis already in the first wave and those who never had been diagnosed with the condition), the variables are set to zero. For three incidences (namely heart attack, stroke, and cancer diagnosis), the respondents were asked in the second and fourth wave whether they had experienced them since the last interview. If a respondent said she did not know whether she experienced one of the incidences or if she refused to report information, she was treated as if there had been no such event. We notice that the SHARELIFE interview took place between the second and fourth wave and that it is not clear from the wave-four questionnaire to which last interview this question referred. Hence, we applied the same rationale as for diabetes to infer the advent of one of these shocks between waves two and four. Results were not sensitive to changing the definition.

For sources of psychological stress, we use information on whether the respondent lost a partner, a sibling, a child, or a grand-child in between the interviews and whether the household's financial situation deteriorated since the last interview. We will confine the analysis to binary measures only. In the second and fourth wave, participants were explicitly asked whether their marital status had changed since the last interview.[7] For the latter three measures there is no explicit information which necessitated the construction of a proxy: if the reported number of living brothers, sisters, children or grand-children of the respondent declined, this was counted as the death of a family member; since coding the absence of such an event with the reverse rationale resulted in many missing values, we decided to interpret missing information as the absence of an event. The number of family members may change due to factors unrelated to death. Given the fact that the sample consists of relatively old

[7] There are a lot of missings for the follow up question, what the marital status of the respondent is at wave II. Thus the "widowed" dummy is set to one when a respondent reported a change in marital status and declared that she was widowed in the second wave; if this information is missing, the dummy takes the value one if the respondent reported that the year of bereavement was larger or equal to the year of the first interview and the respondent is not recorded as being widowed in the first wave.

individuals, the number of underreported cases should be small with respect to the death of a direct offspring or a sibling. For the death of a grand-child, the probability of an error should be higher since it is more likely that an additional grandchild is born, i.e. it may be the case that a grandchild might have died but the number of grandchildren reported in the surveys does not decline. Notice, however, that the measures are nevertheless conservative in that an incidence in the data set should actually be valid but that it is possible that there are persons that experienced such an event but are categorized erroneously as "untreated".

We define an indicator for a household's financial situation to have deteriorated by using information on whether the household was able 'to make ends meet'. If the household's respondent to this module (typically only one for couples) answered 'with great difficulty' or ' with some difficulty' in a later wave but did not do so in the preceding wave, we defined this to be the onset of financial problems.

Figure 2.1 provides a means of inspecting the distributional properties of the outcome variables. Table 2.1 presents the descriptive statistics. It is noteworthy that sample sizes for the numeracy score and the financial situation at wave four are missing for a large part of the sample; analysing numeracy and including financial deterioration is thus expected to shrink the sample. We also note that most adverse shocks are rare events; the event that happens most often is the death of a sibling (about 16 percent of the sample experience the loss of a sibling). In general, health shocks occur only in about 3 percent of all cases but become a little more frequent between waves two and four which could either be due to the advanced average age at this point of time or the larger time interval between the interviews (recall that the SHARELIFE interview took part in between the two waves).

2.3.2 Macro-economic measures

As already noted above, there is an endogeneity problem emerging from the fact that good individual level economic conditions at birth as well as health may be jointly determined by unobservable factors. In order to overcome this issue, we follow van den Berg et al. (2006, 2010) in considering aggregate business cycle fluctuations which are not related to individual-specific unobserved components but nevertheless should exert an influence on the economic well-being of the individual. Using business cycle fluctuations as exogenous shocks is also appealing since they are relatively mild shocks compared to, e.g. famines or epidemics; hence, there are arguably no problems of selective mortality or selective fertility.[8]

[8]Evidence for business cycles not having an effect on the composition of birth cohorts with respect to socio-economic status can be found, e.g., in van den Berg et al. (2009b) or Kareholt (2001).

Table 2.1: Summary statistics for SHARE respondents

Variable	N	mean	s.d.	Min.	Max.
Background (for all respondents in a given wave)					
Age at wave I	14743	69.816	6.903	59.000	85.000
Age at wave II	13650	71.848	6.747	61.000	87.000
Age at wave IV	14408	75.784	6.603	66.000	91.000
Female	23930	0.535	0.499	0.000	1.000
Available in all waves	23930	0.289	0.453	0.000	1.000
Born in Boom Period	23930	0.533	0.499	0.000	1.000
Outcome levels at waves I, II and IV (for all respondents in a given wave)					
Numeracy Score at wave I (0-5)	14635	3.153	1.147	1.000	5.000
Immediate Recall Score at wave I (0-10)	14502	2.967	1.938	0.000	10.000
Delayed Recall Score at wave I (0-10)	14498	4.472	1.800	0.000	10.000
Fluency Score at wave I	14420	17.671	7.097	0.000	88.000
No. of depressive symptoms at wave I (0-12)	14406	2.350	2.252	0.000	12.000
Numeracy Score at wave II (0-5)	13537	3.208	1.165	1.000	5.000
Immediate Recall Score at wave II (0-10)	13421	3.180	1.966	0.000	10.000
Delayed Recall Score at wave II (0-10)	13417	4.616	1.804	0.000	10.000
Fluency Score at wave II	13331	17.802	7.290	0.000	100.000
No. of depressive symptoms at wave II (0-12)	13269	2.273	2.251	0.000	12.000
Numeracy Score at wave IV (0-5)	5406	3.069	1.171	1.000	5.000
Immediate Recall Score at wave IV (0-10)	13464	4.546	1.853	0.000	10.000
Delayed Recall Score at wave IV (0-10)	13456	3.162	2.119	0.000	10.000
Fluency Score at wave IV	13402	17.140	7.023	0.000	99.000
No. of depressive symptoms at wave IV (0-12)	13394	2.576	2.303	0.000	12.000
Shocks between waves					
Child or Grandchild died between wave I and II	9914	0.061	0.240	0.000	1.000
Sibling died between wave I and II	9914	0.154	0.361	0.000	1.000
Widowed between wave I and II	9769	0.031	0.173	0.000	1.000
Onset of financial problems between wave I and II	8503	0.091	0.288	0.000	1.000
Heart Attack between waves I and II	9761	0.021	0.145	0.000	1.000
Stroke between waves I and II	9761	0.021	0.143	0.000	1.000
Cancer Diagnosis between waves I and II	9499	0.022	0.148	0.000	1.000
Onset of Diabetes between wave I and II	9671	0.039	0.193	0.000	1.000
Child or Grandchild died between wave II and IV	8079	0.047	0.211	0.000	1.000
Sibling died between wave II and IV	8079	0.193	0.395	0.000	1.000
Widowed between wave II and IV	8384	0.046	0.210	0.000	1.000
Onset of financial problems between wave II and IV	5339	0.089	0.284	0.000	1.000
Heart Attack between waves II and IV	8366	0.038	0.191	0.000	1.000
Stroke between waves II and IV	7468	0.033	0.178	0.000	1.000
Cancer Diagnosis between waves II and IV	7441	0.037	0.188	0.000	1.000
Onset of Diabetes between wave II and IV	7480	0.056	0.230	0.000	1.000

Above cited studies solve the problem of the secular trend by taking into account only cyclical components of the business cycle, i.e. explicitly abstracting from the overall level of the trend. The cyclical component is thus used as a proxy for an exogenous income shock that may raise or lower the living standard of an individual relative to the idiosyncratic base level. The approach presented below parallels this rationale in that we back out the cyclical components for every country in the sample and thereby obtain cyclical shock components that can be used as proxies for exogenous shocks on the individual level independent of the national context.[9] Other issues emerging from country-specific or individual-specific components are then solved by specifying a robust empirical model as it is presented in the next section.

For the quantitative analysis, we will use suitable GDP data to measure aggregate economic conditions. The data are from Maddison (2010) and freely available on the internet. The GDP figures are in Geary-Khamis-Dollars of 1990, i.e. they are denoted in common currency and at a comparable price level. We essentially proceed as, e.g. van den Berg et al. (2006, 2010), i.e. we take the logarithmic GDP and detrend it by applying a Hodrick-Prescott filter with a smoothing parameter of 500 (see Hodrick and Prescott (1997)).[10] Doing so it is possible to obtain a detrended log-GDP series for each country in the sample by subtracting the secular trend from the log-GDP; this series is referred to as the cyclical component of the log-GDP.

We next define a variable that depicts whether a year correponds to a boom or not; the definition is based on a relative comparison in that we take all cyclical components of all years by country and define all the years where log-GDP is larger than the trend value to exhibit a trend. This corresponds to the definition in van den Berg et al. (2010) and makes the analysis conducted later on more tractable. Doblhammer et al. (2013a) use another definition by defining booms as years where the cyclical component is in the upper quartile of the country specific distribution while recessions refer to the lower quartile. While the latter definition should be more precise and relatively sharp in separating bad from good economic conditions, it is rather impractical when interest is in heterogeneous effects of adverse shocks.

[9] Here, a problem may arise if the the impact of an aggregate shock on the single individual is different for the countries. This might well be the case but is also a problem when analysing only one single country since over the course of time the institutional features are likely to change such that a, say, five percent drop in GDP at one point in time is likely to have a different impact on the single individual than it might have been the case 30 years earlier.

[10] The program used for filtering the time series is the one used in van den Berg et al. (2006) and can be found in the supplementary material to the article (see http://www.aeaweb.org/aer/data/mar06_data_20040207.zip).

Figure 2.2 provides a means of visual inspection of the aggregate data. Every subfigure plots the logarithmic GDP as well as the trend. Over the range from 1920 to 1945, most countries experienced both harsh recessions and large booms. The figures imply that the trend is indeed not continuously upward sloping: for most countries involved in World War II on the side of the Allied Forces, the trend-level declines from the end of the 1930s onwards. For most of the neutral countries in World War II, the trend keeps rising almost everywhere.

As noted in van den Berg et al. (2009a) it is advisable to select an observational window where countries start out in a boom and end in a recession since this makes cohorts born under good economic conditions older and thereby produces relatively conservative estimates with respect to the influence of age and the time-dependent improvement in overall conditions (living standards, hygiene, medical progress etc.): if a cohort born in a recession is affected more severely by life events, this should not be due to them being older on average or being born in relatively underdeveloped periods. Our observational window does not really satisfy this criterion while, however, the number of individuals in the sample who were born before 1925 or even 1930 should constitute a very small fraction. In order to address the sensitivity of the results to this issue, all major models haven been re-estimated using only cohorts born after 1925; figure 2.2 implies that most countries are in a boom at this point in time. All major results were not affected by this while, of course, sometimes statistical precision suffered due to somewhat smaller samples.

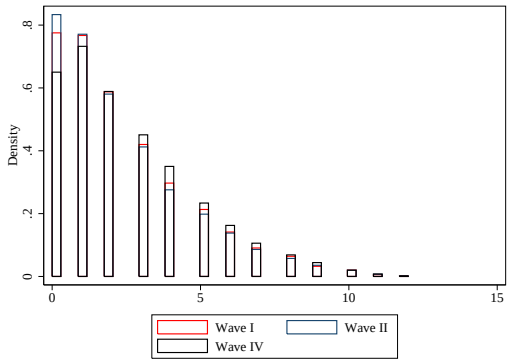
2.4 Empirical model and identification

The goal of this section is to set up an empirical framework that allows to analyse if and to what extent the economic conditions at birth have an influence on the quantitative impact of adverse life events on health outcomes. The statistical model we use in the analysis is essentially an adaption of the one used in Lindeboom et al. (2002) which starts from the primitive of a linear health production function:

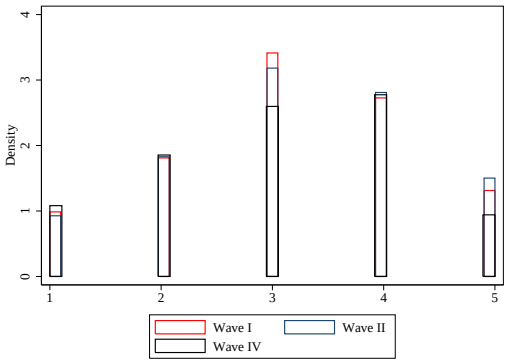
$$H_{it} = X'_{it}\beta_0 + Z'_{it}\beta_1 + \mu_i + \epsilon_c + \tau_{ct} + \gamma_t + u_{it}. \tag{2.1}$$

In this equation, the health status of a person i in country c at time period t is determined by individual background factors X_{it}, life events Z_{it}, a time-invariant and unobserved individual fixed effect μ_i, a time-invariant country fixed effect ϵ_c, a time-specific fixed effect γ_t, and time dependent country specific shocks τ_{ct}. β_0 and β_1 are vectors of coefficients, u_{it}

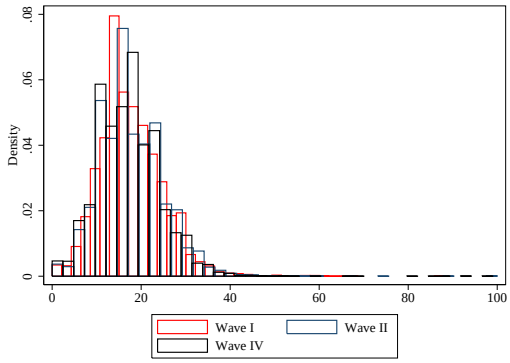
Figure 2.1: Distribution of outcome variables



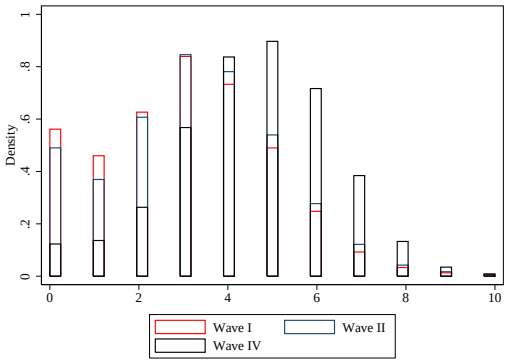
(a) Num. depression symptoms (0-12)



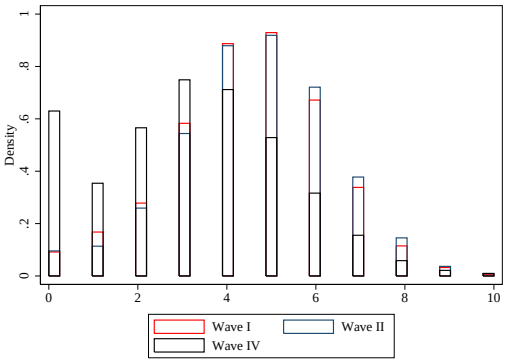
(b) Numeracy score (0-5)



(c) Fluency score



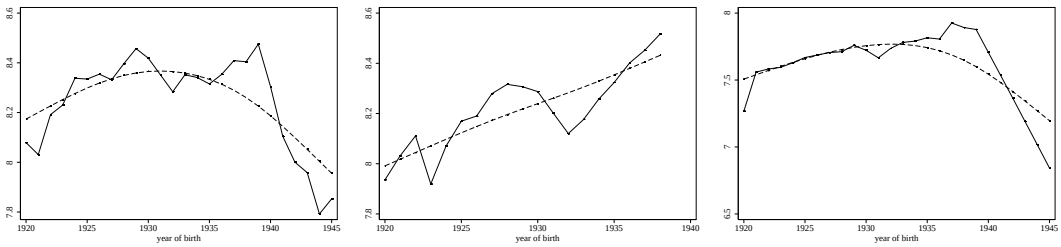
(d) Immediate recall score (0-10)



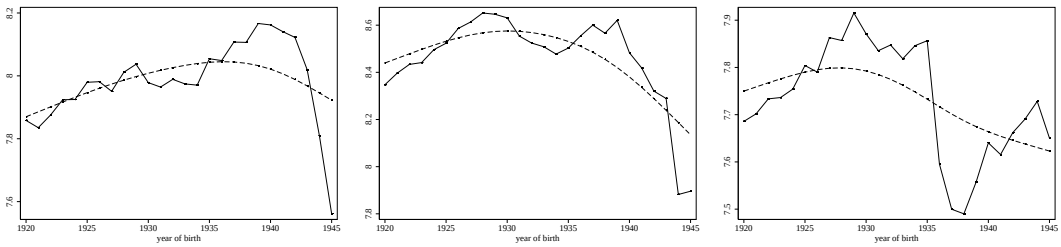
(e) Delayed recall score (0-10)

Figure 2.2: Logarithmic GDP and Hodrick-Prescott filtered trend by country (using a smoothing parameter of 500)

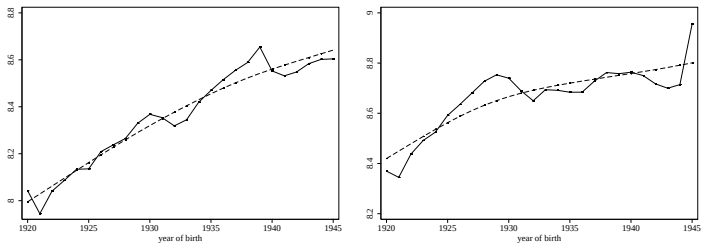
(a) Austria (b) Belgium (c) Denmark



(d) France (e) Germany (f) Greece



(g) Italy (h) Netherlands (i) Spain



(j) Sweden (k) Switzerland

denotes an idiosyncratic shock component which is assumed to be orthogonal to the set of regressors $\{X_{it}, Z_{it}\}$. Despite the fact that our outcomes are inherently discrete scores, we treat them as continuous throughout.

Denoting changes in variables between the two subsequent waves with Δ, (2.1) implies that the health change follows following process:

$$\Delta H_i = \gamma_1 + \Delta X'_{it}\beta_0 + \Delta Z'_{it}\beta_1 + \widetilde{\tau}_{ct} + \Delta u_{it} \tag{2.2}$$

where $\widetilde{\tau}_{ct}$ is a new country-time fixed effect that emerges after taking first differences.

Using this differencing method eliminates all time-invariant background information (education, overall childhood socio-economic status and so forth), individual-specific fixed effects and thereby all possible confounding factors that would emerge from time-invariant unobservable determinants of health in a cross-sectional analysis. Imposing additional structure on the process of dynamic selection which could, e.g., be due to the fact that several individuals are not observed in the two or more waves since they die, equation (2.2) can be consistently estimated by regressing the difference in health outcomes on the differences in explanatory variables and a constant, country-time fixed effects and country fixed effects. In order to assess whether shocks on the individual level have heterogeneous effects along early life conditions, we estimate a variation of the above stated relationship including interaction terms of all first difference regressors with the boom dummy, i.e. in (2.2) we let $\beta_0 = \overline{\beta}_0 + \alpha_0 \times Boom$ and $\beta_1 = \overline{\beta}_1 + \alpha_1 \times Boom$ (we also include interactions of the constant with these indicators while not allowing the country-specific trends to be heterogeneous accross economic conditions at birth).

More specifically, the problem of non-random attrition is solved by differencing if one assumes that any selection rule is non-randomly related only to the unobserved time-invariant individual fixed effect but not the unobserved transitory shocks. For convenience, let us repeat the formal argument of Lindeboom et al. (2002). Suppose there is a latent single index representation of the selection rule $S_i^* = Q'_{i0}\alpha + \nu_i$ that governs whether an individual is observed in all waves. This rule determines the observed attrition status $S_i = 1, 0$ where $S_i = 1$ indicates that individual i does not drop out from the sample. The factors in Q_{i0} are assumed to be exogenous and observed in the first wave; ν_i is a shock component. The

selection rule then takes the form

$$S_i = \begin{cases} 1 & \text{if } S_i^* \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

As noted by Lindeboom et al. (2002), it is common to assume that the idiosyncratic shock of the selection rule ν_i is systematically related to the individual fixed effect d_i from equation (2.1) while not being related to the shock u_{it}. Conditional on the event of an individual being observed in all waves (this is the event one implicitly conditions on when using a complete sample of wave one participants who do not drop out at some point), equation (2.1) becomes

$$\mathbb{E}[H_{it}|S_i = 1] = X'_{it}\beta_0 + Z'_{it}\beta_1 + \mathbb{E}[\mu_i|S_i = 1] + \epsilon_c + \tau_{ct} + \gamma_t. \tag{2.3}$$

Hence, one may allow health outcomes to be systematically different among drop-outs and those observed in all waves; however, this heterogeneity is time-invariant since all components in the term $\mathbb{E}[\mu_i|S_i = 1]$ are independent of time. This implies that the health change among the subset of individuals not dropping out of the sample is not different from equation (2.2). For different degrees of attrition as they are possible with more than two waves (being observed in two waves, not in all three), the argument can be extended analogously. A problem, however, emerges if ν_i and u_{it} are not independent; in this case it is in general not true that $\mathbb{E}[\mu_i + u_{it}|S_i = 1] = \mathbb{E}[\mu_i|S_i = 1]$ which would make this unobserved component time-varying. We solve this issue by following the advice of Lindeboom et al. (2002) who argue that this problem can be solved by including an additional dummy that takes the value one if a participant in the survey takes part in all waves.

Still, there are further caveats to identification which should be kept in mind for interpretation of the results. Notice that for stroke (and arguably also heart attack) one might encounter a endogeneity problem if the decline in outcomes is driven by temporary influences that also evoke a stroke. This could happen if a patient's blood vessels to the brain are blocked by some latent temporary factor that would be incorporated in the unobserved factor Δu_i. Notice also, that for some conditions such as diabetes a diagnosis may not coincide with the onset of the condition since it might have been latently affecting the functioning of the human body for longer already.

2.5 Results

2.5.1 General results

This section presents the results of the empirical approach outlined in the preceding section. By including interaction effects of the boom indicator with all adverse life events, the following results correspond to the estimates of van den Berg et al. (2010) where similar models are estimated on subsamples defined by good and bad economic conditions, respectively.[11]

Cognitive skills. Table 2.2 shows the results for the general working sample; tables 2.3 and 2.4 show the results when stratifying by gender. We first turn to the baseline effects of later life shocks. For cognitive outcomes, there is no indication for a faster deterioration in cognitive skills when born under adverse or good economic conditions as we do not find the interaction of the constant with the boom indicator to be systamtically related to any outcome. In our pooled analysis we find that the loss of a partner increases cognitive skills in the form of numeracy. The death of an offspring or a sibling are both not found to be statistically relevant for cognitive performance. Evidence on the effects of conjugal loss on cognition is scarce and rather mixed. For instance, Ward et al. (2007) show that bereavement decreases cognitive skills but that this statistical association mainly acts via symptoms of depression. In our case, stratifying by gender reveals that this finding is only present among men. It could be explained by mentally stimulating coping strategies, for instance more direct involvement into organizational matters of everyday living.

Other findings confirm that the onset of diseases impacts negatively on various dimensions of cognition among the elderly. Stroke, a heart attack, and also diabetes lower cognitive performance. The latter two mainly impact on the verbal fluency score while stroke has a statistically strong effect on the numeracy score. Stratification by gender shows that the cognitive performance of females is in general hit more heavily by those diseases. Cancer has a distinct impact on verbal fluency among men only.

[11] Scholte et al. (2014) augment their analysis by estimating the fixed effect in equation (2.1) after having consistently estimated the effects of later life events on the outcome using models in first differences. These estimates can in principle be regressed on early life conditions to measure the impact of those on the level of the outcome. We refrain from doing so for the following reasons. First, with an only restricted longitudinal dimension in our data, the fixed effects are hardly estimated consistently. Second and most importantly, by estimating the models in first differences, we can only estimate the impact of time-varying determinants of health. Hence the impact of factors like education or childhood socio-economic status will be included in the fixed effect estimate. Like the overall level of the health outcomes, those factors may be determined by early life conditions. This means that in a regression of the fixed effects on the measure of early life conditions, the results cannot be straightforwardly interpreted.

We now turn to our results regarding the impact of early life conditions on the quantitative impact on later life shocks. Strong evidence for effect heterogeneity along the dimension of early life conditions is rare. Most notably, the adverse effect of diabetes on recall and verbal fluency is cushioned when an individual is born under favorable economic conditions (heterogeneity is, however, only marginally significant). The effect of stroke on numeracy skills is also counteracted in case an individual was born in times of an economic boom.

In order to assess whether men and women suffer differently from adverse shocks and whether early life conditions interact differently with these shocks, we again turn to subsample analyses by gender. For men, this analysis suggests that cancer has a large and significant adverse effect on executive functioning as measured by verbal fluency. Bereavement is found to exert a positive effect on numeracy. Heterogeneity across our measure of early life conditions is not prominent among men.

Effect heterogeneity is more common among women. The results suggest that the adverse effects of diabetes on verbal fluency and recall found in the pooled analysis are driven by women; the same holds for the compensating effect of favorable birth conditions for this health shock (while here the coefficients lack high levels of significance). Stroke has a large and negative impact on all measures of cognition on average while, however, the coefficient is insignificant when verbal fluency is used as an outcome measure; this may be due to stroke being a rare event. For numeracy we find that the adverse effect of stroke is alleviated when a woman is born under good economic conditions. For all other health shocks, there is no or only imprecisely measured evidence for heterogeneity in effects along early life conditions.

We re-estimated all models including the cyclical component of the business cycle at birth explicitly in the models and replacing the boom indicator with this variable. This was done in order to see whether not dichotomizing this variable and not using the binary transformation added relevant information, possibly at cost of statistical precision. We found that doing so introduced noise into the estimations. Not surprisingly, the baseline effects often gained precision while the interaction terms lost significance. We concluded that the specifications detailed above may yield a less flexible distinction between different early life conditions but are statistically more practical in showing the relevant associations.

Depression. In comparison to cognitive skill measures, symptoms of clinical depression are more malleable to shocks at baseline. In the pooled analysis we find strong evidence for

adverse mental health effects of stroke, diabetes, financial problems, bereavement, and to a lesser extent the diagnosis of cancer. The effects of bereavement on symptoms of clinical depression are in line with the literature (see, e.g., Turvey et al. (1999), Cole and Dendukuri (2003), and the survey of Parkes (1992)). The effects of stroke and probably also other health conditions on depressive symptoms could be explained by resulting mental or physical limitations since disability in general has been found to be a decisive risk factor for later life depression (Cole and Dendukuri (2003)).

Again, we split the sample and performed analyses for men and women separately. The mental health of men suffers more from health shocks like stroke and cancer while for women we additionally find large effects for sources of psychological stress such as bereavement and financial deterioration. Hence, men seem to be able to cope with pure psychological stressors while their risk for depression is significantly affected by physical conditions and potentially resulting limitations. For women, psychological stressors such as bereavement and financial hardship turn out to be more influential.

Turning to the relevance of early childhood conditions for the magnitude of the impact of later life shocks on depression, we find that treatment effect heterogeneity is again not present both in the pooled and in the gender specific analyses. It is conceivable that a bad economic environment around birth triggers psychological stress on a family level that also impacts on a newborn or a fetus via a stress reaction of the mother. In our case we might not find an effect since the increase in psychological stress triggered by an business cycle downturn might be too small on an individual level to produce a measurable effect.

2.5.2 Regional variation

The analyses have not explicitly exploited the presence of observations from multiple countries in the data so far. In this section we will thus conduct country- and region-specific estimations in order to see whether the results differ between distinct regional environments. First, we use a sample of Dutch elderly. This will show whether there are differences in comparison to the results of van den Berg et al. (2010). Second, we will infer on possible regional variation by dividing the sample into a northern, a central, and a southern region. This division will show whether there is a "north-south-gradient" in effects which could be driven by economic, institutional or cultural differences. All of these factors are expected to exhibit a north-south-gradient.

Table 2.8 show results for models estimated using only individuals from the Netherlands. In principle, the sample sizes in the SHARE survey do not allow meaningful country-

Table 2.2: Estimation results for pooled sample

Variable	Depression	Recall	Numeracy	Fluency
Δ Age	0.032***	0.051***	0.016***	0.095***
	(0.008)	(0.010)	(0.006)	(0.021)
Δ Age sq.	-0.000***	-0.000***	-0.000**	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Born in Boom	-0.043	0.052	-0.024	0.182
	(0.044)	(0.063)	(0.029)	(0.123)
Child or Grandchild died	-0.016	0.157	0.045	-0.160
	(0.145)	(0.207)	(0.073)	(0.382)
Sibling Died	-0.087	0.057	-0.020	0.276
	(0.087)	(0.123)	(0.057)	(0.306)
Widowed	-0.597***	0.312	0.275***	0.024
	(0.182)	(0.237)	(0.098)	(0.426)
Cancer Diagnosis	-0.382*	0.157	-0.073	-0.937*
	(0.203)	(0.274)	(0.122)	(0.552)
Stroke	-0.770***	-0.679*	-0.430**	-1.082
	(0.258)	(0.355)	(0.178)	(0.720)
Heart Attack	-0.139	-0.543*	-0.031	-1.378**
	(0.205)	(0.325)	(0.168)	(0.544)
Onset of Diabetes	-0.540***	-0.412*	-0.056	-1.303***
	(0.161)	(0.224)	(0.103)	(0.439)
Financial Problems	-0.236**	0.034	-0.015	-0.135
	(0.120)	(0.168)	(0.079)	(0.285)
Child or Grandchild died × Boom	0.069	-0.059	-0.118	-0.125
	(0.196)	(0.278)	(0.101)	(0.549)
Sibling Died × Boom	0.177	0.028	0.068	-0.512
	(0.116)	(0.165)	(0.074)	(0.388)
Widowed × Boom	-0.339	-0.404	-0.214	-0.405
	(0.261)	(0.355)	(0.136)	(0.584)
Cancer Diagnosis × Boom	-0.308	-0.208	-0.023	-0.119
	(0.299)	(0.372)	(0.169)	(0.774)
Stroke × Boom	0.217	0.436	0.459**	0.256
	(0.346)	(0.488)	(0.226)	(0.893)
Heart Attack × Boom	-0.360	-0.014	-0.113	0.841
	(0.287)	(0.415)	(0.212)	(0.706)
Onset of Diabetes × Boom	0.444**	0.454	-0.001	1.327**
	(0.226)	(0.308)	(0.136)	(0.594)
Financial Problems × Boom	0.019	-0.075	-0.031	-0.284
	(0.159)	(0.226)	(0.098)	(0.378)
Constant	-0.410*	0.603*	-0.173	0.123
	(0.245)	(0.326)	(0.129)	(0.655)
R^2	0.032	0.018	0.007	0.023
N	10817	10939	6908	10895

Note: *,**,*** indicates significance at the 10%, 5%, and 1% level, respectively. Individual level clustered standard errors in parentheses. All regressions contain country fixed effects, country-year fixed effects, and an indicator for attrition status.

specific analyses. There is some indication that we are able to reproduce the findings in van den Berg et al. (2010) (even though they use a different sample and select different birth cohorts for their analysis) since we also find adverse effects for stroke on cognitive skills that are worse in case an individual is born under bad economic conditions; we, however, lack the statistical power to establish significance of the findings. We dig deeper into this issue and additionally construct a composite measure of cognitive skills by taking all scores described above, dividing by the maximum attainable score (for verbal fluency, we truncate the variable at 30 – which is the 95th percentile for this variable in both waves – and divide

Table 2.3: Estimation results for male sample

Variable	Depression	Recall	Numeracy	Fluency
Δ Age	0.043***	0.086***	0.013	0.121***
	(0.011)	(0.016)	(0.009)	(0.030)
Δ Age sq.	-0.000***	-0.000***	-0.000	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Born in Boom	-0.026	0.010	0.002	0.045
	(0.059)	(0.093)	(0.042)	(0.188)
Child or Grandchild died	0.220	0.375	0.078	0.622
	(0.171)	(0.294)	(0.109)	(0.536)
Sibling Died	-0.096	0.079	-0.114	-0.360
	(0.122)	(0.192)	(0.087)	(0.368)
Widowed	-0.405*	0.704*	0.456***	0.211
	(0.241)	(0.377)	(0.139)	(0.642)
Cancer Diagnosis	-0.551**	0.200	-0.124	-1.734**
	(0.272)	(0.387)	(0.151)	(0.780)
Stroke	-1.045***	-0.397	-0.222	-0.870
	(0.358)	(0.436)	(0.274)	(0.800)
Heart Attack	0.206	-0.714	-0.285	-0.993
	(0.281)	(0.461)	(0.226)	(0.833)
Onset of Diabetes	-0.395*	-0.204	-0.242	-0.675
	(0.212)	(0.303)	(0.154)	(0.649)
Financial Problems	0.027	0.089	0.052	0.204
	(0.173)	(0.241)	(0.119)	(0.437)
Child or Grandchild died × Boom	0.020	-0.059	-0.182	-1.261
	(0.248)	(0.394)	(0.154)	(0.771)
Sibling Died × Boom	0.112	0.124	0.095	0.102
	(0.164)	(0.258)	(0.113)	(0.489)
Widowed × Boom	0.103	-0.918	-0.228	-1.103
	(0.378)	(0.600)	(0.214)	(0.989)
Cancer Diagnosis × Boom	-0.202	-0.621	-0.046	1.326
	(0.384)	(0.510)	(0.215)	(1.003)
Stroke × Boom	0.226	0.182	0.225	-0.132
	(0.458)	(0.625)	(0.339)	(1.138)
Heart Attack × Boom	-0.403	-0.307	0.132	0.547
	(0.374)	(0.588)	(0.292)	(1.054)
Onset of Diabetes × Boom	0.300	0.221	0.023	1.283
	(0.297)	(0.426)	(0.197)	(0.852)
Financial Problems × Boom	-0.353	-0.228	0.013	-0.639
	(0.239)	(0.327)	(0.149)	(0.619)
Constant	-0.792**	-0.008	-0.207	-0.472
	(0.335)	(0.492)	(0.190)	(0.887)
R^2	0.034	0.032	0.017	0.030
N	5008	5069	3230	5041

Note: *,**,*** indicates significance at the 10%, 5%, and 1% level, respectively. Individual level clustered standard errors in parentheses. All regressions contain country fixed effects, country-year fixed effects, and an indicator for attrition status.

by 30), and then sum up all single scores. Our score consequently ranges from zero to four.[12] For this score, we find that stroke has a negative and highly significant effect at baseline and that this effect is significantly compensated by a good economic environment at birth (results not shown here). All other effects are zero also for this overall measure. The negative but insignificant effects of stroke on all measures of cognitive skills are thus strong enough to carry over to the composite measure which is at least suggestive evidence in favor of the results of van den Berg et al. (2010).

[12]Compared to the composite score used by Doblhammer et al. (2013a), we do not use quantiles of the scores at a given wave, i.e. individual performance is not measured relative to the overall sample's performance at a given point in time.

Table 2.4: Estimation results for female sample

Variable	Depression	Recall	Numeracy	Fluency
Δ Age	0.022**	0.022	0.019**	0.069**
	(0.011)	(0.014)	(0.008)	(0.030)
Δ Age sq.	-0.000**	-0.000***	-0.000**	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Born in Boom	-0.062	0.092	-0.045	0.296*
	(0.063)	(0.086)	(0.039)	(0.163)
Child or Grandchild died	-0.274	-0.068	0.019	-0.801
	(0.228)	(0.289)	(0.098)	(0.537)
Sibling Died	-0.092	0.053	0.056	0.825*
	(0.121)	(0.161)	(0.075)	(0.470)
Widowed	-0.740***	-0.035	0.090	-0.195
	(0.277)	(0.290)	(0.138)	(0.569)
Cancer Diagnosis	-0.216	0.103	-0.011	-0.071
	(0.299)	(0.390)	(0.198)	(0.747)
Stroke	-0.426	-1.119*	-0.661***	-1.513
	(0.358)	(0.597)	(0.226)	(1.262)
Heart Attack	-0.492*	-0.374	0.352	-1.648**
	(0.299)	(0.448)	(0.228)	(0.677)
Onset of Diabetes	-0.658***	-0.610*	0.041	-1.849***
	(0.235)	(0.324)	(0.130)	(0.596)
Financial Problems	-0.453***	0.029	-0.048	-0.346
	(0.164)	(0.232)	(0.106)	(0.376)
Child or Grandchild died × Boom	0.156	-0.019	-0.065	0.804
	(0.296)	(0.385)	(0.133)	(0.774)
Sibling Died × Boom	0.239	-0.055	0.061	-1.060*
	(0.160)	(0.214)	(0.097)	(0.587)
Widowed × Boom	-0.629*	0.077	-0.129	0.182
	(0.365)	(0.424)	(0.179)	(0.724)
Cancer Diagnosis × Boom	-0.367	0.303	0.008	-1.734
	(0.466)	(0.536)	(0.264)	(1.195)
Stroke × Boom	0.257	0.816	0.783***	0.878
	(0.522)	(0.778)	(0.252)	(1.405)
Heart Attack × Boom	-0.434	0.395	-0.478*	0.981
	(0.445)	(0.568)	(0.283)	(0.890)
Onset of Diabetes × Boom	0.523	0.633	0.064	1.313
	(0.336)	(0.439)	(0.182)	(0.836)
Financial Problems × Boom	0.309	-0.002	-0.078	-0.090
	(0.213)	(0.310)	(0.131)	(0.474)
Constant	-0.069	1.086**	-0.171	0.721
	(0.353)	(0.435)	(0.176)	(0.951)
R^2	0.041	0.017	0.012	0.025
N	5809	5870	3678	5854

Note: *,**,*** indicates significance at the 10%, 5%, and 1% level, respectively. Individual level clustered standard errors in parentheses. All regressions contain country fixed effects, country-year fixed effects, and an indicator for attrition status.

Tables 2.5, 2.6 and 2.7 show the results of estimations using three different subsamples comprising northern, central, and southern European countries, respectively. We follow Bolin et al. (2008) (who also use the SHARE data) in assigning Sweden and Denmark to the northern sample, Spain, Italy, and Greece to the southern sample, and the remaining countries (Austria, Germany, France, Switzerland, Belgium, and the Netherlands) to the central European sample. This classification is expected to divide the sample along different sources of heterogeneity. First, there are institutional differences. Second, overall economic prosperity both at the periods of birth as well as at the point when the health shocks occurred is expected to exhibit a north-south-gradient. Third, cultural differences should

differ between the regions; Bolin et al. (2008) and references therein, for instance, study the differences in family ties which are expected to be stronger in southern countries.

As the tables show, mean effects differ between the regions. The patterns we observe seem slightly ad hoc. For the subsample of observations from central European countries, the subset with the largest number of observations, results resemble those of the general analysis quite closely. In particular since the sample of central European countries groups together many heterogeneous countries which should introduce more noise in this sample than in the more specific ones, it is not obvious how cultural background can produce the strongest effects for the central part of Europe and different results for the extremely heterogeneous southern and northern parts. It thus seems that one should take these results with caution since any conclusions drawn regarding effect heterogeneity may in fact be driven by sample size rather than differences in regional features.

2.6 Conclusion

In this chapter we have performed an empirical analysis to uncover heterogeneity in effects of adverse life events and health shocks on cognitive skills and clinical depression across individuals born under different economic conditions. In order to overcome problems of dynamic selection and endogeneity issues, longitudinal data and an aggregate business cycle measure to proxy for economic conditions at birth have been employed. We use a multi-national dataset as it helps to disentangle cohort and business cycle effects. The chapter augments the analysis of van den Berg et al. (2010) who use a conceptually related framework to analyse a sample of elderly Dutch individuals.

Impacts of adverse life events on cognitive skills are only rarely found to be heterogeneous along the dimension of early life conditions. However, women seem to benefit from good early life conditions when exposed to an adverse shock while for men we do not find similar evidence; this heterogeneity along the dimension of gender has also been noted in van den Berg et al. (2010). Since most of the health shocks are only rarely observed in our data we remain skeptical as to what extent the lack of evidence for heterogeneous effects is due to sample size problems or the true absence of them in the underlying process we estimate; for instance, point estimates for the effect of stroke on different components of cognitive skills are often found to be very large but to lack statistical significance. Nevertheless it is remarkable that we find the adverse effect of stroke on cognitive skills among women to be present and significantly so only when a woman is born in a bad

economic environment. For diabetes we find similar but less pronounced results when analyzing a specific component of cognitive skills, namely recall. Recall or episodic memory is a cognitive skill that is probably most susceptible to general aging and thus may also be most malleable to adverse shocks. For depression, we find rather mixed results but no evidence for heterogeneity among men.

The fact that we find that relatively good early life conditions cushion adverse effects of stroke stresses the special importance of developmental origins for cerebrovascular accident which is well-documented in the medical literature. We do not find evidence supporting the relevance of early life conditions for the effect of cardiovascular diseases which has also been a focal point of the literature.

Our results lend additional support to the hypothesis that not only overall level of mental health is decisively influenced by early life conditions (as most of the literature has documented so far) but that also the robustness of mental outcomes, i.e. their vulnerability to major later life events, may have a component that is determined to some extent early in life. Since both accumulated mental health outcomes and their susceptibility to common adverse shocks eventually determine elderly individuals' economic decisions, our results should put additional emphasis on the importance of early life conditions also from policy makers' perspectives.

2.A Additional tables

Table 2.5: Estimation results for northern European sample

Variable	Depression	Recall	Numeracy	Fluency
Δ Age	0.039***	0.051**	0.030*	0.216***
	(0.013)	(0.022)	(0.017)	(0.044)
Δ Age sq.	-0.000***	-0.000***	-0.000	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Born in Boom	-0.072	0.272*	-0.024	0.199
	(0.085)	(0.148)	(0.073)	(0.302)
Child or Grandchild died	0.112	-0.126	0.038	0.237
	(0.248)	(0.399)	(0.147)	(0.642)
Sibling Died	-0.129	-0.353	-0.036	0.203
	(0.161)	(0.265)	(0.112)	(0.580)
Widowed	-0.247	0.303	-0.020	-1.344*
	(0.299)	(0.386)	(0.189)	(0.693)
Cancer Diagnosis	-0.290	0.247	-0.025	-1.556
	(0.389)	(0.511)	(0.199)	(0.966)
Stroke	-0.047	0.548	0.246	1.341
	(0.392)	(0.585)	(0.427)	(1.375)
Heart Attack	0.377	0.814	0.682	-0.361
	(0.428)	(0.732)	(0.438)	(1.051)
Onset of Diabetes	-0.186	-0.415	0.192	-0.367
	(0.309)	(0.443)	(0.194)	(0.727)
Financial Problems	-0.005	-0.064	0.196	-0.275
	(0.275)	(0.452)	(0.216)	(0.734)
Child or Grandchild died × Boom	0.102	-0.210	-0.282	0.898
	(0.351)	(0.587)	(0.233)	(1.288)
Sibling Died × Boom	0.008	0.184	0.170	0.043
	(0.222)	(0.363)	(0.165)	(0.749)
Widowed × Boom	-0.079	-0.232	0.205	2.750**
	(0.501)	(0.838)	(0.282)	(1.314)
Cancer Diagnosis × Boom	-0.026	-0.773	-0.087	-0.252
	(0.514)	(0.717)	(0.299)	(1.395)
Stroke × Boom	-1.068*	-0.771	-0.105	-2.491
	(0.547)	(0.967)	(0.489)	(1.958)
Heart Attack × Boom	-0.751	-1.596*	-1.090*	-0.136
	(0.579)	(0.938)	(0.557)	(1.533)
Onset of Diabetes × Boom	0.056	-0.256	-0.206	-1.591
	(0.397)	(0.719)	(0.354)	(1.231)
Financial Problems × Boom	0.126	-0.067	0.064	-0.100
	(0.368)	(0.651)	(0.294)	(1.041)
Constant	-0.611	0.838	-0.376	-1.898
	(0.386)	(0.632)	(0.403)	(1.232)
R^2	0.025	0.030	0.016	0.029
N	2185	2196	1283	2197

Note: *,**,*** indicates significance at the 10%, 5%, and 1% level, respectively. Individual level clustered standard errors in parentheses. All regressions contain country fixed effects, country-year fixed effects, and an indicator for attrition status.

Table 2.6: Estimation results for central European countries

Variable	Depression	Recall	Numeracy	Fluency
Δ Age	0.028***	0.085***	0.021**	0.073**
	(0.011)	(0.016)	(0.010)	(0.032)
Δ Age sq.	-0.000***	-0.000***	-0.000**	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Born in Boom	0.018	0.021	-0.048	0.194
	(0.061)	(0.096)	(0.045)	(0.186)
Child or Grandchild died	-0.056	0.084	0.018	-0.548
	(0.188)	(0.287)	(0.109)	(0.545)
Sibling Died	0.035	0.376**	-0.006	-0.133
	(0.122)	(0.187)	(0.099)	(0.366)
Widowed	-0.665***	0.487	0.446***	0.368
	(0.251)	(0.337)	(0.130)	(0.622)
Cancer Diagnosis	-0.463*	0.136	-0.199	-1.016
	(0.253)	(0.417)	(0.193)	(0.908)
Stroke	-0.787**	-0.988**	-0.710***	-2.501**
	(0.312)	(0.499)	(0.250)	(1.042)
Heart Attack	-0.114	-0.703	0.104	-1.604*
	(0.279)	(0.481)	(0.198)	(0.872)
Onset of Diabetes	-0.715***	-0.440	-0.025	-1.268*
	(0.222)	(0.354)	(0.153)	(0.676)
Financial Problems	-0.189	-0.149	0.023	0.144
	(0.153)	(0.239)	(0.106)	(0.405)
Child or Grandchild died × Boom	0.100	0.222	-0.061	-0.174
	(0.252)	(0.401)	(0.151)	(0.736)
Sibling Died × Boom	-0.116	-0.406	0.060	-0.202
	(0.160)	(0.259)	(0.125)	(0.532)
Widowed × Boom	0.125	-0.585	-0.261	-1.770**
	(0.338)	(0.500)	(0.183)	(0.793)
Cancer Diagnosis × Boom	-0.498	-0.080	0.074	0.013
	(0.387)	(0.533)	(0.246)	(1.191)
Stroke × Boom	0.509	0.937	0.601*	1.833
	(0.447)	(0.705)	(0.347)	(1.272)
Heart Attack × Boom	-0.232	-0.043	-0.197	1.182
	(0.399)	(0.616)	(0.273)	(1.089)
Onset of Diabetes × Boom	0.481	0.846*	-0.018	2.059**
	(0.306)	(0.469)	(0.188)	(0.932)
Financial Problems × Boom	-0.155	0.142	-0.130	-0.418
	(0.210)	(0.334)	(0.134)	(0.564)
Constant	0.004	-0.534	0.065	-1.433
	(0.332)	(0.529)	(0.193)	(1.027)
R^2	0.026	0.021	0.012	0.028
N	5148	5210	3166	5182

Note: *,**,*** indicates significance at the 10%, 5%, and 1% level, respectively. Individual level clustered standard errors in parentheses. All regressions contain country fixed effects, country-year fixed effects, and an indicator for attrition status.

Table 2.7: Estimation results for southern European countries

Variable	Depression	Recall	Numeracy	Fluency
Δ Age	0.004	0.023	0.007	0.113***
	(0.015)	(0.017)	(0.009)	(0.033)
Δ Age sq.	-0.000**	-0.000**	0.000	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Born in Boom	-0.130	0.045	0.004	0.265
	(0.085)	(0.102)	(0.041)	(0.202)
Child or Grandchild died	-0.077	0.632	0.021	0.052
	(0.397)	(0.439)	(0.127)	(0.912)
Sibling Died	-0.237	-0.068	-0.027	0.954
	(0.179)	(0.204)	(0.083)	(0.706)
Widowed	-1.107**	-0.135	0.226	1.155
	(0.484)	(0.582)	(0.214)	(0.939)
Cancer Diagnosis	-0.379	0.015	0.084	0.132
	(0.528)	(0.524)	(0.244)	(0.821)
Stroke	-1.435**	-0.759	-0.305	0.238
	(0.697)	(0.746)	(0.294)	(1.008)
Heart Attack	-0.468	-1.164**	-0.493	-1.599*
	(0.412)	(0.525)	(0.315)	(0.888)
Onset of Diabetes	-0.505	-0.444	-0.179	-2.066**
	(0.320)	(0.373)	(0.182)	(0.829)
Financial Problems	-0.399	0.378	-0.169	-0.624
	(0.247)	(0.262)	(0.135)	(0.496)
Child or Grandchild died × Boom	0.008	-0.637	-0.012	-0.628
	(0.496)	(0.510)	(0.162)	(1.146)
Sibling Died × Boom	0.669***	0.410	0.016	-1.342*
	(0.231)	(0.263)	(0.109)	(0.803)
Widowed × Boom	-0.857	-0.062	-0.567**	-0.358
	(0.627)	(0.707)	(0.275)	(1.207)
Cancer Diagnosis × Boom	-0.006	0.357	-0.056	-0.426
	(0.840)	(0.800)	(0.357)	(1.337)
Stroke × Boom	0.891	0.076	0.486	-1.360
	(0.875)	(0.908)	(0.372)	(1.275)
Heart Attack × Boom	-0.397	1.181*	0.428	0.829
	(0.586)	(0.664)	(0.361)	(1.106)
Onset of Diabetes × Boom	0.655	0.158	0.100	1.789*
	(0.450)	(0.492)	(0.246)	(0.973)
Financial Problems × Boom	0.273	-0.433	0.105	-0.067
	(0.306)	(0.339)	(0.161)	(0.591)
Constant	0.579	0.882*	-0.266	-1.685
	(0.449)	(0.506)	(0.194)	(1.158)
R^2	0.049	0.012	0.013	0.022
N	3484	3533	2459	3516

Note: *,**,*** indicates significance at the 10%, 5%, and 1% level, respectively. Individual level clustered standard errors in parentheses. All regressions contain country fixed effects, country-year fixed effects, and an indicator for attrition status.

Table 2.8: Estimation results for Dutch sample

Variable	Depression	Recall	Numeracy	Fluency
Δ Age	0.028	0.038	0.009	0.265***
	(0.031)	(0.058)	(0.026)	(0.089)
Δ Age sq.	-0.000	-0.000***	-0.000	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)
Born in Boom	0.117	-0.008	0.147	0.252
	(0.132)	(0.241)	(0.109)	(0.375)
Child or Grandchild died	0.290	0.399	-0.072	-2.151
	(0.524)	(0.749)	(0.287)	(1.896)
Sibling Died	0.159	0.408	0.178	0.255
	(0.212)	(0.370)	(0.166)	(0.596)
Widowed	-0.672	0.940	0.387	0.122
	(0.447)	(0.736)	(0.285)	(1.028)
Cancer Diagnosis	-0.429	-0.484	0.050	-1.608
	(0.530)	(0.886)	(0.481)	(1.093)
Stroke	-0.801	-1.081	-0.425	-4.477
	(0.666)	(1.288)	(0.456)	(3.118)
Heart Attack	0.622	0.157	0.181	0.634
	(0.568)	(0.935)	(0.344)	(1.540)
Onset of Diabetes	-0.178	-0.525	0.156	-1.381
	(0.373)	(0.731)	(0.260)	(1.072)
Financial Problems	-0.329	-0.235	0.005	0.176
	(0.411)	(0.641)	(0.246)	(0.948)
Child or Grandchild died × Boom	0.427	0.540	-0.288	1.170
	(0.665)	(1.120)	(0.365)	(2.142)
Sibling Died × Boom	0.024	-0.314	-0.093	-0.806
	(0.292)	(0.529)	(0.214)	(0.834)
Widowed × Boom	-0.134	-1.919	-0.228	-1.977
	(0.650)	(1.188)	(0.360)	(1.470)
Cancer Diagnosis × Boom	-0.818	1.191	-0.575	0.777
	(0.744)	(1.167)	(0.591)	(1.405)
Stroke × Boom	1.591*	1.806	0.309	3.989
	(0.944)	(1.885)	(0.891)	(3.600)
Heart Attack × Boom	-1.705**	-0.835	-0.449	1.159
	(0.843)	(1.231)	(0.514)	(2.128)
Onset of Diabetes × Boom	-0.776	1.414	-0.246	2.552
	(0.555)	(1.005)	(0.372)	(1.614)
Financial Problems × Boom	-0.225	-0.009	0.232	0.288
	(0.547)	(1.049)	(0.312)	(1.247)
Constant	-0.267	1.068	-0.100	-5.012*
	(0.939)	(1.710)	(0.608)	(2.641)
R^2	0.043	0.032	0.020	0.041
N	1028	1037	631	1032

Note: *,**,*** indicates significance at the 10%, 5%, and 1% level, respectively. Individual level clustered standard errors in parentheses. All regressions contain an indicator for attrition status.

Chapter 3

Is there a Stress Component in Famine Effects? Disentangling Long Term Health Effects of Adverse Early Life Conditions using German Data on Exposure to Famine and Air Raids[1]

3.1 Introduction

Recently, social sciences have witnessed an increasing interest in the long-run effects of early life conditions on adult health and socio-economic outcomes. For a concise overview on the impact of early life conditions on later life outcomes, we refer the reader to e.g. Almond and Currie (2011a). What has started as an epidemiological strand of literature – and has most famously been theorized about in the seminal work of Barker (1994) – has triggered interdisciplinary academic work mostly for its universal relevance: If early life conditions are found to have significant impacts on late life health outcomes they may induce long-run costs for the affected individuals and a public health system as a whole that may be underestimated when focusing on short term effects only.

Analysing the impact of early life conditions on later life health is a challenging empirical task. A standard approach is to use aggregate level shocks that affected a fraction of individuals during an early life period. The shock-like variation in early life circumstances

[1]This chapter is joint work with Gerard J. van den Berg, Anna Hammerschmid, and Katharina Walliczek.
Data Acknowledgement: Part of the data used in this publication were made available to us by the German Socio-Economic Panel Study (SOEP) at the German Institute for Economic Research (DIW), Berlin. We thank the DIW Berlin, especially Christine Kurka and Jan Goebel, for hospitality and help with the SOEP data.

can then (a) either be used as an instrument for an individual-level variable if available, or (b) it can be used to compute a reduced form effect of this shock on later life outcomes. The most frequently used approach is the second, mainly for its simplicity and its lower demand on data availability. Within this subset of studies one can find a bulk of work studying the long run effects of early life exposure to famine as a proxy for a nutritional shortage in early life; see Lumey et al. (2011) for a recent survey. Since nutritional deprivation and food insecurity are in principle economic problems, famine studies have also become of special interest in the economics literature (see, e.g., Lindeboom et al. (2010), van den Berg et al. (2011), Neelsen and Stratmann (2011), Jürges (2013)).

Outcomes found to be affected by famine range from chronic diseases and physical indicators such as type II diabetes risk and adult height to mental disorders such as schizophrenia as well as economic outcomes. Given the broad spectrum of outcomes affected by early life famine one question has yet to be conclusively answered: Is an individual-level nutritional shortage the only channel via which famine affects those outcomes? Or is there another component, such as famine induced psychological stress or stress evoked by factors paralleling the famine, that simultaneously influences long run health? If this was in fact the case, the validity of reduced form famine effects to empirically test theories of fetal programming through nutritional hardship had to be challenged and existing results had to be reinterpreted.

Two main arguments support this concern. First, most of the modern day famines that have been used to measure reduced form famine effects took place in the context of World War II or its aftermath. World War II and war activities in general, however, have been found to have a distinct impact on health and economic outcomes of individuals exposed to these conditions early in life. Second, like malnutrition, the hormonal reaction to stress early in life may have a profound impact on the development of metabolic features or gene expression (Maric et al. (2010)). Moreover, since early life stress and malnutrition are oftentimes found to be correlates of a low socio-economic environment which in general predicts worse later life health outcomes and metabolic features (Miller et al. (2009), Miller and Chen (2010), Morozink et al. (2010)) it is evident that disentangling the effect of stress and malnutrition is a hard task not only on an aggregate but also on an individual level.[2] Exogenous variation that plausibly shifts only one of these components, i.e. evokes either stress or malnutrition on an individual level, may be a remedy to this problem.

[2]Krabbendam et al. (2005) show that demographic factors pick up part of the effect of self-perceived stress on adverse pregnancy outcomes which again emphasizes the connection between the two variables.

Our study addresses this concern and uses World War II (WWII) and its aftermath in Germany as a unique natural experiment to disentangle the effects of early life famine and early life stress. We combine a detailed micro data set that includes the place of birth of a respondent on a level as fine as the exact municipality with city-level information on the exact timing and intensity of Allied air raids on Germany during WWII and regional-level information on food rations that were distributed during the German famine in the aftermath of the war. Under the assumption that air raids that were experienced very early in life or in utero evoke stress and that famine may have an effect via both stress and malnutrition we can use our empirical findings to trace out the distinct effect of each component on a range of outcomes. For instance, if we find one outcome to be significantly affected by stress induced by bombings but not significantly so by exposure to a period of famine, our finding hints at the stress component of famine to be negligibly small. Conversely, if famine exposure has an effect on an outcome while the experience of air raids does not, it follows that a nutritional shortage is likely to be the relevant biological channel for this outcome.

Up to now, conclusive evidence on this matter is scarce even though the problem we outlined above has been acknowledged in the literature on, for instance, the Dutch famine where psychological stressors may have been prevalent also due to war-related factors paralleling the famine (e.g. Roseboom et al. (2001)). Even if a specific stressor may have been identified, it requires a large degree of cross-sectional and temporal identifying variation in this stress inducing factor that is independent of famine exposure to come up with a solid empirical result. Our unique data set provides this information, with the raw data of air raids being measured on a very narrow local level and with daily precision.

Our findings are important for several reasons. First, knowing which outcomes are influenced by a certain type of early life condition may help to determine potential threats to late life health more narrowly and select adequate compensatory measures. Second, we provide reliable empirical evidence on the relevance of severe stress early in life on later life health and economic outcomes. Third, our findings aim to help researchers to interpret existing results of famine studies and to design studies in the future.

To preview our results, we find that male adult height is adversely influenced solely by prenatal exposure to famine whereas female adult height is influenced by postnatal exposure to famine. Exposure to air raids in the first six months after birth has an adverse impact on hypertension risk among men. The economic importance of our findings for reduced adult height among men when prenatally exposed to famine is strengthened

by evidence for famine to be relevant also for pension income, a measure of life time productivity. In the light of our theoretical exposition, those reduced form estimates imply that male and female adult height is solely influenced by malnutrition whereas hypertension among men is likely to be determined by stress only. Moreover, in the context of our study the stress component associated with famine exposure is unlikely to be quantitatively relevant.

The results are robust to various sensitivity analyses. Most importantly, we show that our main results remain robust when controlling (i) for the potential endogeneity of the place of birth of a respondent; (ii) for selective mortality during the war; and (iii) for exposure to famine and air raids at later ages. The endogeneity of the place of birth is of importance since children and mothers were evacuated from cities on a large scale by state programs during the war. Using additional data on moves of mothers during pregnancy or the first six months of life, we show that we can at least partially control for the degree of selection by including a variable that measures the overall exposure of a particular city to air raids until just before the estimated month of conception. Selective mortality can be addressed since almost half of the casualties directly related to air raids occurred at particular dates in particular cities, namely four so-called 'firestorms'. Dropping observations born in those cities does not alter the results. The last issue is of particular importance since individuals have been found to be vulnerable also at later stages of their childhood (e.g. van den Berg et al. (forthcoming), Sparén et al. (2004)). Moreover, observations affected by air raids early in life are those exposed to a famine later in childhood. Explicitly controlling for exposure to adverse conditions later in childhood does not alter the results.

The remainder of this chapter is structured as follows. Section 3.2 reviews the theoretical and empirical literature on the effects of stress and malnutrition. Section 3.3 describes the historical context of our study. Section 3.4 presents our data sources and discusses the key variables. Section 3.5 describes our analytical framework, the empirical strategy, and discusses potential threats to identification. Section 3.6 presents the empirical findings and various robustness checks. The final section discusses the results and concludes.

3.2 Explanatory frameworks and previous evidence

Infants and fetuses are generally regarded as particularly susceptible to outside influences. The fetal programming hypothesis by David Barker links those outside influences to permanent alterations in the human metabolism and long-run health risks, such as a high

risk for type II diabetes and cardiovascular diseases (see, e.g., Barker (1994)).[3] We are interested in the long-run consequences of prenatal and postnatal stress and nutritional deprivation on long-run health. In this section we present an overview of the literature on the effects of early life stress (3.2.1) and malnutrition (3.2.2).

3.2.1 The effects of early life stress

Conditions of stress are known to lead to the release of the hormone cortisol and other glucocorticoids. Cortisol enhances brain functioning and boosts energy metabolism. Under chronic stress, however, high levels of cortisol are known, among others, to increase the risk of hypertension and obesity. For expectant mothers, high cortisol levels have been found to reduce birth weight which hints at the adverse effect of prolonged stress on the intrauterine environment (Bolten et al. (2011)). Maric et al. (2010) cite two possible physiological relationships that could explain this finding. First, cortisol and other maternal glucocorticoids could enter the placental environment in times of prolonged stress and thereby alter the catabolism of the fetus. Second, stress situations may change blood circulation and reduce the blood supply to the uterus, which might be an evolutionary mechanism to protect the mother's life at the expense of the fetus. Both of these routes may lead to a 'predictive adaptive response' of the fetus' metabolism to prolonged maternal stress (i.e. an adaption to an anticipated stressful or dangerous environment, see e.g. Gluckman et al. (2005b)). For newborns, a similar adaptive reaction to outside stressors and high cortisol levels (now no longer via the mother's higher release of cortisol and thus no longer filtered by the placental barrier) may be possible. Hence, prenatal and postnatal exposure to stress may have distinct effects on late life health.

Apart from an increased susceptibility to developing chronic diseases, prolonged early life stress can have consequences for the functioning and development of the human brain and thereby affect cognitive outcomes. Lupien et al. (2009) give a comprehensive overview of animal and human studies and also provide theoretical explanations (see also Talge et al. (2007)). Since brain regions are rapidly developing in utero and in early childhood (up until puberty), they are especially malleable and responsive to adverse outside influences. Lupien et al. (2009) cite evidence for stress-triggered increased hypothalamus-pituitary-adrenal (HPA) axis activity, i.e. enhanced release of glucocorticoids, which ultimately may lead to alterations in the development of brain regions responsible for controlling the HPA axis.

[3] See also Almond and Currie (2011b) for a comprehensive survey of fetal programming studies with economic scope.

The hippocampus is one such brain region that is rapidly developing early on and may thus be especially vulnerable during the fetal and the post-neonatal period of life. Consequently, increased and prolonged early life stress may trigger cognitive deficits (especially regarding memory) and emotional disorders (see the studies cited in Lupien et al. (2009)). Evidence pointing into this direction can also be found in Aizer et al. (2012) who find negative effects of elevated maternal cortisol levels during pregnancy on children's cognitive and educational outcomes. Since they control for mother fixed effects, their results are arguably causal. In contrast, Black et al. (2014) find only short-term negative effects of stress caused by the death of a grandparent during pregnancy in utero on birth weight and neo-natal child health. In the medium to long-run, outcomes such as school choice, IQ score or earnings are not found to be affected.

Medical studies have also analyzed the consequences of maternal psychological stress and of exposure to stressful situations for newborns in the short-, medium-, and long-run. Khashan et al. (2008) show that schizophrenia risk is increased in offspring born to mothers who suffered from severe psychological stress, such as the death of a close relative. Also considering schizophrenia, van Os and Selten (1998) and Malaspina et al. (2008) show that this condition is more prevalent in individuals who were in utero during the German invasion of the Netherlands in May 1940 or during the 1967 Arab-Israeli war, respectively. In a similar vein, other large-scale man-made disruptions, such as events related to war or the 9-11 terrorist attacks but also natural disasters, have been used as arguably exogenous variation in expectant mothers' stress levels. For instance, Maric et al. (2010) show that pregnant women living in Belgrade during the 1999 bombings gave birth to significantly lighter babies while no effects are found on the probability for giving preterm birth or other complications during pregnancy. These findings are corroborated by the extensive literature review in Harville et al. (2010) who find that fetal growth may be reduced by stressful events but that length of gestation is typically not.

3.2.2 The effects of early life malnutrition and famine

Long-run health effects of malnutrition around birth are a major focus in medical as well as in economic studies. Barker (1990) finds that intrauterine environment and conditions very early in life explain regional differences in cardiovascular mortality much later in life, where maternal nutrition, birth weight and infant mortality rates approximate early-life conditions. Other studies find a relationship between in utero exposure to a famine and adult height,

diabetes, and schizophrenia. For an overview of the research on early-life conditions see e.g. Lumey et al. (2011), Almond and Currie (2011b) or Pollitt et al. (2005).

A potential mechanism behind these associations may be epigenetic changes during gestation which may permanently alter the phenotype: A fetus is able to adapt to a variety of environments due to a high developmental plasticity. Especially during so-called critical periods (mostly periods of rapid organ development), the human body adaptively responds to biological or social cues (Kuzawa and Quinn (2009)). During gestation, the mother's metabolism transfers information about environmental conditions, e.g. nutrient availability, to her baby through the placenta. If the mother suffers from malnutrition, the baby reacts by adapting to a world with food scarcity. That may involve modifications in hormone production, metabolism, development of organs, and fetal growth rate. By these modifications, the fetus gains survival advantages under the adverse situation which, however, come at a cost: If nutrient supply increases after birth, the famine-programmed individual may be more likely to suffer from cardiovascular disease, hypertension, type II diabetes, higher mortality, or obesity (see, e.g., Cameron and Demerath (2002); Barker et al. (1990); Gluckman et al. (2008); Schulz (2010)).

Besides biological pathways, also behavioral mechanisms may explain long-run effects of early life malnutrition. Parents might try to compensate early-life food scarcity in a later stage of childhood, when the nutritional situation improves. As stated above, however, nutrient-rich food might be harmful for the famine-programmed child. A hunger period can also induce a higher stress level in the family through a fight for resources within and between households. For non-biological pathways from early life famine to later life health effects see, e.g., Hamelin et al. (1999), Hadley and Patil (2006) and Whitaker et al. (2006). In the context of the German famine, Kesternich et al. (2013) present evidence for behavioral channels to be relevant for the connection between early life hunger and later life adult health.

Several famines have been studied to separate the effects of malnutrition from socio-economic status, which is in general strongly correlated with experiencing hunger; see Lumey et al. (2011) for a comprehensive survey of the famine literature. The "Dutch Hungerwinter" in 1944/1945, caused by an embargo initiated by the German occupying forces, often serves as an example of adverse living conditions, see e. g. Banning (1946), Lumey et al. (2011), and Painter et al. (2005). Studies found cardiovascular diseases, lower self-rated health, obesity and diabetes to be among the consequences of malnutrition. Susser

and Stein (1994) also find a decrease in adult stature after exposure to the Dutch famine shortly after birth.[4]

3.2.3 The interconnection between stress and malnutrition

After having discussed the effects of stress and malnutrition separately, we note that there may also be a connection between both conditions. Stress and malnutrition may be interconnected and may also influence health outcomes to some degree via a common channel.

The biology literature suggests that adversity in early life alters the phenotype and thereby influences health later in life: low socioeconomic status in childhood has been associated with a pro-inflammatory phenotype that might foster late life diseases in the long run (see, e.g., Miller et al. (2009); Miller and Chen (2007)). Moreover, there is indication for harsh family climates to also trigger pro-inflammatory phenotypes (see, e.g., Miller and Chen (2010)). Along these lines, it is conceivable that stress induced by the exposure to food shortages or bombings early in life may lead to similar epigenetic mechanisms. Thus, stress due to nutritional shortages and bombing exposure could influence later life health, to a certain degree, through such a common channel. Additionally, in a recent review Kuzawa (2012) points towards some similar health consequences that are potentially triggered by nutritional and psychological stress in utero.

Moreover, there is indication for malnutrition to have an inherent stress component. For instance, fasting has been shown to increase cortisol concentrations in healthy men and women (see, e.g, Bergendahl et al. (1996) and Fichter et al. (1986), respectively).

[4] Also the "Great famine" in China 1959-61 has been analyzed with respect to long-run effects. Chen and Zhou (2007) use regional excess death rates as a proxy for famine intensity and find adult height to be lower by 3.03 cm among those severely exposed to famine. Susser and St Clair (2013) compare studies on Dutch and Chinese famine with respect to mental disorders. Increased risk of schizophrenia seems to be a common result of both famines. Li et al. (2010) find a higher risk of metabolic syndrome among those individuals exposed to the famine during fetal life or early childhood, which is aggravated by Western dietary habits or overweight during adulthood. According to Royer and Witman (2013), sex ratio, literacy, employment and marriage rate were also affected in children exposed to the famine around birth.

3.3 Historical background[5]

3.3.1 Allied air raids on Germany

British air raids on German targets were deliberately kept minor in the first years of World War II since the British did not want to provoke retaliatory actions by the German forces which they were not able to counter with the insufficient material they had at hand in the early phase of the war.[6] Moreover, before Germany occupied the neutral states Belgium and the Netherlands, many of the more important industrial targets, mostly located in the Ruhr region, were only accessible via either dangerous or cumbersome routes if the neutral air space along the North Sea was to be respected. After the Dutch surrendered to Germany in May 1940, the British started their offensive against Germany, focusing almost entirely on the bombing of oil infrastructure and communication targets. While such kind of strategic bombing of industry had appeared to be a means of limiting front line warfare and the oftentimes barbaric mass killing of troops as it had occurred in World War I, it soon evolved to exhibit another little known dimension, namely the direct involvement of civilians.

The precision bombing campaigns turned out to be relatively ineffective and thus gradually, starting with an official order in October 1940, a list of some 20 to 30 cities and their civilian areas explicitly became targets of British air raids which were to be attacked in case the major industrial targets could not be located due to a lack of visibility. The first city bombing took place only about two months later, when the town center of Mannheim was the target of an attack carried out on 16 December 1940. In the following months and until the end of the war, such attacks, aimed at undermining the morale of the German population, were not only gradually extended but eventually became the major part of the campaign against Germany. In early 1942, area bombing, i.e. efforts not to hit specific targets but larger geographical areas and thus not discriminating between industry and civilian areas, became the British Bomber Command's top priority. It is important to note that the selection of targets was not exclusively based on economic importance of a city (see, e.g., Friedrich (2002) and Vonyó (2012)).

The area bombing offensive was paralleled by an expansion in Bomber Command staff and air craft as well as technical innovation, mostly with respect to navigation devices,

[5] We note that World War II in general, German aggression and the horrendous actions taken by the Nazi regime affected many people in different nations. Our focus on the effects of World War II on the German civilian population is due to the focus of our study on the German context and is not be interpreted as a normative stance.

[6] This section mainly draws from Davis (2006).

which increased the precision of bombing operations. The strategic paradigm shift quite naturally also translated into higher absolute bomb loads dropped on targets and a higher load per sortie. It is symptomatic that during the period prior to the area bombing campaign, from September 1939 until December 1941, only about 50,000 tons of bombs were dropped, representing only about 5 percent of the bomb load dropped over the course of World War II (Davis (2006)). In June 1943, the British Bomber Command joined forces with the US Eighth Air Force starting with the 'Pointblank' directive. While the British bombers continued area bombing operations, the US forces almost exclusively focused on strategic precision bombings, even though the US officials' renunciative attitude towards the bombing of civilian targets and the US Air Force's actual restraint from such operations may have been overstated due to image concerns during the war (Schaffer (1980)).

The area bombing efforts of the Allied airborne forces cannot be regarded as unambiguously successful. Kirby and Capey (1997) cite official documents painting a rather negative picture in that both the intended adverse effects on German production and the population's morale were minor and always rather short-term. For instance, German armament production was estimated to have exhibited a positive trend during the period of the Pointblank offensive (mid 1943 to March 1944) and German production peaked mid 1944 when it was about 2.3 times the level of what it had been in early 1942 (Werrell (1986)). Regarding the time dimension in the severity of bombing activity it should be noted that about 70 percent of the bombs on Germany were dropped after the Pointblank offensive (Werrell, 1986).

3.3.2 Civilian life in Germany during World War II

For the purposes of the present study it is vital to shed some light on the living conditions of the civilians during the air raids on Germany. In general, bombings impacted on many facets of German civilian lives. Our exposition is mainly focused on three major sources of selection, namely selective survival, internal migration, and health effects of bombings for the survivors.

With regards to the first, it has been established that up to 600,000 people lost their lives in air raids (Werrell (1986)). According to Hewitt (1983), half of those casualties occurred in the four firestorms over Hamburg, Kassel, Darmstadt, and Dresden, the first two taking place at the end of 1943, the latter in late 1944 and early 1945, respectively. Those figures have to be compared to the overall pre-war population which is estimated to be around 80 million. Since targeted cities and areas therein were oftentimes (but not exclusively) larger

in scale and more industrialized, it is conceivable that a large proportion of the casualties had been of a working-class background. In fact, Hewitt (1983) documents that the British officials explicitly declared working-class neighborhoods as the number one civilian targets. This could lead to a positively selected sample of survivors in our analysis. Therefore, we control for socio-economic background of the parents in our analyses. Regarding excess child mortality, The United States Strategic Bombing Survey (1947, ch. 6) states that there was no evidence for the observed increase to be attributable to bombing attacks. Still, we test robustness by excluding people potentially affected by the four firestorms mentioned above.

With respect to the second selection issue it must be noted that large scale evacuation programs started in 1943 and a total of about 10 million people were evacuated from cities; see Heineman (1996). The main focus of these programs was to protect women and children, whereas employed women without children or with older children had to remain in the cities. With regards to the evacuation of children, the German government organized large-scale programs which, however, were not unanimously accepted by parents. The number of children evacuated peaked in 1941 and remained rather low thereafter, despite increased bombing activity of the Allied forces and evacuation programs gradually becoming non-voluntary from 1943 onwards. This may imply that parents deliberately opted in and out of evacuation programs, preserving the status quo regarding location choices. Very young children were accompanied by their mothers and it is generally reported that resentment was smaller for such special sub-programs. In total, more than two million children were evacuated during the war, about half a million below the age of six, i.e. jointly with their mothers (Kock (1997)). We take care of the issue of selective migration in a robustness check.

The third issue deals with health effects of bombings for the survivors. Judging from historical accounts, the experience of air raids triggered psychological stress. Contemporary Allied observers find increased heart disease mortality rates as well as increases in stress-related diseases as peptic ulcer, cardio-vascular diseases, neurodermatitis, glaucoma, urticaria and excema in cities shortly after air attacks (The United States Strategic Bombing Survey (1947, p. 99-149)). Also subjective well-being seems to have suffered already in the first years of the war, since consumption of sleeping pills and pain killers increased substantially as well as medication-induced chronic gastric damage (Süß (2003, p. 398)). This affected also young children, since mothers suffered psychosomatically especially from the double burden of waged work in arms manufacturing or agriculture and housework and childcare (Süß (2003, pp. 403)).

Apart from stress, however, Allied air attacks do not seem to have specifically affected other health risks for children – in spite of bombings drastically affecting the civilian life: 800,000 civilians were injured in area bombings, more than 2 million housing units were destroyed, making 7.5 million people homeless (Hewitt (1983)) so that many had to live in crowded emergency accommodations (Süß (2003, p. 391)). Scarcity of housing stock persisted for years after the war and was aggravated by the bulk of refugees from the former eastern parts of the German Reich. Additionally, a major part of medical care infrastructure was destroyed by the air raids (Süß (2003, p. 391)). Since British Bomber Command raids took almost exclusively place in the night, people furthermore had to spend nights in crowded bombing shelters. Yet, all those factors did not lead to major epidemics of infectious diseases (see The United States Strategic Bombing Survey (1947, ch. 4) and Süß (2003, pp. 391)). This does not mean that there was no increase in disease rates at all. Still, in most contemporaneous reports it is explicitly noted that infectious diseases were not directly associated with air raids (The United States Strategic Bombing Survey (1947, ch. 4)).[7]

For our empirical strategy it is most important to note that health quality of toddlers and children did not suffer from adverse conditions directly associated with air raids other than through psychological stress. Namely, the adverse conditions of the war affected age groups very differently (Süß (2003, p. 399)). Whereas health status of adults decreased much earlier during the war, most health offices reported good to satisfying health quality for young children as late as 1943 and even 1944. Special health care facilities and e.g. synthetic vitamins to counteract deficiency symptoms may have been responsible for their surprisingly good health (Süß (2003, p. 400)). Correspondingly, infant mortality rose very little until 1943. Only in 1944, it rose substantially to around 10% in big cities (Süß (2003, p. 387, 447)). As noted above, this late increase cannot be attributed to air attacks alone, but rather to the cumulative adverse conditions (The United States Strategic Bombing Survey (1947, p. 126)).

3.3.3 Nutrition in Germany around World War II

In this section, we briefly summarize the food situation in Germany around the Second World War. The exposition is partly based on a summary of the German post-war famine

[7]For instance, diphtheria, a disease mostly prevalent in children, was endemic already before the war and only exhibits a slight increase in the infection rate during the war. Apparently, most of this surge in diphtheria cases can be attributed to excess adult cases (Galazka and Dittmann (2000)). After declining between 1932 and 1938, also tuberculosis incidence rose again from 1940 onwards (Süß (2003, pp. 382)). Here, similarly, mostly young men were affected, and the rise in incidence rates was most pronounced in rural regions, which were less affected by air raids than cities.

in van den Berg et al. (2011). For more detailed descriptions, see also e.g. Klatt (1950), Farquharson (1985) or Trittel (1990).

In August 1939, the German government installed a food rationing system which determined the maximum amount of food each person was eligible for. Consumers were differentiated by intensity of physical work, e.g. coal miners earned extra supplements, but also children got a specific composition of food (extra amount of milk). Before food rationing started, nutritional quality in Germany was high. Authorities were able to hold it constant during the first war years (food rations during the war amounted to about 2,500 kcal per day and per capita, Klatt (1950)). In 1942 the food supply got slightly more difficult. Rations were reduced due to bad harvest and restricted imports in April 1942, but the reduction was almost fully called off again in October. In 1943 and 1944, the bombing of German cities and evacuation programs became the major challenges. Although some rations had to be reduced or could be handed to the citizens only some weeks late, the nutritional situation altogether was still not critical. In 1944 nutrition worsened due to the loss of agricultural areas to the Allied Forces and infrastructure problems. In the beginning of 1945, food transports via train were largely impossible, the Allied forces increasingly occupied agricultural areas and nutrition became a major problem for the civilian society (see Schmitz (1956, pp. 357)).

By the time it surrendered to the Allied Forces, Germany had come to the end of its economic capabilities. Cities and infrastructure had been destroyed by Allied bombardments, many workers had died or were – as prisoners of war – unavailable to the labor market. Additionally, much of the Eastern part of the Reich was lost to the Soviet forces which implied great losses in agricultural capacities and a huge influx of refugees into the newly established occupation zones. Besides transportation means, also qualified labor and fertilizer were missing for a sufficient food supply and distribution. The centrally organized food administration system was set inactive directly after the war. Food supply in the weeks directly after the end of war was badly organized, until the occupational powers had installed their own systems. The Allied forces took care of food supply in order to avoid epidemics and riots. Their explicit aim was to hold the German living standard below the average of all European countries excluding Great Britain and Soviet Union and including South and Eastern European countries (Schmitz (1956, p. 358)). In the first time after the war, local authorities had to run the ration system independently of any central organization, and imports from the Allies were often insufficient.

Hunger and starvation became a part of German day-to-day living as, at some stages of the famine, people had to live off rations around 1,500 kcal per day and person or even less (Cornides (1948)). Due to adverse weather conditions causing bad harvests, deficits in transportation infrastructure and a low food production worldwide the situation only improved from spring 1948 onwards (Schmitz (1956, p. 365-368)). Especially during the harsh winter of 1946/47, the so-called 'hungerwinter', the food situation deteriorated to catastrophic levels. Since also the crop of 1947 largely failed to meet the needs of the increasingly starving population, rations dropped even below 1,000 kcal in many parts of Germany.

Rations, however, differed across zones, since each occupational power had its own regime and its own food administration. Although the US and the UK occupation zone united quite quickly in early 1947, the French and the Soviet zone had completely independent systems until the end of the famine. But also within zones, food availability varied depending e.g. on the share of agriculture in the region, which led to black market activities.

Although the rationing system was still, at least partially, in place until April 1950 in the Western zones and even longer in the Soviet zone, nutritional quality improved a lot around the currency reform in the Federal Republic of Germany (FRG) in June 1948. This can be seen in the rations data but also in contemporaneous literature for all four zones. The currency reform is one factor that was responsible for this improvement (making it more beneficial to sell agricultural products on official instead of black markets or keeping it in stock). Other reasons given in the literature are a surprisingly good harvest in 1948 worldwide and the Marshall Plan (starting in early 1948) providing high quality food imports (Rothenberger (1980, pp. 210)). From July 1948 onwards, the official calorie level was at an average of 1,800 kcal. However, individual spending capacities increasingly determined real caloric intakes (Schmitz (1956, p. 370)), which eventually enabled most of the population to achieve a normal nutritional condition relatively quickly.

3.3.4 Effects of World War II and the post-war famine

The short-run effects of warfare on the German civilian population as surveyed by different contemporary observers have been reviewed in Section 3.3.2. Other empirical studies consider the long-run effects of World War II or similar stressful events. Individuals who were affected by World War II are found to exhibit worse economic and health outcomes. Akbulut-Yuksel (2009) shows that city-level war time destruction of housing stock has negative effects on educational attainment, adult height, and health satisfaction

of individuals who were of school-age during the war.[8] While impacts on educational attainment may point to a rather mechanical effect of war destruction, effects on height may hint at biological channels. Hence, the reduced form effects measured in that paper cannot be straightforwardly interpreted since effects of war time destruction on health can be triggered by direct experience of heavy air raids or more detrimental living conditions after the war.

In their study of the effects of World War II on health outcomes and socio-economic measures, Kesternich et al. (2014) advance on this issue by using retrospective individual level reports on war time suffering and thus providing evidence on the channels via which war exposure might affect long-run outcomes.[9] In general both studies show that World War II may have evoked a whole variety of adverse factors, ranging from psychological stress to nutritional deprivation, which may all have long-run health consequences and can hardly be disentangled.

In psychology, there is also a strand of literature analyzing the effects of World War II trauma. For instance Glaesmer (2013) provides a recent overview on long-run psychological effects in a German sample. The author also points to a recent interest in the effects of trauma on physical health. As an example, Glaesmer et al. (2011) analyze the relationship between post-traumatic stress disorder (PTSD), traumatic events and physical health measures in a sample of elderly Germans. They find a significant relationship between health and PTSD/trauma (war-related as well as civilian). Their findings suggest a significant correlation between traumatic experiences and all 20 medical conditions under study (including hypertension, diabetes), whereas 12 of 20 outcomes are significantly related to PTSD.

The post-war famine in Germany 1945-1948 has been analyzed already in other studies (see above). Contemporary public health reports document that nutritional deprivation had a profound immediate and midterm impact on the German population: children as well as newborns were smaller and lighter than normal and suffered from a higher risk of infections and infant mortality (Klatt (1950); Dean (1951); Droese and Rominger (1950)), adults were underweight (Report of the Military Governor No. 49 (1949)) and showed symptoms of psychological stress (Davis (1951)). Studies that measure long-run effects on those born or conceived during the famine suggest that those cohorts are worse off in terms of higher

[8] See also Ichino and Winter-Ebmer (2004) for additional evidence on the effects of World War II on educational attainment.

[9] In the context of the effects of famine on long-run health, van den Berg et al. (2011) stress the non-causality of reduced form effects of famine exposure and use individual level reports of nutritional deprivation in childhood in order to measure a causal local average treatment effect.

education and occupational status (Jürges (2013)). Additionally, regional variation in early life adversity during the famine is found to be associated with differences in height among men (Akbulut-Yuksel et al. (2013)).

3.4 Data

3.4.1 Survey data

Our analysis is based on the German Socioeconomic Panel (SOEP) and the respondents of its most recent wave (2012).[10] We restrict our analysis to individuals born between 1930 and 1950 within today's borders of Germany since we have full contextual information only for birth places in Germany. An important advantage of our data set is that it allows us to identify a respondent's place of birth on a level as detailed as the municipality of birth. Valid information on the place of birth is another sample selection criterion.[11] This enables us to merge in contextual information about exposure to air raids and to famine around birth. Since the place of birth entered the questionnaire only in 2012, the respondents of the 2012 wave constitute the basis of our analyses.

In addition, the data provide information on the month and year of birth. Since our analysis will be concerned with the very first months of life of a respondent and the prenatal period, background information referring to any period thereafter may be endogenous with respect to early life circumstances. We thus use only a limited set of individual level control variables, i.e. the father's level of education, since any information determined after birth may be an outcome of the early life experiences of interest and thus endogenous.

Our outcome variables refer to different domains of health. They are all measured late in life, namely, when the respondents were on average 71 years old. First we have physical indicators that have been found to be sensitive to nutritional shocks (diabetes and height) or to stress (hypertension). In the SOEP, data concerning health conditions like diabetes or hypertension are collected by asking whether a doctor ever diagnosed a respondent with the respective disease. Second, as an economic outcome variable we use monthly pensions (in Euros) for men, as they reflect life-time income and thus productivity.Since the pension system was different in Eastern Germany until re-unification, we do not use respondents born in Eastern Germany when analysing this variable. Also, we drop respondents born 1949 and 1950 when analysing this variable as they are comparably young and thus a

[10]Concerning data quality and attrition issues of the SOEP, see, e.g., Kroh (2013), Lipps (2009).

[11]For some respondents the place of birth was ambiguous due to multiple cities having the same name. In this case, we declared them to be born in the largest of the candidate cities.

Table 3.1: Descriptive Statistics by gender

	Men			Women		
Variable	N	Mean	s.d.	N	Mean	s.d.
Bombing in pregnancy	1994	0.064	0.244	2046	0.074	0.262
Bombing in postnatal period	1994	0.074	0.261	2046	0.084	0.277
Tons p.c. dropped during gestation	1994	0.000	0.001	2046	0.000	0.004
Tons p.c. dropped in 1st 6 months	1994	0.000	0.007	2046	0.000	0.002
Tons p.c. dropped during gestation if positive	119	0.003	0.005	149	0.005	0.014
Tons p.c. dropped in 1st 6 months if positive	144	0.007	0.024	166	0.004	0.007
Less than 1500 kcal during gestation (1/0)	1988	0.158	0.365	2043	0.164	0.371
Less than 1500 kcal 1st 6 months (1/0)	1987	0.147	0.355	2043	0.147	0.355
Year of Birth	1994	1941.546	5.699	2046	1941.680	5.630
Place of birth: pop. >= 10000	1992	0.574	0.495	2041	0.554	0.497
Father's education high	1994	0.196	0.397	2046	0.172	0.377
Hypertension/High blood pressure	1989	0.460	0.499	2045	0.428	0.495
Diabetes	1989	0.188	0.390	2045	0.137	0.344
Height	1993	175.903	6.489	2046	163.427	5.843
Pension, Euro (if retired, from West, born bef. '49)	1117	2030.765	1336.183	–	–	–

relatively large fraction has not yet retired. Table 3.1 gives an overview on background variables and the outcomes.

3.4.2 Contextual variation: air raids on Germany

We supplement the micro-data set with data on air raids on Germany. The data accompany the detailed description on Allied air raids by Davis (2006). The data provide information on day, month and year of all US and RAF attacks on Germany as well as the name of the city the target was located in, the type of target, and the tons of bombs dropped. For early attacks of the Royal Air Force on German facilities, the data are only crude in the sense that they do not provide information on the exact target or the exact date. Thus, we use this data set starting with the attacks from June 1941 only. Unless noted differently, we take a bombing operation as the observational unit throughout this section, which we define as an air raid on a specific city on a specific day.

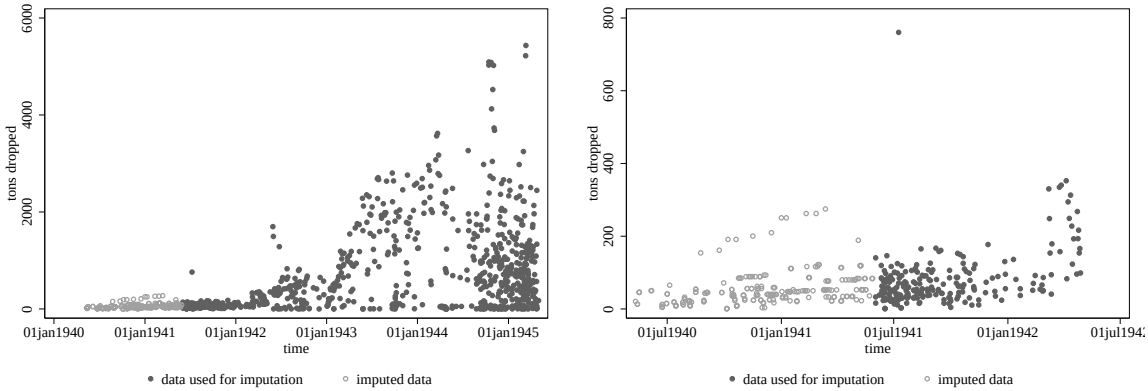
The most important task is to identify the targeted cities and merge in the regional identifier (*Gemeindekennziffer*, GKZ) information in order to identify survey respondents as being born in an air raid stricken city. This task is complicated by wrong spelling of city names in the bombings data, multiple cities having the same name, reforms that merged different localities into one, and also the fact that some localities do not have an own GKZ since they are too small to be independent municipalities. For unmatched targets that have been bombed more than once and with a sufficiently large bomb load, we went through the data set and took care of them receiving the correct GKZ. Then we applied probabilistic matching algorithms exploiting the 'reclink' package in STATA to match the

remaining targets which, however, yielded sufficiently good matches in only a small number of cases. In total, we start with data on 7261 bombing operations in Germany, whereof we ignore a fraction since they were mining operations or the data does not specify their target region. Eventually, we succeed in matching 6075 bombings, which leaves us with 423 unmatched bombings on German municipalities. Unmatched bombings are much smaller than matched bombings in terms of tons of bombs dropped, both on average as well as along the distribution. Targets therein were never hit by bombings exceeding 200 tons or their name is ambiguous or they lie in non-German areas. Furthermore, more than half of them were targets of opportunity. Due to factors detailed below, missing information on the remaining bombings is not likely to bias our results.

Since the RAF attacked German cities already before June 1941, we collected information on British Bomber Command attacks for the period prior to this date. The Bomber Command war diary information of Middlebrook (1985) provides such information alongside detailed verbal descriptions of operations for the whole war period. For the period from May 1940 to May 1941, this gives additional information on 251 bombings with sufficiently detailed information on the targeted city. As a qualitative measure of how strong the attacks have been, this data source only provides information on the number of sorties per attack, dichotomized in four categories. This information is inferior to our primary source's information on bombs dropped since the number of sorties has often not been equal to the number of attacking air craft and the strength of an attack has been primarily dependent on the type of air craft and the possible bombing capacity. Moreover, the bomb load dropped by a given air craft has been strongly time dependent since later in the war, with operations focused on destroying vast areas and not specific industry targets or facilities as in the early phase, tonnage was increased.

In order to overcome this missing data problem, we impute bomb loads for the early period exploiting the overlap between the two data sources, i.e. common entries in both the Bomber Command War Diary data and the data set by Davis (2006) from June 1941 to the end of the war. For those bombings that turn up in both data sets, we have both the tons of bombs dropped on a target and the four category score, giving information on the attacks strength. We regress the logarithm of tons dropped on a set of dummies depicting the number of sorties, quarter of year dummies, a linear time trend and a second order polynomial in the number of targets attacked on the same day, and the 1939 population size of a given city. Using the coefficient estimates, we predict conditional expectations for the

Figure 3.1: Fit of imputation: bomb loads dropped per air raid before June 1941



(a) All data

(b) Data around the cut off

tons of bombs dropped in the early period of the war, i.e. prior to June 1941. Figure 3.1 gives a graphical overview of the imputation fit.

We sum the tons of bombs dropped on a city in a given month which will provide us with our main contextual variation. We scale this variable by the 1939 population size of a city as measured in the 1939 census (see Zeitler and Mewes (1941)). This information is available only for cities above 10,000 inhabitants.[12] Hence, in a small number of cases we declare individuals who lived in a bombed city in one of our critical periods as not exposed to bombings because they were born in a small municipality with a number of residents below 10,000. We notice that this renders our results as conservative. Last, in order to merge the information on the bombings and our micro-data in a meaningful way, we aggregate the bombings to a monthly level.

Figure 3.2(a) gives an overview over the geographical variation in our data. For this graph we aggregate our data to a region level, which is a coarser level than we have in our data; since for many municipalities in one region we do not have population size, we do not scale the overall bomb load by population size. Figure 3.2(b) abstracts from the overall magnitude of bombs dropped and depicts the number of attacks per region (recall that an attack is defined as an air raid on a distinct city on a given day). Both maps show that the air raids exhibited some geographical dispersion but that they were mostly concentrated in some major areas around Berlin and at the Ruhr and that southern parts of Germany were

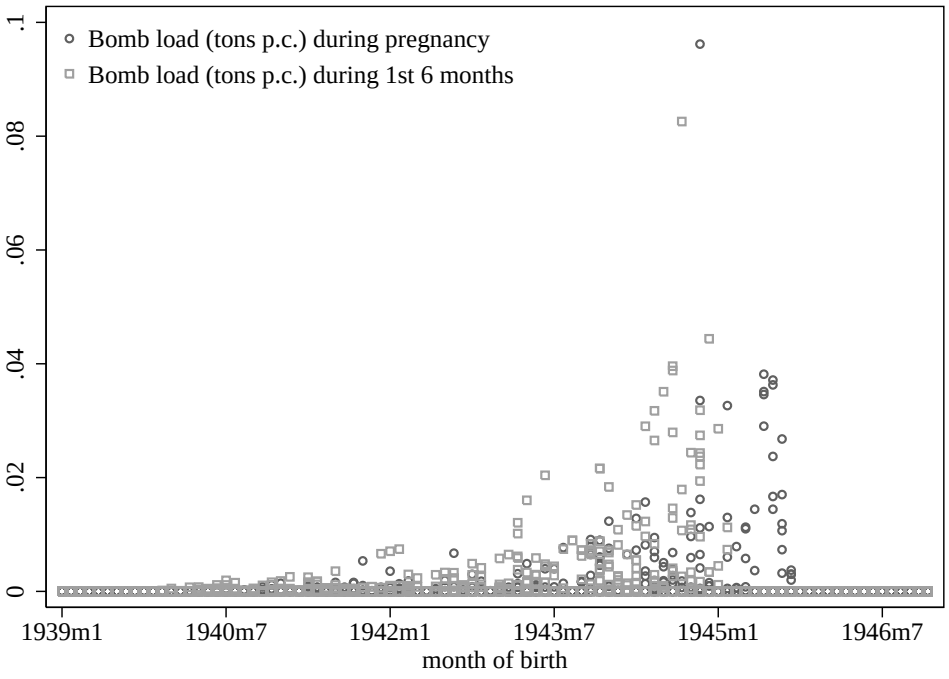
[12] We digitized this information manually and oftentimes merged the location identifier to this file by hand, too. Due to multiple places having the same name and no way of discriminating one against the other in a few cases, we expect small measurement error here (some identifiers having the wrong population size). For municipalities that were merged into one, we added up the population size of all entities subsumed therein for which we have information on the number of inhabitants.

(a) Overall tons of bombs dropped by region

(b) Number of attacks (distinct city on a given day) by region

Figure 3.2: Regional variation in air raids

Figure 3.3: Exposure of SOEP respondents to air raids in tons per capita (without 1 large outlier > 0.1)



less affected by bombings. Figure 3.3 shows the variation in the data when merged into the sample of SOEP respondents. The figure shows that the majority of the sample has not been exposed to bombings over the course of the war. Those who have been exposed to the war are found to have faced increasingly severe circumstances in the last year of the war.

3.4.3 Contextual variation: regional-level data on food rations

Judging from historical accounts, the food situation can be declared as non-critical until the last months before the Nazi regime collapsed and after the currency reform in June 1948. We define a 'normal' food situation to exhibit a baseline caloric level of 2,500 kcal. To account for the deterioration in the food situation in the last months of the war, we set the ration levels to 2,000 kcal for January 1945 to May 1945. In order to depict the nutritional situation in post-war Germany during the famine, we use information on official food rations for the so-called normal consumer from different historical sources. This specification has the

advantage that it is most widely available and that it is comparable over space and time.[13] Food rationing cards were distributed every four weeks ("Zuteilungsperiode", ZP) until December 1947, when the system changed to calendar months.

Using our raw data we have to tackle three challenges: First, area-level time series are almost never complete. Second, the 'normal consumer' ration is oftentimes not a good measure of the de facto food availability since shortages made it impossible to meet the ex ante promised rations, but also black market purchases or supplements are unobserved. Third, historical accounts show that there was variation in food availability also within zones, namely between urban and rural areas (see the section on the historical background).

Augmenting our basic data on official normal consumer food rations with different types of data, we address all of these challenges using an imputation model that exploits not only (a) official area-level ration data for the normal consumer, but also (b) official area-level data on actual call-ups, i.e. the rations that were de facto distributed, and (c) city-level rations from different sources which we also assume to be call-up data rather than an official ration.[14] We use all these data to estimate a regression model that explains variation in rations with a function of time (season and year fixed effects) and intercepts for specific data characteristics (a baseline effect for call-up-data and a baseline effect for city-level data). Moreover, we include occupational zone fixed effects. The function in time allows us to inter- and extrapolate food rations for each data source type. Furthermore, notice that an average call-up effect can be identified as the average difference between the official food rations and the call-up data. The city effect (recall that we define all city data to be call-up data) can be identified as the average difference between the official rations and the city data net of the already computed average call-up effect. Table 3.2 gives an overview of the imputation model.

Using this imputation we construct two time series per occupational zone. First, we compute a complete time series of official food rations which on average has a relatively high level and which we attribute to all non-urban areas. Second, by subtracting the 'call-up' effect from the city-level series predicted by the imputation model, we deduce a series of

[13] To operationalize and homogenize all our information, we adopt the following conventions: (a) If sources refer only to month and year and do not indicate the rationing period, we infer the ration period to be the one covering most of the month indicated. (b) We distinguish between called-up food rations and official food rations as well as between area-level data that gives rations for countries or zones as whole and city-level data which refers to the specific situation in a given city. (c) Since some food rations are given in daily calories, while others describe the composition of bread, potatoes, meat etc., we standardize all of them using a contemporaneous calorie calculation table (see Schmitz (1956, pp. 469)). Although the official calculation method in the Bi-zonal Area changed between ration period 100 and ration period 101, we stick to the early system.

[14] The rationale behind this is that the data on city level fluctuate a lot between single ration periods and across cities within a zone which is indicative of those numbers to be call-ups.

Table 3.2: Imputation model for predicting food rations

	b	se	t	p
City data (1/0)	-80.16314	23.5164	-3.408818	.0007189
Call-up data (1/0)	-92.62795	25.89259	-3.577393	.0003897
US occupational area (1/0)	-203.3938	81.40028	-2.498687	.0128667
French occupational area (1/0)	-76.80201	85.84196	-.8946907	.3714918
Soviet occupational area (1/0)	-394.8507	91.37267	-4.321322	.0000196
1st quarter of year	-174.5002	21.50112	-8.115867	6.04e-15
2nd quarter of year	-114.8753	23.13227	-4.96602	1.02e-06
3rd quarter of year	31.78962	24.68265	1.287934	.1985154
Year = 1946	-398.7337	317.8371	-1.254522	.2103867
Year = 1947	-240.0152	317.5353	-.7558692	.4501736
Year = 1948	426.0263	86.1683	4.944118	1.13e-06
R^2	0.605			
F-Stat.	25.436			
N	424			

Note: Model also includes zone $\times$ year fixed effects (omitted for brevity).

food rations we attribute to urban areas. This series can be regarded as an 'official' city ration since it is net of a call up effect but still accounting for a 'city-penalty' inherent in our data. Figure 3.4 shows the result of this exercise and provides a means of visual inspection of our identifying variation. In the next step, this information is collapsed to a monthly level and combined with the micro-data.[15] Figure 3.5 shows the variation of the contextual treatment indicators in our micro sample. We can infer that there is a sharp drop in average rations after the war during the pregnancy period and the first year after birth. A relatively large fraction of respondents was actually exposed to rations below 1,500 kcal during either of these critical periods.

Keeping the broad idea in mind of how to align all possible types of data on food rations, we now turn to our sources for area- and city-level ration data.

British and US zone (BBZ and USBZ). The Report of the Military Governor No. 39 (1948) (OMGUS) describes the nutritional quality in US and UK zone, and in the Bi-zonal area after unification. It is the only official source available for the individual nutritional situation. Beside the official rations, also call-up data are given for parallel time spans. This is necessary since food shortages in the Bi-zonal area made it impossible to maintain the official ration level from March 1947 onwards (ZP 99 to ZP 112). As described above, we use this information on both call-ups and official food rations to infer the average difference between them. Thereby, we can also isolate the city effect when using call-up data.

[15] If two separate allocation periods followed each other in one month, we took the value of the one that lasted the longest in this specific month.

Figure 3.4: Imputed rations by occupation zone

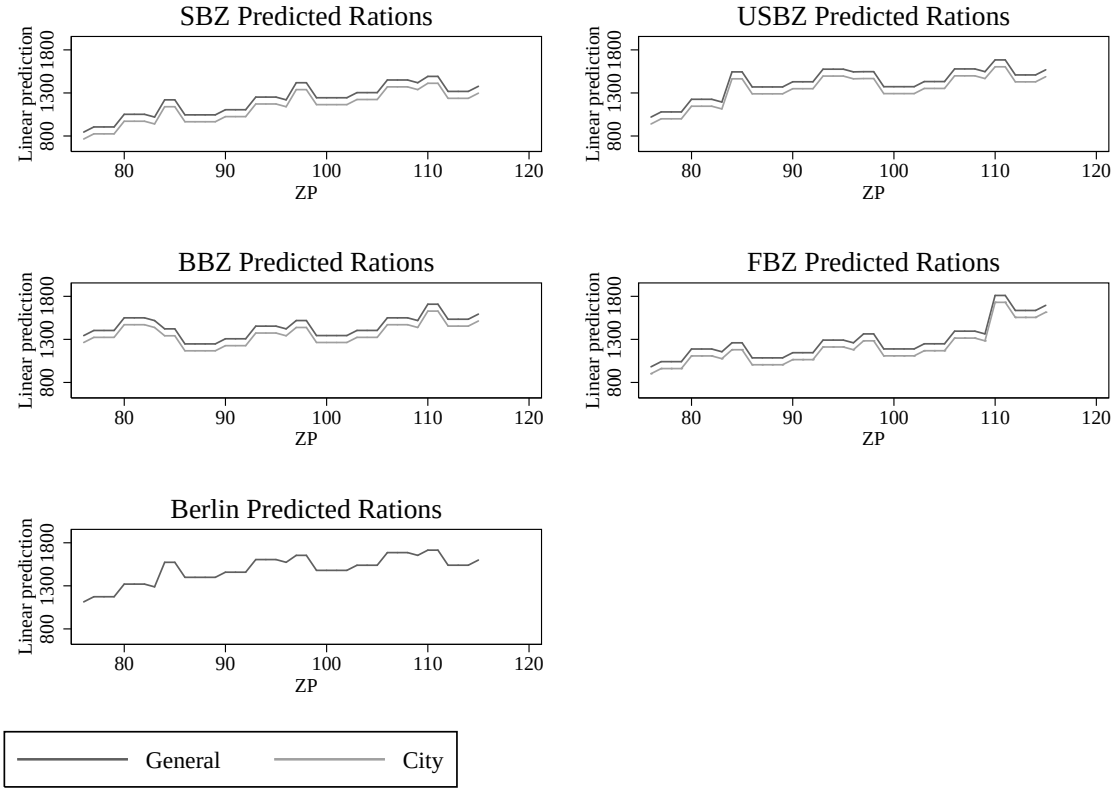
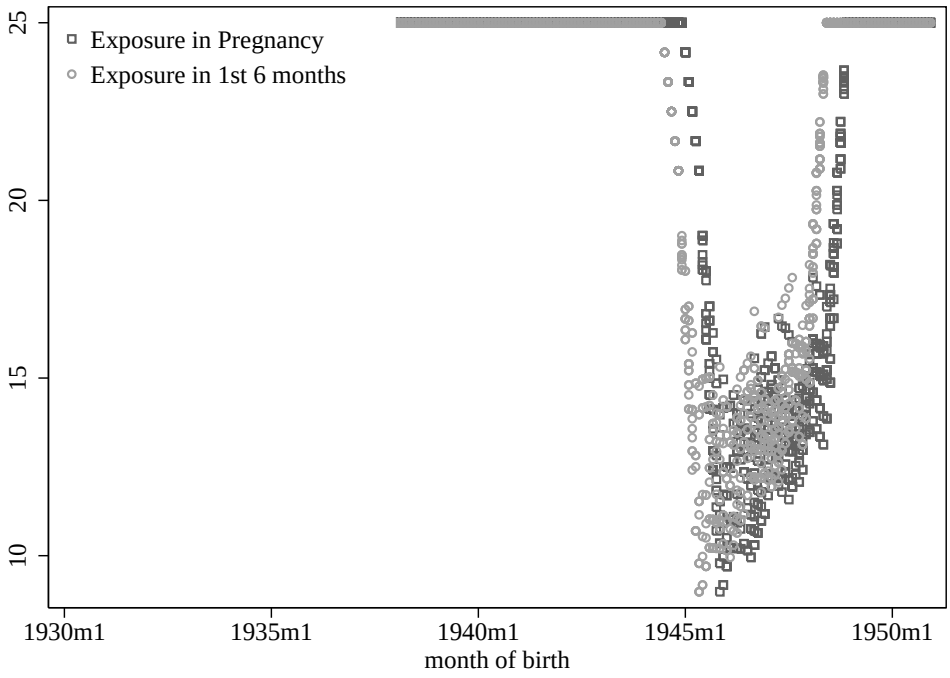


Figure 3.5: Average rations by critical period for SOEP respondents



To extend the time series to periods before March 1946, we refer to Schlange-Schöningen (1955). This source yields monthly averages of daily ration caloric values from April 1945 to December 1949. Due to slightly different definitions, however, the time series for Schleswig-Holstein, Essen, British zone, American zone are only partly comparable especially in the very first time after the breakdown. Additional imprecision might occur due to regional differences in the food quality. We assume that these differences are small enough not to affect our analyses.

French zone (FBZ). The French military government for Germany followed a completely different strategy with respect to the German economy than US and UK. Collaboration between the three powers was thus very difficult in the first years after 1945, and hence also the nutritional situation of the German population was highly unequal. For the food rations in the French zone, we refer to Manz (1985). The nutrition data therein stem from the "Bulletin Statistique, No. 2 - No. 9" and covers October 1945 to December 1948. Since October 1946, the official food rations were dependent on municipality size. Unfortunately, we only have data for cities with less than 20,000 residents.

Soviet zone (SBZ). To our knowledge, the most comprehensive source for nutrition in the Soviet zone is Boldorf (1998), who gives detailed information on the living standard in the soviet occupational zone, later GDR. He describes the official food rations between November 1, 1945 and May 28, 1958. The category of a 'normal consumer' of the Western zones did not exist in the Soviet zone. Instead, the food rationing level discriminated between five ration card groups, depending on occupation and eligibility of the single person.[16] Also the city size mattered for the food rationing level, with rations in cities being higher than in rural regions due to the supposedly easier access to additional resources and higher ability to at least partly provide for their own needs. According to Boldorf (1998, p. 76), the population in small cities or rural regions, which did not have regular access to subsistence food production, was hit hardest by the famine. Other channels of food supply were not accessible for most consumers, also due to extremely high prices. Thus, the food ration level gives a quite accurate picture of the average consumption level. Still, there were periods of crisis until 1949, when food ration cards could not be fully called up.

[16] According to Boldorf (1998, p. 63, 71), we select card type IV as the best proxy for a 'normal consumer' after February 1947, because in 1947/1948 almost half of the rations cards belonged to this type. For the ration periods before, we refer to card type V as 'normal consumer'.

Cities. We use information on the call-up of the official rations in several cities (i.e. with how much food the urban population was actually supplied). Officially, food rations in the cities were as high or even higher as in the countryside, due to the special needs of industrial workers, who were more frequent in urban areas. Furthermore, the low share of people growing their own food as well as the potentially long distance to the next farmer led to sometimes higher rations for city residents (e.g. partly in the French and in the Russian zone). As described in our review of the historical background before, however, the shortfall in food to be transported to the cities led to a much worse food supply in the cities.

The ration data we use to identify the city effect are taken from several sources. First and foremost, we have ration data for Berlin from the Hauptamt für Statistik Berlin-Wilmersdorf (1947, p. 236). Schmitz (1956) gives a very detailed picture on the technical and practical consequences of rationing food and other consumer items for the example of Essen. He collected complete time series data on rations in Essen and Duesseldorf and single data points for Hamburg, Munich, Nuremberg, Frankfurt am Main, Mainz, Kiel, Stuttgart, Wuppertal, Muenster, and Cologne (Schmitz (1956, p. 469)). Nutrition in Hamburg is also described by Boldorf (1998, p. 78). Seel (1950, p. 22) depicts the food situation at the level of single cities of the British zone, i.e. Bochum, Duisburg, and Remscheid. As a last source we consult Martini (1947, p. 31) who shows ration data for Bonn.

3.5 Theoretical exposition and empirical methods

3.5.1 Theoretical exposition

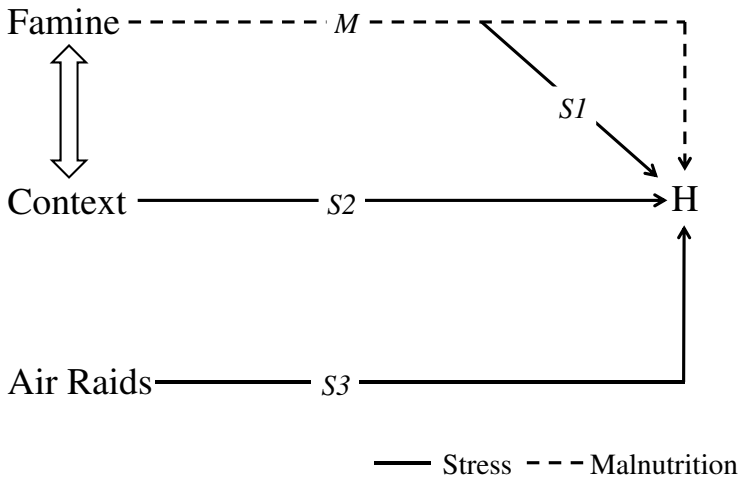
To interpret the findings of our empirical analysis, it is helpful to consider the following stylized theoretical model, in which a health outcome H is a function of bombing and famine exposure:

$$H = f(Bombing, Famine). \tag{3.1}$$

For the sake of simplicity, we restrict our exposition to one single critical period since similar considerations can be easily extended to the whole spectrum of potential critical phases of childhood.

The potential causal channels between famine, air raids, stress, malnutrition at a given critical age and later life health are depicted in figure 3.6 for illustrative purposes. We assume that bombing exposure is a source of stress (S_3) only and famine involves both malnutrition (M) and two possible stress components S_1 and S_2. S_1 is the stress that might be directly

Figure 3.6: Schematical overview over possible effect components



evoked by malnutrition. This stress component could consist of a direct physiological stress reaction as well as a stress reaction to food insecurity (worries etc.).[17] Concerning the post-war famine, Allied observers state that food insecurity and insufficiency were a major cause for worries and cares among the German families, even after food rations had substantially risen to normal levels, and even in the American zone, which was best supplied with food, see Merritt and Merritt (1970). This indicates that the famine did not only affect the population physiologically, but also put a psychologic burden on them. S_2 is a stress component that is evoked by contextual factors that may parallel the famine, such as general social upheavals or crime rates. Assuming additivity of both famine related stress components, it follows that

$$H = f(S_3, (S_1 + S_2, M)). \tag{3.2}$$

We assume an overall linear functional form. Moreover, the different components of stress S_1, S_2, S_3 are occurrences of an overall, homogeneous stress variable S. The only dimension these stress variables are allowed to vary in is quantity, but not quality. This assumption implies that the magnitude of stress originating from different sources might

[17] An empirical analysis that aims to disentangle stress from nutrition channels in famine effects might not want to detach the impact of S_1 from the nutrition effect. Being directly (physiologically or psychologically) related to the individual experience of hunger, it is actually part of the causal effect of malnutrition. However, to the extent that S_1 is an occurrence of a uniform underlying stress concept, our analysis will detach this hunger-related stress from the nutrition effect.

vary. But for a given magnitude all three stress channels are assumed to trigger the same physiological and psychological effects. In terms of our theoretical model, this assumption implies that the impact of S_1, S_2, S_3 on Y is homogeneous for given magnitude (β_1).

Literature on Acute Stress Disorder (ASD) suggests an association of very different kinds of stressful events with this diagnosis (see e.g. Harvey and Bryant, 1999; Biggs et al., 2012).[18] Thus, there might be some comparable immediate stress reaction in response to different stressful events. However, the prevalence rates of ASD differ across traumatic event types (see, e.g., table 6.1 in Biggs et al. (2012)). For instance, Isserlin et al. (2008) suggest that future research needs to examine the role of stressor types for ASD and ASR (Acute Stress Response). Biggs et al. (2012) point to specific risk factors (on the individual level and trauma characteristics) related to ASD.

Assuming a homogeneous impact of all stress component on the outcome, we can rewrite the model to

$$H = \beta_0 + \underbrace{\beta_1 S_3}_{\text{Bombardments}} + \underbrace{\beta_1(S_1 + S_2) + \beta_2 M}_{\text{Famine}} \tag{3.3}$$

and thus

$$H = \beta_0 + \underbrace{\beta_1(S_1 + S_2 + S_3)}_{\text{Stress}} + \underbrace{\beta_2 M}_{\text{Nutrition}} . \tag{3.4}$$

Hence, the outcome is determined by overall exposure to stress and malnutrition early in life.[19]

[18] For a description of ASD we closely follow the American Psychiatric Association (2000, p. 469-472): ASD is characterized by increased arousal directly after experiencing an extreme traumatic situation. The traumatic event often involves actual or threatened death or threatened violation of physical integrity for the person him- or herself or others, causing horror, fear, helplessness. Typical symptoms are the reexperience of the traumatic situation, avoidance and anxiety with respect to similar situations or recollection, distress that is clinically visible, impairment of normal functioning. These symptoms may not be due to physiological or medical effects of the traumatic situation. If the symptoms last for longer than 4 weeks, the diagnosis might be adapted to Post-traumatic Stress Disorder. ASR follows a very similar definitions, see again the description of the American Psychiatric Association (2000) for details.

[19] In the model with multiple critical periods, equation 3.4 can be written as

$$H = \beta_0 + \sum_{a=0}^{A} \underbrace{\beta_{1a}(S_{1a} + S_{2a} + S_{3a})}_{\text{Stress at period } a} + \sum_{a=0}^{A} \underbrace{\beta_{2a} M_a}_{\text{Nutrition at period } a}$$

where periods 0 to A are potentially critical periods in childhood within which an individual is particularly vulnerable to stress and malnutrition. All theoretical considerations can be straightforwardly extended to this extended model.

Model predictions. There are four potential constellations of the effects of bombardments and of famine on later life health. In the following, we elaborate on these cases. Each one has a different implication for the underlying relationships in the context of our model.

First, consider two cases, in which the effect of bombings (stress) is not different from 0 ($\beta_1 = 0$). It follows that the stress component of hunger, no matter of which magnitude it is, has no effect on H either, according to our model. Thus, the effect of famine is a pure nutrition effect: If the famine is found to have an impact (no impact) on the health outcome, only the nutritional shortage has an impact (no impact) on the outcome of interest.

Next, consider the case that famine does not have an effect, but bombings do have an impact on outcome H (i.e. $\beta_1 \neq 0$). In the light of our model, this result would imply that stress does affect H. Since the overall famine effect (nutrition component + stress components) is not different from 0, and we know that stress does have a non-zero effect on the outcome, we can infer that the magnitude of the stress components of famine ($S_1 + S_2$) must be negligibly small in this example if we can assume that the nutrition effect should go in the same direction as the stress effect.[20]

Last, if both bombardments and famine have a non-zero effect on the outcome variable of interest, we cannot draw reliable conclusions on the relevance of stress- or nutrition-related channels for the famine effect. We can only infer that the overall famine effect, i.e. combination of stress and nutrition channels, is different from zero. However, this result can potentially be caused by three candidate constellations. First, the magnitude of $S_1 + S_2$ might be small, leading to the conclusion that the famine effects are only driven by nutrition. The second possibility is that $S_1 + S_2$ are of large magnitude and $\beta_2 M = 0$ (i.e. the famine effect is only stress-driven). The last possible constellation would be a famine effect that is a combination of both, nutrition and stress components.

3.5.2 Empirical model

To disentangle the effects of nutrition and stress exposure in early life, we run a reduced form analysis of health outcomes in late life (in particular adult height) on indicators of prenatal and postnatal exposure to hunger and stress through bombardments. The last section has discussed the imposed structure and theoretical implications of the underlying model. Equation (3.3) has already hinted at the relationship between the structural health production function and the reduced form model that is the focus of our empirical analysis. In general, we want to emphasize that our empirical results only provide suggestive

[20] If there is no theoretical justification for assuming this, an insignificant famine effect could also be explained by counteracting nutrition and stress components which cancel each other out.

evidence on the structural associations of the underlying model because we estimate reduced form relations.

We estimate the following equation:

$$y_i = \alpha + \theta F_i + \phi B_i + \delta X_i + \epsilon_i \tag{3.5}$$

where F_i is a vector that comprises two indicators measuring exposure to famine within six months before birth and within six months after birth, respectively. Exposure to famine is determined by the average food ration to be below 1,500 calories in the respective time frame.[21] Due to the high level of imputation for food rations, we use a binary famine variable to limit the impact of measurement error and noise. B_i denotes a vector that depicts the overall per capita amount of bombs (in tons) on the place of birth within the last two trimesters of pregnancy and within the first six months of life, respectively. In contrast to the food rations, here we can make use of the actual bomb load, since data quality is very high. X_i is a collection of control variables: a linear trend in the individual's birth year, an indicator for whether the individual was born in a city, and a dummy that indicates whether the father's education is/was high.[22] Last, ϵ_i is an i.i.d. error component.

Using this estimation approach, we exploit temporal and regional variation in dropped bomb loads and food rations to identify the reduced form effect of hunger and stress exposure on health in late adulthood. Our identification strategy relies on the assumption that, given the covariates that are included in our model, the place and month of birth are exogenous. Including the information on growing up in a city, we control for systematic differences between rural and urban areas.

3.5.3 Identification issues

As mentioned above, our identification strategy relies on the exogeneity of the place of birth. However, this exogeneity assumption could be violated if parental moving decisions before the birth of their children were made in a systematic way. As an example, parents could choose their place of living according to the city's destruction level. We address these potential concerns in a sensitivity analysis in Section 3.6.2.

[21] The prenatal and postnatal age window should be of the same length since we define famine exposure as a function of the average ration over those intervals. If they were of a different length it would be possible that we would define an individual as not famine exposed after birth while another individual being in utero at the very same instant would potentially be defined as famine exposed (and vice versa).

[22] The variable takes the value 1 if the father's education is Realschule, Fachhochschule, or Abitur. It takes the value 0 if the father's education is indicated as unknown, Hauptschule, other degree, or no degree.

Moreover, parental moving decisions after the birth of a child could also cause a problem for identification. If e.g. heavy bombing influenced the decision to move to a place with better living conditions (e.g. the countryside) in the first year after the birth of a child, our estimated treatment effects would be biased. Such movings could have taken place within the framework of evacuation programs, as described in Section 3.3.2, or be privately organized. In particular, a tendency to move away from heavily bombed cities to the countryside after birth could lead to underestimated treatment effects in the study at hand.

As pointed out in, e.g., van den Berg et al. (2011), further potential threats to validity in famine studies are selective fertility and selective mortality. van den Berg et al. (2011) argue that fertility reduction during famines is presumably more prevalent among families with lower socio-economic status. Controlling for a proxy of socio-economic status during childhood (father's education), we already address this endogeneity problem that is caused by differential fertility reduction in our main analysis.

Moreover, as discussed in Section 3.3.2, our sample might be positively selected since working-class neighborhoods were common civilian bombing targets. Thus, it is likely that the survival under air raids depends on socio-economic background. Controlling for the father's education level, a proxy of socio-economic status, we also address this concern in all our analyses. However, we cannot address the general issue of selective mortality (see, e.g., the discussion of selectivity in famine studies in van den Berg et al. (2011)) that is present in samples of elderly people. Since the frailest individuals are likely to have died already before the survey was conducted, our estimates are likely to understate (in absolute terms) the true treatment effects.

3.6 Regression results

3.6.1 Main results

Results of our estimations are given in Table 3.3 for males and Table 3.4 for females. Considering the regression results for men, we find evidence for adult height being significantly reduced when rations were low during pregnancy. This finding is common in the famine literature and with a reduced form effect of about -1.3 centimeters comparably large in our study. Bombings do not exert a significant impact on height even though the point estimates are relatively large. This, however, is a mechanical effect of the small scale of our bombings variables and the relatively large scale of adult height. In summary,

Table 3.3: Males – Main results

Variable	Hypertension	Height	log(Pension)	Diabetes
Less than 1500 kcal during gestation (1/0)	0.050	-1.344*	-0.189**	0.029
	(0.052)	(0.734)	(0.087)	(0.050)
Less than 1500 kcal 1st 6 months (1/0)	-0.016	0.515	0.102	-0.012
	(0.054)	(0.729)	(0.085)	(0.053)
Tons of bombs p.c. dropped during gestation	-5.084	-63.931	6.527	12.588
	(8.338)	(103.475)	(9.445)	(8.174)
Tons of bombs p.c. dropped in 1st 6 months	2.965***	-0.925	2.864	-2.463
	(0.932)	(8.814)	(6.248)	(2.638)
Year of Birth	-0.006***	0.219***	0.008*	-0.005**
	(0.002)	(0.027)	(0.004)	(0.002)
Father's education high	-0.026	2.324***	0.292***	-0.072***
	(0.029)	(0.369)	(0.052)	(0.023)
Place of birth: pop. >= 10000	-0.006	0.580*	0.127***	0.014
	(0.023)	(0.296)	(0.035)	(0.021)
Constant	12.509***	-248.990***	-7.620	8.924**
	(4.074)	(51.947)	(7.928)	(3.605)
R^2	0.006	0.059	0.068	0.011
F-Stat.	3.009	17.038	10.137	2.836
N	1981	1985	1111	1439

Note: Robust standard errors in parentheses. *,**,*** indicates significance at the 10%, 5%, and 1% level, respectively.

Table 3.4: Females – Main results

Variable	Hypertension	Height	Diabetes
Less than 1500 kcal during gestation (1/0)	-0.060	-0.077	-0.010
	(0.048)	(0.543)	(0.029)
Less than 1500 kcal 1st 6 months (1/0)	0.069	-0.998*	-0.006
	(0.050)	(0.555)	(0.030)
Tons of bombs p.c. dropped during gestation	0.101	6.295	-0.387
	(3.218)	(27.098)	(1.132)
Tons of bombs p.c. dropped in 1st 6 months	2.351	70.273	-1.989
	(5.108)	(53.363)	(1.822)
Year of Birth	-0.009***	0.140***	-0.005***
	(0.002)	(0.024)	(0.002)
Father's education high	-0.091***	1.235***	-0.033
	(0.029)	(0.342)	(0.021)
Place of birth: pop. >= 10000	0.036	-0.267	0.004
	(0.022)	(0.266)	(0.018)
Constant	18.795***	-107.979**	8.993***
	(4.030)	(46.490)	(3.073)
R^2	0.020	0.026	0.010
F-Stat.	5.871	8.227	2.278
N	2038	2039	1390

Note: Robust standard errors in parentheses. *,**,*** indicates significance at the 10%, 5%, and 1% level, respectively.

our results imply that a reduction in height captures malnutrition rather than stress. Since height has been found to correlate positively with economic outcomes and cognitive skills (see, e.g., Case and Paxson (2008, 2010)), it is conceivable that famine and thereby implied malnutrition has a profound impact on the overall economic prosperity of an individual. In fact, this hypothesis is corroborated by our findings for pension income of individuals. We find that pensions are reduced by almost 20 percent in case of prenatal exposure to famine. This also hints at the importance of famine induced health conditions or deficits in other dimensions of human capital for the productivity level over the life cycle. The effects of air raids and thus stress early in life are not significant which may point at stress being of rather low importance for overall productivity. A possible explanation is that early life stress related conditions may also be prevalent only very late in life, i.e. without productivity reducing effects during the professional life of an individual. Since famine exposed individuals could have lower pensions by the mere fact that a selected group with lower pension levels retired already due to famine exposure, we checked whether the probability to be retired differs for famine exposed individuals (results are available upon request). In these analyses we do not find a significant relationship between famine exposure and the retirement decision.

Another striking result is that we find postnatal experience of bombings to raise the risk for hypertension among men. This finding supports the biological explanatory frameworks discussed before. The experience of famine around birth is not significantly related to hypertension risk even though the coefficient on the dummy indicating prenatal exposure to famine is relatively large (5 percentage points). The coefficients on our bombing measure cannot be straightforwardly interpreted. Taking the average per capita bomb load for our sample (conditional on having experienced an air raid in the first six months of life) and multiplying the coefficient found in our analysis yields an average air raid-induced excess hypertension risk of about 2 percentage points. Relative to the overall hypertension prevalence among men in our sample (46 percent), this amounts to an average increase of 4 percent.

Moreover, we may quantify the effects of air raids with the help of further illustrative 'back-of-the-envelope' calculations. To this end, consider the cities of Berlin and Cologne which were heavily bombed during the war. We then compute the absolute per capita bomb load dropped on both cities for the six months following a given month during the war. Multiplying this measure with the coefficient found in our analysis then gives the excess hypertension risk for a hypothetical individual born in one of those cities during the war.

Figure 3.7: Predicted effects of air raids on hypertension risk over the course of the War for Berlin and Cologne

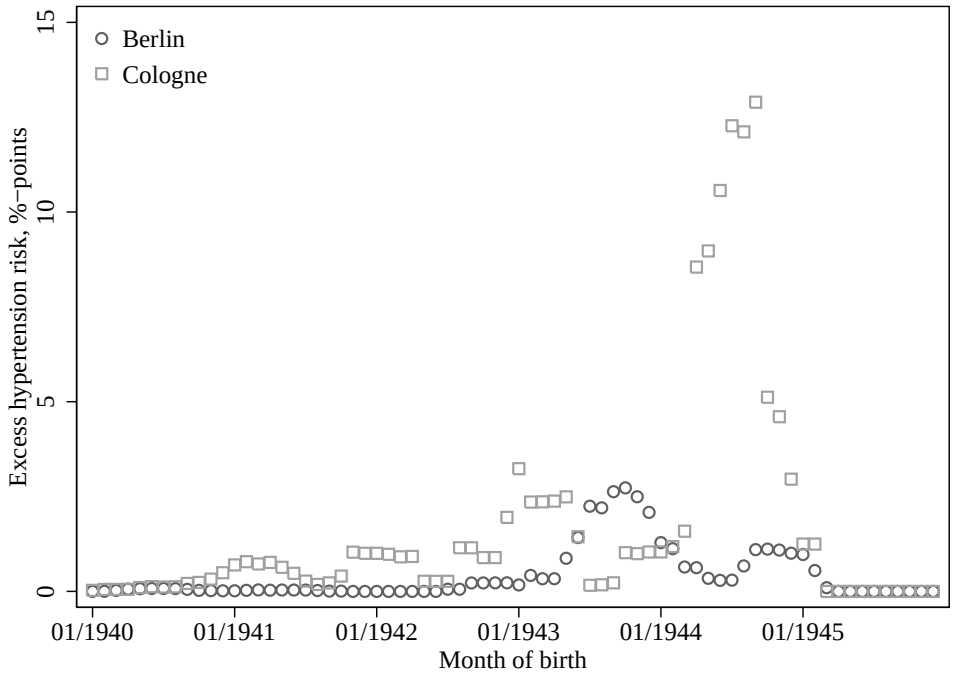


Figure 3.7 plots the result of this exercise. The figure implies that, depending on the month of birth, air raids may have induced an elevated hypertension risk of up to 3 percentage points for Berlin and 14 percentage points for Cologne. Relative to the overall average prevalence of hypertension in our male sample of 46 percent, these effects amount to an increase in hypertension risk of about 6.5 percent for Berlin and 30 percent for Cologne. This shows that the effects on hypertension can be sizable.

Turning to the results for women, we find that they are less frail to adverse early life conditions. This finding regarding the gender differences in early life impacts on later life morbidity and mortality has been noted throughout the medical literature (e.g. Waldron (1983), Kraemer (2000)). Still, we find that women who experienced the famine in the first six months of their lives are about one centimeter shorter. As we also find famine exposure to be particularly influential on adult height among men, this again points at the importance of the nutritional situation early in life for adult height. Again, this might be indicative of malnutrition within this age window to be a decisive factor also for cognitive skills and overall productivity.

Let us now consider the results for diabetes. Running analyses on the full samples gave counter-intuitive results regarding the association between air raid exposure and diabetes

risk (the event of air raids was negatively related with diabetes risk). It turns out that diabetes prevalence in Germany exhibits a significant west-east gradient with the Eastern German population having higher rates of type II diabetes (Moebus et al. (2008)). Our original counter-intuitive result may thus be driven by regional factors. We thus decided to drop individuals born in Eastern Germany which yields the results displayed in tables 3.3 and 3.4. Here, we do not find evidence for either of our measures of early life shocks to be significantly related to diabetes.

It is crucial to note that the factors inherent in the Eastern German sub-sample that drove the counter-intuitive effects of air raids on diabetes risk only apply to this outcome. Consider tables 3.5 and 3.6 where we run all analyses on the Western German sample. In this restricted West German sample, we still find evidence for a height effect of prenatal famine on male adult height and of bombing exposure during the first six months of life on hypertension risk among men. Moreover, for women we still find a negative impact of postnatal famine on adult height. This means that all other main results remain unaffected (despite the smaller sample size, the effects of prenatal famine on height in the male sample and the effect of postnatal famine on female height are now estimated even more precisely).

The fact that other results remain unaffected by this stratification exercise can be explained by other conditions not exhibiting an east-west gradient. Moebus et al. (2008) show that metabolic syndrome is more prevalent among the Eastern population while, however, risk factors other than diabetes (like hypertension) are not. It has to be noted that the study population in Moebus et al. (2008) is younger than in our sample and mostly born after the post-war famine.[23] It is thus likely that the regional differences we observe for diabetes are not driven by our main explanatory variables but rather by factors that were present in Eastern Germany later in life and specifically impacted on the diabetes risk of post-war cohorts (but not on other health conditions).

The estimated coefficients for the control variables are in line with the literature: The younger the individuals and the better educated their fathers (i.e. the higher the socio-economic status), the healthier the respondents appear to be. Moreover, the education of a respondent's father is a decisive factor for health risks and pension income. All relationships are such that individuals with a more highly educated father are healthier.

[23] Their sample was drawn in 2005 and men (women) in the East and West were on average 54 (52) and 52,8 (50,6) years old, respectively, i.e. born in 1951 or later.

Table 3.5: Males – Main results excluding observations from Eastern Germany

Variable	Hypertension	Height	log(Pension)	Diabetes
Less than 1500 kcal during gestation (1/0)	0.057	-1.704**	-0.189**	0.029
	(0.060)	(0.862)	(0.087)	(0.050)
Less than 1500 kcal 1st 6 months (1/0)	-0.066	0.533	0.102	-0.012
	(0.062)	(0.854)	(0.085)	(0.053)
Tons of bombs p.c. dropped during gestation	-11.196	6.994	6.527	12.588
	(7.252)	(132.200)	(9.445)	(8.174)
Tons of bombs p.c. dropped in 1st 6 months	7.239**	-33.522	2.864	-2.463
	(3.124)	(50.584)	(6.248)	(2.638)
Year of Birth	-0.007***	0.217***	0.008*	-0.005**
	(0.002)	(0.030)	(0.004)	(0.002)
Father's education high	-0.011	2.431***	0.292***	-0.072***
	(0.033)	(0.425)	(0.052)	(0.023)
Place of birth: pop. >= 10000	-0.007	0.759**	0.127***	0.014
	(0.027)	(0.349)	(0.035)	(0.021)
Constant	13.292***	-246.091***	-7.620	8.924**
	(4.703)	(59.201)	(7.928)	(3.605)
R^2	0.008	0.063	0.068	0.011
F-Stat.	2.351	13.387	10.137	2.836
N	1439	1442	1111	1439

Note: Robust standard errors in parentheses. *,**,*** indicates significance at the 10%, 5%, and 1% level, respectively.

Table 3.6: Females – Main results excluding observations from Eastern Germany

Variable	Hypertension	Height	Diabetes
Less than 1500 kcal during gestation (1/0)	-0.047	0.609	-0.010
	(0.056)	(0.603)	(0.029)
Less than 1500 kcal 1st 6 months (1/0)	0.090	-1.449**	-0.006
	(0.059)	(0.626)	(0.030)
Tons of bombs p.c. dropped during gestation	-1.228	-3.324	-0.387
	(3.492)	(30.531)	(1.132)
Tons of bombs p.c. dropped in 1st 6 months	3.767	48.717	-1.989
	(5.161)	(54.522)	(1.822)
Year of Birth	-0.012***	0.151***	-0.005***
	(0.002)	(0.028)	(0.002)
Father's education high	-0.052	1.395***	-0.033
	(0.034)	(0.414)	(0.021)
Place of birth: pop. >= 10000	0.020	-0.110	0.004
	(0.027)	(0.319)	(0.018)
Constant	23.518***	-129.316**	8.993***
	(4.793)	(54.970)	(3.073)
R^2	0.022	0.033	0.010
F-Stat.	4.403	7.051	2.278
N	1390	1390	1390

Note: Robust standard errors in parentheses. *,**,*** indicates significance at the 10%, 5%, and 1% level, respectively.

3.6.2 Robustness checks

Endogeneity of place of birth. One of the major threats to our empirical strategy is that the place of birth might be selective. This is problematic especially for the people born during the war and thus potentially being selectively evacuated to evade bombings. We first tackle this matter by analyzing the probability of being born in a city which has been bombed during the last two trimesters of pregnancy. Using linear probability models we regress this variable on a set of regressors, most notably a function in the cumulative bomb load dropped on this specific place of birth until 9 months prior to a respondent's birth. Notice that this variable varies over time for a given city, depending on how often and how strongly this city has been bombed. This variable is non-zero for post-war cohorts and we use this measure as a general proxy for war-time destruction of a municipality in later analyses. We also add those variables interacted with a dummy for being born during the war to differentiate between war and post-war cohorts.

Table 3.7 shows the results of this exercise. We can infer that there is a statistically significant association between cumulative bomb loads dropped on a respondent's place of birth just prior to conception and the probability to be born in a city that has been air raid stricken during pregnancy and that this relationship is non-linear. The non-linearity could be indicative of a saturation effect: for large cumulative bomb loads dropped on a city and later periods of the war, capacities and infrastructure for evacuating expectant mothers might have been exhausted. Selective fertility might play a role here as well.

Second, we use supplementary data from a specific SOEP module that deals with the early life circumstances of most of the SOEP respondents born 1935 to 1950. From this module we have information on whether the respondents' mother moved during pregnancy or whether they lived in a place different from their place of birth during the first 6 months of their lives. Note that the probability for a move during pregnancy should actually correlate negatively with the bomb load measure associated with the actually realized place of birth (an outcome realized only after the move). If they are mobile, mothers should select their residence based on security concerns. The first three columns of Table 3.8 confirm this reasoning. Note that here the model fit is better when including the selection variables linearly. The last set of results shows the correlation between overall bomb loads dropped on the place of birth until birth and the probability for a move within 6 months after birth. Here, the results may be easier to interpret and we expect mothers to move away from dangerous places. This is in fact the case since the coefficient on the variable measuring the

Table 3.7: Results of linear probability models estimating the probability to be born in a city that has been bombed during pregnancy

Variable	Males	Females
Overall tons p.c.× born in war	91.213***	276.552***
	(12.224)	(33.546)
Overall tons p.c. squared × born in war	-1312.803***	-1.47e+04***
	(210.998)	(2829.235)
Overall tons p.c. before pregnancy	-1.284***	-1.279***
	(0.147)	(0.247)
Overall tons p.c. before pregnancy squared	2.482***	2.436***
	(0.301)	(0.637)
Born during war	0.142***	0.125***
	(0.013)	(0.013)
Year of Birth	0.003***	0.003***
	(0.000)	(0.000)
Father's education high	0.012	0.022
	(0.014)	(0.016)
Place of birth: pop. >= 10000	0.088***	0.104***
	(0.010)	(0.010)
Constant	-5.910***	-6.532***
	(0.766)	(0.941)
F-Stat.	29.788	39.797
N	1992	2041

Note: Robust standard errors in parentheses. *,**,*** indicates significance at the 10%, 5%, and 1% level, respectively.

Table 3.8: Results of linear probability models estimating the probability to move given bomb load dropped on place of birth

	Moved during pregnancy			Moved 6 months after birth		
Variable	Pooled	Males	Females	Pooled	Males	Females
Overall tons p.c.× born in war	-1.748**	-2.203***	0.501			
	(0.775)	(0.830)	(4.927)			
Overall tons p.c. before pregnancy	0.348	0.656	-0.205			
	(0.251)	(0.446)	(0.235)			
Born during war	0.007	-0.000	0.012	0.032**	0.031	0.026
	(0.009)	(0.012)	(0.012)	(0.015)	(0.022)	(0.021)
Year of Birth	0.003***	0.003**	0.003***	0.004**	0.003	0.004**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Father's education high	0.006	0.015	-0.004	0.056***	0.041	0.070**
	(0.011)	(0.017)	(0.014)	(0.020)	(0.028)	(0.028)
Place of birth: pop. >= 10000	-0.010	-0.002	-0.016	0.009	-0.015	0.032
	(0.009)	(0.013)	(0.013)	(0.015)	(0.022)	(0.020)
Overall tons p.c.× born in war				6.804***	5.387**	12.022*
				(2.269)	(2.565)	(6.895)
Overall tons p.c. before birth				-0.565**	-0.384	-0.902**
				(0.225)	(0.250)	(0.394)
Constant	-5.528***	-4.965**	-6.517***	-7.274**	-6.221	-8.484**
	(1.592)	(2.222)	(2.395)	(2.887)	(4.293)	(3.904)
F-Stat.	3.996	2.583	2.783	5.826	2.345	4.233
N	2608	1271	1337	2608	1271	1337

Note: Robust standard errors in parentheses. *,**,*** indicates significance at the 10%, 5%, and 1% level, respectively.

Table 3.9: Males – Results when including proxies for degree of selectivity

Variable	Hypertension	Height	log(Pension)	Diabetes
Less than 1500 kcal during gestation (1/0)	0.036	-1.276*	-0.198**	0.024
	(0.053)	(0.745)	(0.092)	(0.051)
Less than 1500 kcal 1st 6 months (1/0)	-0.017	0.503	0.099	-0.012
	(0.053)	(0.732)	(0.087)	(0.052)
Tons of bombs p.c. dropped during gestation	-10.532	152.897	4.228	-6.655
	(13.351)	(190.493)	(14.345)	(10.671)
Tons of bombs p.c. dropped in 1st 6 months	3.335***	-9.987	2.933	-2.408
	(1.079)	(12.167)	(6.288)	(2.135)
Year of Birth	-0.006***	0.216***	0.008*	-0.004**
	(0.002)	(0.028)	(0.005)	(0.002)
Overall tons p.c.× born in war	6.513	-222.609*	1.851	16.299**
	(8.886)	(114.482)	(9.651)	(7.349)
Overall tons p.c. before pregnancy	0.262	1.210	0.188	-0.104
	(0.485)	(7.588)	(1.474)	(0.563)
Born during war	-0.031	0.253	-0.011	-0.019
	(0.025)	(0.317)	(0.044)	(0.023)
Father's education high	-0.025	2.303***	0.293***	-0.070***
	(0.029)	(0.369)	(0.052)	(0.023)
Place of birth: pop. >= 10000	-0.009	0.578*	0.126***	0.015
	(0.024)	(0.304)	(0.036)	(0.022)
Constant	12.520***	-244.470***	-8.703	8.528**
	(4.212)	(53.761)	(9.514)	(3.793)
R^2	0.008	0.061	0.068	0.015
F-Stat.	2.363	12.918	7.135	4.355
N	1981	1985	1111	1439

Note: Robust standard errors in parentheses. *,**,*** indicates significance at the 10%, 5%, and 1% level, respectively.

cumulative bomb load and the dummy for being born during the war is found to be positive and significant.

Building on the previous results and having shown that cumulative bomb load is an informative variable regarding the degree of selection, we proceed as follows. We re-run the main regressions including the cumulative bomb load measures in all regressions. We thus compare individuals who may or may not have been bombed during a potentially critical period and who may exhibit similar places of birth concerning cumulative bomb loads, i.e. that may be equivalently selected. Tables 3.9 and 3.10 show the results. We still find that hypertension risk among men is adversely affected by the postnatal experience of air raids. The effects of famine on male adult height and pension income remain robust as well. Also for women, all results noted already in our main analysis remain unchanged.

Exposure to air raids and famine at later childhood ages. In another robustness check we consider exposure to adverse circumstances also at higher childhood ages. This could be problematic since other studies have shown that outcomes like adult height, blood pressure, or cause-specific mortality are affected by exposure to nutritional shortages in

Table 3.10: Females – Results when including proxies for degree of selectivity

Variable	Hypertension	Height	Diabetes
Less than 1500 kcal during gestation (1/0)	-0.074	0.115	0.003
	(0.049)	(0.561)	(0.031)
Less than 1500 kcal 1st 6 months (1/0)	0.067	-0.998*	-0.005
	(0.050)	(0.556)	(0.030)
Tons of bombs p.c. dropped during gestation	0.064	6.954	-0.413
	(3.211)	(26.328)	(1.117)
Tons of bombs p.c. dropped in 1st 6 months	1.179	74.671	-2.674
	(5.480)	(48.732)	(2.027)
Year of Birth	-0.009***	0.135***	-0.004***
	(0.002)	(0.025)	(0.002)
Overall tons p.c.× born in war	15.322	-134.985	-0.848
	(13.276)	(127.181)	(9.238)
Overall tons p.c. before pregnancy	0.266	0.902	-0.345
	(0.606)	(9.558)	(0.522)
Born during war	-0.033	0.505*	0.021
	(0.025)	(0.301)	(0.021)
Father's education high	-0.090***	1.217***	-0.033
	(0.029)	(0.341)	(0.021)
Place of birth: pop. >= 10000	0.035	-0.303	0.007
	(0.023)	(0.285)	(0.020)
Constant	18.784***	-99.606**	8.550***
	(4.162)	(49.225)	(3.211)
R^2	0.021	0.028	0.011
F-Stat.	4.509	6.285	2.090
N	2038	2039	1390

Note: Robust standard errors in parentheses. *,**,*** indicates significance at the 10%, 5%, and 1% level, respectively.

later childhood, most importantly just before puberty (e.g. van den Berg et al. (forthcoming), Sparén et al. (2004)). There is also evidence that exposure to famine or nutritional shortages around this period leads to epigenetic changes and thereby even influences outcomes of offspring which also points at the plasticity around that age (see, e.g., Kaati et al. (2002), van den Berg and Pinger (2014) and references therein). If air raids or famine had a sizable impact on the later life health outcomes under study, our control group's outcomes could be contaminated and the effect measured in our analysis would be biased. This problem is aggravated by the fact that those exposed to air raids around birth were affected by the famine at a later but potentially still vulnerable age.

We conduct the following robustness check. We construct four additional explanatory variables: a first set displaying the bomb load dropped on a respondent's place of birth in the age interval from 7 to 48 months and from age 8 to 10; second, two indicators for whether a respondent was exposed to a six month interval with average food rations below 1,500 kcal sometime in these intervals. Note that exposure to both adverse conditions is determined by the place of birth and not the residence at the specific age. We should also

Table 3.11: Males – Results when including exposure to air raids and famine in later age windows

Variable	Hypertension	Height	log(Pension)	Diabetes
Less than 1500 kcal during gestation (1/0)	0.047	-1.246*	-0.186**	0.027
	(0.053)	(0.740)	(0.087)	(0.051)
Less than 1500 kcal 1st 6 months (1/0)	-0.007	0.084	0.100	-0.011
	(0.057)	(0.768)	(0.086)	(0.054)
Less than 1500 kcal months 7 to 36 (1/0)	-0.016	0.730*	0.010	-0.004
	(0.030)	(0.383)	(0.048)	(0.027)
Less than 1500 kcal age 8 to 10 (1/0)	-0.002	0.286	0.049	-0.002
	(0.032)	(0.394)	(0.046)	(0.031)
Tons of bombs p.c. dropped during gestation	-4.200	-96.512	7.091	12.898
	(8.535)	(100.558)	(9.585)	(8.105)
Tons of bombs p.c. dropped in 1st 6 months	3.164***	-1.437	3.118	-2.324
	(1.063)	(10.335)	(6.366)	(2.645)
Tons of bombs p.c. dropped in months 7 to 36	-0.219	-3.558	-0.030	-0.417
	(0.739)	(8.084)	(1.020)	(0.474)
Tons of bombs p.c. dropped age 8 to 10	0.198	7.146	3.172***	-0.442
	(0.791)	(7.772)	(0.899)	(0.442)
Year of Birth	-0.006***	0.219***	0.010**	-0.005**
	(0.002)	(0.029)	(0.005)	(0.002)
Father's education high	-0.026	2.298***	0.288***	-0.072***
	(0.029)	(0.369)	(0.052)	(0.023)
Place of birth: pop. >= 10000	-0.006	0.591*	0.115***	0.018
	(0.024)	(0.304)	(0.037)	(0.022)
Constant	12.079***	-251.133***	-11.625	9.063**
	(4.353)	(55.653)	(9.018)	(3.928)
R^2	0.007	0.061	0.075	0.012
F-Stat.	1.953	11.151	8.080	1.985
N	1981	1985	1111	1439

Note: Robust standard errors in parentheses. *,**,*** indicates significance at the 10%, 5%, and 1% level, respectively.

note that the effects of exposure to adverse conditions at later ages are not straightforwardly interpretable since non-biological channels become more relevant at later ages. For instance, famine and social unrest around puberty could limit access to education and thereby change intermediate socio-economic outcomes which then may determine health outcomes later in life. As shown in tables 3.11 and 3.12, including those additional variables into our models does not change our main results.

Further robustness checks. As noted in Section 3.3.2, almost half of the 600,000 casualties that can be attributed directly to air raids died in one of the 'firestorms' over Hamburg, Kassel, Darmstadt, and Dresden. Hence, we check whether our results change when dropping all respondents who were born in one of those cities over the whole observational period to address the issue of selective mortality. Tables 3.13 and 3.14 show the results of this robustness check. The results of our main analysis turn out to be robust and oftentimes gain precision.

Table 3.12: Females – Results when including exposure to air raids and famine in later age windows

Variable	Hypertension	Height	Diabetes
Less than 1500 kcal during gestation (1/0)	-0.064	-0.002	-0.009
	(0.048)	(0.546)	(0.029)
Less than 1500 kcal 1st 6 months (1/0)	0.076	-1.408**	-0.023
	(0.054)	(0.594)	(0.036)
Less than 1500 kcal months 7 to 36 (1/0)	-0.014	0.603*	0.025
	(0.030)	(0.349)	(0.027)
Less than 1500 kcal age 8 to 10 (1/0)	-0.023	0.089	-0.004
	(0.032)	(0.372)	(0.027)
Tons of bombs p.c. dropped during gestation	0.114	4.393	-0.496
	(3.235)	(27.397)	(1.131)
Tons of bombs p.c. dropped in 1st 6 months	3.020	52.459	-2.767
	(5.201)	(55.009)	(1.925)
Tons of bombs p.c. dropped in months 7 to 36	-0.237	-5.646	-0.320
	(0.750)	(7.829)	(0.476)
Tons of bombs p.c. dropped age 8 to 10	1.592**	-3.479	-0.405
	(0.725)	(9.742)	(0.474)
Year of Birth	-0.009***	0.135***	-0.005***
	(0.002)	(0.025)	(0.002)
Father's education high	-0.089***	1.221***	-0.033
	(0.029)	(0.341)	(0.021)
Place of birth: pop. >= 10000	0.035	-0.251	0.008
	(0.023)	(0.274)	(0.019)
Constant	18.544***	-99.044**	9.818***
	(4.332)	(49.427)	(3.356)
R^2	0.021	0.028	0.011
F-Stat.	4.500	5.611	1.643
N	2038	2039	1390

Note: Robust standard errors in parentheses. *,**,*** indicates significance at the 10%, 5%, and 1% level, respectively.

Table 3.13: Males – Results when dropping individuals born in a city affected by a firestorm

Variable	Hypertension	Height	log(Pension)	Diabetes
Less than 1500 kcal during gestation (1/0)	0.050	0.029	-1.516**	-0.186**
	(0.053)	(0.053)	(0.737)	(0.089)
Less than 1500 kcal 1st 6 months (1/0)	-0.014	-0.024	0.743	0.094
	(0.055)	(0.054)	(0.735)	(0.088)
Tons of bombs p.c. dropped during gestation	-4.155	10.285	-89.949	12.331
	(8.679)	(9.101)	(102.958)	(11.177)
Tons of bombs p.c. dropped in 1st 6 months	2.870***	-2.433	1.225	2.337
	(0.933)	(2.706)	(8.586)	(6.224)
Year of Birth	-0.006***	-0.004**	0.206***	0.008*
	(0.002)	(0.002)	(0.027)	(0.004)
Father's education high	-0.027	-0.075***	2.406***	0.301***
	(0.029)	(0.023)	(0.378)	(0.053)
Place of birth: pop. >= 10000	-0.003	0.012	0.643**	0.132***
	(0.024)	(0.021)	(0.300)	(0.036)
Constant	11.795***	7.992**	-224.767***	-7.791
	(4.135)	(3.671)	(52.505)	(8.006)
R^2	0.006	0.011	0.058	0.072
F-Stat.	2.821	2.600	16.182	10.497
N	1913	1388	1917	1069

Note: Robust standard errors in parentheses. *,**,*** indicates significance at the 10%, 5%, and 1% level, respectively.

Table 3.14: Females – Results when dropping individuals born in a city affected by a firestorm

Variable	Hypertension	Height	Diabetes
Less than 1500 kcal during gestation (1/0)	-0.056	-0.011	-0.065
	(0.048)	(0.030)	(0.547)
Less than 1500 kcal 1st 6 months (1/0)	0.063	-0.003	-0.957*
	(0.051)	(0.031)	(0.560)
Tons p.c. dropped during gestation	0.182	-0.474	5.612
	(3.154)	(1.143)	(27.402)
Tons p.c. dropped in 1st 6 months	0.519	-1.727	73.411
	(5.311)	(2.022)	(58.449)
Year of Birth	-0.010***	-0.005***	0.146***
	(0.002)	(0.002)	(0.024)
Father's education high	-0.091***	-0.032	1.239***
	(0.029)	(0.022)	(0.353)
Place of birth: pop. >= 10000	0.036	0.007	-0.281
	(0.023)	(0.019)	(0.269)
Constant	19.059***	8.961***	-119.319**
	(4.084)	(3.142)	(46.823)
R^2	0.020	0.009	0.028
F-Stat.	5.729	2.054	8.356
N	1970	1344	1971

Note: Robust standard errors in parentheses. *,**,*** indicates significance at the 10%, 5%, and 1% level, respectively.

A further concern could be that unobserved cohort-specific factors may bias our results. This may be the case if, for instance, cohort specific shocks to selective child mortality or selective fertility are determinants of late life health and on the other hand interact with the exposure to air raids. Including cohort fixed effects in our regression is very demanding on the data and on the power of our identifying variation and we thus suspect these alterations to yield at most suggestive evidence. Our measures of exposure to air raids should be less sensitive to this shortcoming since they exhibit a larger degree of cross-sectional variation within the war years. As a robustness check, we thus run another set of regressions where we change our main specification by replacing the linear trend in year of birth with cohort fixed effects. For this change in model specification our main results qualitatively remain the same (see tables 3.15 and 3.16). The effects of prenatal famine on pensions among men turn out to be insignificant. It is likely that this is due to a loss in identifying variation. Recall that for these regressions we drop respondents born in Eastern Germany and that a relatively large fraction of the later born cohorts, i.e. those exposed to the famine, are not retired yet, i.e. cannot be used for these regressions. Both factors limit the cross-sectional variation in famine exposure within a birth cohort and are thus likely to be responsible for a loss in significance.

Table 3.15: Males – Results when including cohort fixed effects

Variable	Hypertension	Height	log(Pension)	Diabetes
Less than 1500 kcal during gestation (1/0)	-0.039	-1.720*	-0.117	-0.034
	(0.069)	(0.883)	(0.114)	(0.064)
Less than 1500 kcal 1st 6 months (1/0)	-0.008	-0.293	0.063	-0.015
	(0.076)	(1.069)	(0.120)	(0.069)
Tons of bombs p.c. dropped during gestation	-2.459	-65.718	-0.831	13.207
	(9.132)	(104.447)	(8.970)	(8.296)
Tons of bombs p.c. dropped in 1st 6 months	3.051***	0.319	-1.218	-2.286
	(1.009)	(8.451)	(5.885)	(2.775)
Father's education high	-0.027	2.311***	0.295***	-0.075***
	(0.029)	(0.370)	(0.051)	(0.023)
Place of birth: pop. >= 10000	-0.005	0.640**	0.143***	0.016
	(0.024)	(0.299)	(0.036)	(0.022)
R^2	0.013	0.068	0.087	0.019
F-Stat.	1.262	5.361	4.433	1.160
N	1981	1985	1111	1439

Note: Robust standard errors in parentheses. *,**,*** indicates significance at the 10%, 5%, and 1% level, respectively.

Another problem of omitted variable bias could emerge from the fact that season of birth has been found to be related to health and longevity (e.g. Doblhammer and Vaupel (2001)). Both the nutritional situation as well as exposure to bombings could exhibit seasonality. We examine this issue by adding quarter of birth dummies to our regressions. These are not found to be statistically significant predictors of our outcomes and all main results remain unchanged.

Regarding the operationalization of our key contextual variables, the 1,500 kcal threshold may be conceived as ad hoc. Obviously, we face a trade off between statistical precision and effect relevance: Setting the threshold to a very low level may guarantee us to filter out those individuals who suffered from severe nutritional shortages in either of the critical periods but on the other hand, our 'treatment group' may shrink considerably such that an effect is very imprecisely measured. Conversely, setting this threshold to a higher level may yield a more precise estimate at the expense of effect relevance since we now may declare individuals as 'treated' who may not have experienced as adverse early circumstances after all. We assess this matter by running our main regressions again but changing the definition of the nutritional shortage indicators around our original 1,500 calorie level from 1,300 onwards. The effect of prenatal famine on height among men is found to be negative and sizable for thresholds lower than 1,500 calories while, however, significance is lost for larger deviations. The effect of postnatal famine on height of women shows the same pattern.

Regarding the bombing measure, the adequacy of including this variable linearly in our regression models has to be justified. We experimented with a logarithmic transformation of

Table 3.16: Females – Results when including cohort fixed effects

Variable	Hypertension	Height	Diabetes
Less than 1500 kcal during gestation (1/0)	-0.017	0.204	0.012
	(0.066)	(0.798)	(0.056)
Less than 1500 kcal 1st 6 months (1/0)	0.195***	-1.887**	0.056
	(0.069)	(0.736)	(0.046)
Tons of bombs p.c. dropped during gestation	1.072	-4.385	0.120
	(3.432)	(28.729)	(1.117)
Tons of bombs p.c. dropped in 1st 6 months	3.416	45.644	-2.230
	(5.201)	(56.423)	(1.927)
Father's education high	-0.086***	1.234***	-0.032
	(0.029)	(0.343)	(0.021)
Place of birth: pop. >= 10000	0.037	-0.329	-0.005
	(0.023)	(0.269)	(0.019)
R^2	0.029	0.035	0.028
F-Stat.	2.433	3.142	1.478
N	2038	2039	1390

Note: Robust standard errors in parentheses. *,**,*** indicates significance at the 10%, 5%, and 1% level, respectively.

this variable which requires setting zero values to a very small value. It turns out that for our models explaining hypertension risk, a second order polynomial in the logarithmic bomb load provides the best fit. Both the first order and the second order term are estimated to enter the regression function with a positive coefficient. Moreover, the estimated coefficient values are such that on the support of the bombings variable, the predicted relationship between hypertension risk and the bombardment variable is mostly linear. Hence, we conclude that our original specification is supported by the data and the most efficient in terms of parametrization.

3.7 Conclusion

The present empirical study sheds light on the components of famine effects. Famines are aggregate disruptions and arguably provide exogenous variation in the nutritional situation of a fraction of a population under study. Hence, famines have been used to empirically test theoretical predictions of fetal programming of humans by nutritional shortages early in life that may manifest as worse health outcomes later in life. The commonly used study design is to measure a reduced form effect of famine exposure on later life health. However, it is possible that a suchlike measured effect picks up factors different from food shortages, most importantly psychological stress. Such a stress component could be either directly evoked by individual hunger or factors paralleling the famine. It is thus questionable whether famine studies are an adequate study design to measure the long run effects of malnutrition.

Using a unique combination of German survey data and information on both food rations as a measure of famine exposure within an early life age window and the exposure to Allied air raids in World War II, we gain new insights on this matter. As argued in our study, air raids may be assumed to solely evoke a stress reaction while exposure to famine may imply both individual malnutrition and stress. This means that if we find air raid exposure to have an effect on a given outcome, early life stress is likely to be a determinant of this dimension of health. If the same health outcome is not affected by exposure to famine, we may conclude that the stress component of early life famine is relatively small or non-existent or that malnutrition and stress under famine cancel each other out. On the other hand, if air raids do not have an effect on an outcome but famine does, malnutrition should be the driving force.

Our results show that researchers should be cautious with the conclusions drawn from famine studies and that effect components may have differential impacts depending on the outcome of interest. The main findings (all insensitive to a variety of robustness checks) imply, first, that the prevalence of hypertension in male respondents is higher when exposed to air raids in the first year after birth, but does not react to famine. Under the assumption that malnutrition does not have a beneficial effect on hypertension risk but most likely an adverse one (as found in experimental animal studies; see Tonkiss et al. (1998); Ojeda et al. (2008) and references therein), this means that we do not have evidence for the stress component in the first year exposure to famine to be quantitatively relevant. Second, we find that male adult height is significantly affected by famine exposure during gestation but not by air raids, i.e. not by stress. The same applies to the effect of postnatal exposure to famine on height of women. This implies that a famine effect on height is in fact attributable to malnutrition. Notice that our finding that air raids affect hypertension among men is indicative for air raids to be an informative proxy for stress which strengthens our conclusion drawn for the effect of stress on height: since bombings measure stress sufficiently precisely to turn up as a significant reduced form estimate in the case of hypertension, not finding a significant effect of air raids on height is more likely to be due to height being insensitive to stress rather than it is due to air raids not capturing stress in the population sufficiently well. Since we also find evidence for prenatal famine to have significant effects on pensions of male respondents (a proxy for life cycle productivity), our findings hint at the importance of nutritional shocks for correlates of height, namely dimensions of human capital and thus overall productivity. In general, the effects of stress under the German famine seem to be relatively small.

Our study faces some drawbacks such as measurement issues concerning the rations data. Moreover, factors inherent in the historical background such as selective mortality and selectivity of the place of birth during the war may pose threats to identification. Conducting a range of robustness checks shows that those issues are unlikely to drive our results. However, we have to note that our results, though internally valid, are context dependent and that other famines may differ in terms of the size of the effect components we try to disentangle. This should also explain differences between our findings and the effects found in other famine studies. Our work could thus be used as a point of departure for further research in this direction.

Chapter 4

Instrumental Variable Estimation of the Causal Effect of Hunger Early in Life on Health Later in Life[1]

4.1 Introduction

During the past decade, the effect of early-life conditions on health outcomes late in life has become a focal point of research in economics, demography, epidemiology, biology, and related fields.[2] Empirical studies use non-experimental data of real-life individuals and typically relate indicators of early-life conditions to outcomes of the same individuals at high ages. A major concern is that both (i) observed conditions in the parents' household early in life, including conditions in utero as captured by birthweight, and (ii) outcomes later in life, are jointly dependent on unobserved confounders. To be able to detect causal effects, one needs to observe exogenous variation in the early-life conditions, and relate this to outcomes later in life. Moreover, this variation should only affect high-age health by way of the individual early-life conditions and not through secular or cohort-specific changes in society after the early-life period of interest. As a result, popular indicators are often not some unique characteristic of the newborn individual or his family or household, but rather a temporary state of the macro environment into which the child is born. In that case they are also called contextual variables.

[1] This chapter is joint work with Gerard J. van den Berg and Pia Pinger. An earlier version of this chapter has been published online as IZA working paper number 6110 in 2011 (in 2012, a similar version was also published in working paper series at The Institute for Labour Market Policy Evaluation (IFAU), Uppsala; The Centre for European Economic Research (ZEW), Mannheim; and the Department of Economics, University of Mannheim). An updated version has been published in Pia Pinger's doctoral thesis at the University of Mannheim (Pinger (2013)). This version is a third update to the two previous versions.
Data Acknowledgement: This chapter uses data from SHARELIFE release 1 and SHARE release 2.5.0. The SHARE data collection has been primarily funded by the European Commission. See the notes to chapter 2 for more details and www.share-project.org for the full list of funding institutions.

[2] For overviews, see e.g. Pollitt et al. (2005), Barker (2007), Lawlor (2008) and Almond and Currie (2011b).

In the literature, the most frequently used indicator is whether the individual has been exposed to a famine at a specific childhood age, notably in utero. Lumey et al. (2011) provide an excellent overview. The underlying idea is that in many households birth in a famine involves nutritional shortages, in terms of quality and/or quantity of nutrition. Birth outside of a famine is expected to involve better nutritional conditions. A negative association between birth during a famine and a favorable health outcome later in life is interpreted as evidence for a causal effect of early-life conditions on that health outcome. Studies have detected associations with adult height, schizophrenia, mortality, fertility, hospitalization, next generation birthweight and even gene expression.[3] A few studies have also reported associations between exposure to a famine at the onset of puberty and late-life health outcomes.[4]

However, exposure to a famine is not equivalent to exposure to a nutritional shortage. During a famine, a fraction of all households does not face food shortages, for example because the household belongs to the ruling or wealthy class, or because it is self-sufficient in terms of food, or, in a war context, because it is allied to those responsible for the cause of the famine. Similarly, in non-famine eras, a fraction of households faces food shortages because of poverty. This means that the comparison of famine-born individuals to non-famine-born individuals does not provide a consistent estimate of the average causal effect of nutritional shortages around birth. Most likely, the latter effect is under-estimated in absolute value by such a comparison.

To advance on this, it is necessary to observe the nutritional status in the households at the time interval in which the child is in utero or at the childhood age of interest. However, this is often impossible. After all, in part of the sample, the critical period must have occurred during a famine. During famines, societies are in a state of disruption, so the relevant data are typically not collected.[5]

[3]Results from the famine-based studies are qualitatively confirmed by studies using alternative contextual indicators of conditions around birth, such as business cycles and seasons. See e.g. van den Berg and Lindeboom (2014) and the references in footnote 2 for overviews.

[4]For example, Sparén et al. (2004) find that boys who had been exposed around age 9 to the Leningrad siege famine of 1941 have higher rates of cardiovascular morbidity and mortality much later in life. Oppers (1963) finds an effect of the Dutch Hunger Winter famine exposure at ages 7-14 on adult height. Other studies have related the calcium intake around age 9 to adult height. Lindeboom et al. (2010) use immigrant siblings to study the effect of economic conditions at later childhood ages, and they find a significant effect on adult height. They also survey other studies concerning later childhood ages.

[5]If the study outcomes concern health at high ages then an additional problems arises. To observe such outcomes, the individuals in the data need to have been born a long time ago. This means that the critical period of interest necessarily occurred say before 1950. At the time, data on household conditions were not systematically collected.

In this chapter we deal with these problems by using self-reported retrospective data on the individual occurrence of a period of severe hunger at certain childhood ages. Specifically, we relate these data to the occurrence of famines, and we use Instrumental Variable (IV) estimation techniques to estimate average causal effects of nutritional shortages during certain childhood ages on adult height. Height is a universally accepted proxy of adult health and observable for individuals of all ages, while chronic diseases manifest only relatively late in the life cycle. In terms of the IV treatment evaluation literature, our instrumental variable is the exposure to a famine early in life, our treatment is the experience of a nutritional shortage early in life, and our outcome is adult height. If the effect of malnutrition is heterogeneous in the population, the estimated effect is an average for individuals who have been affected by a respective famine, which is famine-specific. We acknowledge this heterogeneity in the effect of famine-induced malnutrition on height by estimating local average treatment effects (LATE; see Imbens and Angrist (1994)).

The magnitude by which the causal effect exceeds the reduced-form effect hinges on the percentage of individuals alive who have been affected by malnutrition during childhood. Lumey et al. (2011) show that famine studies often provide varying reduced-form effect sizes.[6] This chapter contributes to this literature by showing that differences in famine exposure give rise to heterogeneity in reduced-form findings. The larger the fraction of famine-affected individuals and the higher the degree of nutritional deprivation, the higher the estimated reduced-form effect. Only at the extreme, if "compliance" to the instrument is perfect, does the reduced-form effect conform to the causal effect of nutritional shortage. More in general, to obtain this causal effect, the reduced-form effect must be scaled by the rate of compliance.

This approach is novel. We are the first to use the occurrence of famines to obtain an estimate of average *causal* effects of *nutritional shortages* during specific childhood ages on health later in life. The approach is not specific for famines as early-life indicators but also applies to other contextual indicators such as business cycles, seasons, weather, and infant mortality rates. As a by-product, our study provides estimates of the strength of the association between the famines we investigate and an actual hunger episode.

Our individual data are from the Survey of Health, Aging, and Retirement in Europe (SHARE), a European longitudinal survey based on a random sample of individuals aged 50+. The third wave asks respondents for retrospective accounts of specific aspects of their

[6] Reduced-form findings for male adult height range from zero for the Leningrad siege (Koupil et al. (2009)) to around 3 cm for the 1959-1961 Chinese famine (Chen and Zhou (2007)).

lives. The birth cohorts in the data include cohorts that were exposed to the famines in the Netherlands, Germany, or Greece, in various time intervals in the 1940s. These are the three famines in the countries participating in SHARE in its observation window.[7] Their origins are well-established in the literature, and they have been used to study long-run effects (see the literature discussion in Section 4.2). Notice that the combined evidence based on multiple famines is potentially less sensitive to culture or cohort-specific conditions.

Strictly speaking, it is the cause of the famine that is the ultimate instrumental variable, instead of the famine itself. After all, any period or era in which many people are hungry may be called a famine. What matters for the study of long run effects is that the famine is an exogenous event with no long-run impact apart from the effect running through household-specific nutritional conditions. This is more likely if the famine is due to an external intervention in society and if the famine is short and is not anticipated. The three famines in our data satisfy these requirements, as they are all due to trade blockades in combination with government rule by foreign occupying forces.

Nutritional shortages that only took place in utero will not be reported as episodes of hunger. Even with a perfect recollection of past periods of hunger (e.g. if the individual obtained this information from parents), a spell in utero will only be reported if it stretches past the day of birth. More in general, recall of a period of severe hunger may be difficult if this period took place around birth. Indeed, in our data, the reporting of hunger during a famine is low if the age during the famine was close to zero. We deal with this by using responses from individuals who were children aged above 5 but not yet adult during the famine in order to estimate the association between famine exposure and the experience of hunger. So when we consider long-run effects of nutritional shortages for newborns, we relate famines around birth to health later in life, and we use sample members who were older children during the famine to estimate the nutritional severity of the famine. In effect, we apply two-sample instrumental variable estimation methods, notably the two-sample nonparametric Wald estimator (2SIV) and two-sample two-stage least squares estimator (2S2SLS). This approach has been developed by Angrist and Krueger (1992)) and Arellano and Meghir (1992); see Ridder and Moffitt (2007) for an integrated overview. In our setting, the approach requires the assumption that the connection between famine and hunger is on average the same for children at different ages. This is not innocuous. As we shall see, there is anecdotal evidence of special food support for young children during famines that was

[7] Barring famines that affected only a few respondents, like the German 1916/18 famine; see van den Berg and Pinger (2014). Our analysis restricts attention to births in 1920-1955.

not available for older children. In that case, our estimates provide a lower bound for the average causal effect.

The SHARE data have established a high reputation in terms of quality, and by now many studies have been published using these data (see e.g. Börsch-Supan et al. (2008)). Our particular empirical analysis faces two data design limitations. First, for our purposes, the sample is not large. The number of respondents per country is around 1500, and to be exposed to a famine at a specific age, the respondent needs to have been a child (or in utero) in one of three countries in a birth cohort interval with a length equal to the famine, where the actual famine duration ranges from a few months to at most a few years. Secondly, the survey questions concerning the period of severe hunger ask for an interval in terms of full calendar years. In each analysis we therefore need to define and align three different intervals early in life: the relevant age interval in which nutritional shortage may cause long-run effects, the calendar time interval for which the individual reports severe hunger, and the calendar time interval in which the famine took place. Inevitably, we have to make several shortcuts, and it is important to address the sensitivity of the results with respect to this. It turns out that moderate changes in the definitions of these intervals only affect the status of relatively few respondents, and the results are usually insensitive to this.

Since we aim to estimate a causal effect of nutritional shortages in general, the relevance of our findings should stretch beyond famine-stricken societies. From an economic point of view, it matters to find out at which age of the children exposure to nutritional shortages is most detrimental in the long run. This helps to address which policy measures are most efficient and cost-effective in preventing adult health problems.

The chapter proceeds as follows. In Section 4.2 we review the explanatory frameworks to understand the long-run effects. Section 4.3 describes the three famines in our observation window and summarizes the evidence obtained so far for those famines. In Section 4.4 we describe our data. Section 4.5 formally presents the econometric methods. Here we also examine selectivity issues associated with the famines. Section 4.6 presents our results. We do not only study effects of adverse conditions around birth but we also identify whether periods in early adolescence are critical with respect to nutritional shocks. We also consider the strength of the association between famines and actual nutritional shortages. This is of importance for the empirical literature in which the three famines have been used as indicators of early-life conditions. In Section 4.6 we also carry out placebo estimations using cohorts from countries that were not affected by famines, as well as a variety of other robustness checks. Section 4.7 concludes.

4.2 Explanatory frameworks for causal long-run effects of conditions early in life

4.2.1 Conditions around birth

Most explanations for long-run effects of nutritional conditions around birth build on Barker's fetal origins or fetal programming hypothesis (see e.g. Barker (1994)). Effects of fetal undernutrition on metabolic adaptation in utero may affect the phenotype such that the risk of cardiovascular disease later in life is increased (Hales and Barker (1992), Bateson (2001), Gluckman and Hanson (2004)). Underlying this model is the idea that several critical periods early in life influence the development of humans. During these periods, developing systems modify their settings in response to social and biological cues (Kuzawa and Quinn (2009)). This includes durable epigenetic changes that modify gene expressions.

Along this way, adverse conditions are known to influence inflammation, measured in terms of interleukin-6 production, in adolescence and adulthood, plausibly through changes in gene expressions (see e.g. Morozink et al. (2010)). An episode of hunger early in life may thus engender a proinflammatory phenotype. Over time, this takes an allostatic toll on the body, resulting in a higher risk of chronic diseases later in life (Morozink et al. (2010), Miller and Chen (2010), Miller et al. (2009), Zhang et al. (2006), Cole et al. (2011)), notably cardiovascular diseases, diabetes and hypertension. Such mechanisms should be seen as a predictive adaptive response to the future environment (Gluckman and Hanson (2004), Cole et al. (2011)). Indeed, the long-run effects of reduced nutrition in utero are stronger if the affected individuals are exposed to a much more favorable environment in childhood (Schulz (2010)). Of course, severely adverse nutritional conditions may also directly affect the build-up of organs and other body parts, in utero or in the postneonatal period (see e.g. the survey by Cameron and Demerath (2002)).

The above causal pathways are all biological after the initial nutritional shortage. However, non-biological mechanisms are also possible. An episode of hunger may lead to a permanent change in the role of food in the household. After such an episode, adults may cook more greasy food or force their children to finish their plates by any means (see e.g. Hamelin et al. (1999), for changes in household behaviors in response to food insecurity). This may cause adverse long-run health outcomes of the children. Also, an episode of hunger may induce fights for resources and thus a higher level of stress, changing within-household relationships and leading to a prolonged exposure to a higher stress level which may affect health at high ages (Hadley and Patil (2006), Whitaker et al. (2006)).

4.2.2 Conditions in later childhood

Recently, interest has increased in long-run effects of nutritional conditions beyond infancy. Gluckman et al. (2005a) and Barker (2007) give overviews of the underlying medical mechanisms. Particular attention has been given to the onset of puberty as a sensitive period. As documented and surveyed by Marshall and Tanner (1986), Gasser et al. (1994), and Zemel (2002), the earliest manifestation of puberty concerns the so-called "fat spurt" around age 9-10. In this spurt, the body collects resources in anticipation of the adolescent growth spurt. Sparén et al. (2004) argue that nutritional distortions and stress at this stage may lead to a permanent disruption of blood pressure regulation, leading to long-run cardiovascular health problems.

A nutritional shortage at certain childhood ages may also have an instantaneous adverse effect on schooling decisions and outcomes (see e.g. Jyoti et al. (2005)) and thus affect health outcomes via realized education or adult socioeconomic status (Leigh (1983)). Lastly, adverse conditions may affect children positively by inducing more responsible behavior. For example, Elder (1999) investigates the impact of the Great Depression on children born in 1920-1921 and finds that the experience of economic hardship around the age of 10 led to more resilience and psychological strength.

4.2.3 Gender differences

The literature on the effects of early-life conditions on cardiovascular morbidity and mortality has found stronger effects for males than for females (see the surveys in Poulter et al. (1999), Rasmussen (2001), Lawlor et al. (2004), Huxley et al. (2007), van den Berg et al. (2009a)). For over-all mortality this is echoed by the studies of the Finnish famine from 1866-1868 (Doblhammer et al. (2013b)) and the Dutch Potato Famine from 1846-1847 (Lindeboom et al. (2010)). Men with certain health problems that might originate earlier in life are more likely to die than women with the same conditions (Case and Paxson (2005)).

Focusing on adult height as an outcome, we should first point out that studies on height often do not consider females. This is because these studies use military enlistment register data that were collected among young men only. Studies using different data sources tend to find that male stature is more responsive to changes in early-life conditions than female stature (Kuh et al. (1991), Eveleth and Tanner (1990). Indeed, historically, gradients in growth and variability in stature over time are lower for females than for males even if expressed in percentage terms (Kuh et al. (1991)).

In sum, it seems that boys tend to be biologically more sensitive to early-life nutritional conditions than girls (Kraemer (2000), Low (2001)). This may be explained by the so-called "male vulnerability" that arises because males are the heterogametic sex (see. e.g. Low (2001)). Males have an unprotected Y chromosome and therefore they may be more vulnerable to adverse conditions giving rise to long run effects. Another factor is that boys are born with less mature respiratory systems than girls (Waldron (1983)). More in general, Eriksson et al. (2010) find that intra-uterine growth strategies differ between men and women. For a given low birthweight, male placentas are smaller than those of their female counterparts. Smaller placentas often result in compensatory expansions in boys, which, while benefiting brain development, is at the cost of other body parts. The authors suggest that in its extreme form this growth strategy of boys in response to adverse uterine conditions may explain the higher prevalence of adverse adult health outcomes (such as a higher mortality rate) among men compared to women.

Concerning adult height, an additional explanation for why female adult height is comparatively less responsive to malnourishment is that hunger leads to a delay in menarcheal age. This delay lengthens the period of linear growth (Koupil et al. (2009)).

4.3 The famines in European countries in the 1940s

4.3.1 The Dutch famine

The Dutch famine has been studied for decades as a cause of adverse living conditions. Therefore, the following account can be brief. The relevant literature starts with contemporaneous studies, notably Dols and van Arcken (1946), who provide a detailed description of the famine and report data on rations and agricultural production, and Banning (1946), who focuses on public health issues.

Prior to World War II, food standards had been high in the Netherlands, both in terms of caloric value as well as composition of the diet. There were no notable disruptions in food availability during the first years of the occupation of the Netherlands, which started in May, 1940. In September 1944, parts of the South of the country were liberated, and the London-based Dutch Government in exile called out a railroad strike in the occupied parts of the Netherlands as an attempt to display authority over the inhabitants of the occupied parts. In response, the occupying forces moved away the means of transportation from the occupied parts, effectively initiating an embargo on food transports to the densely populated western part of the country, i.e. the provinces of North and South Holland and

Utrecht. In combination with the early onset of the harsh winter of 1944/45, the freezing of waterways, and the generally bad state of transport infrastructure, this effectively closed off the western part of the country from any imports of food, fuel, medication etc. This triggered the Dutch "hunger winter'. Individuals had to live on rations as low as 500 kcal per day. For school children, average rations amounted to 664 kcal in the first quarter of 1945. The situation lasted until the end of the occupation which coincided with the end of World War II (early May 1945). Immediately, rations rose to 2,400 kcal per day. Following most of the literature on the Dutch famine, we take November 1944 to be the onset of the famine spell. This is later than the onset of the strike in mid-September 1944.[8]

The excess death rate in the first half of 1945 over the rate in 1944 amounts to 269 percent for men and 173 percent for women (Dols and van Arcken (1946)). Banning (1946) reports a higher incidence of tuberculosis and hunger oedema and an increased infant mortality rate. Banning (1946) notes that in small towns mortality rates rose to a level almost as high as those in large cities. Special aid was targeted at starving children and young adults by the "Inter-Church Council", an organization formed of different clerical associations (first, the focus was on children aged 5 to 16; the inclusion of children aged 3 to 5 followed later on). Help was provided in the form of food. In addition, a fraction of the adversely affected children was sent to districts where the food situation was somewhat better. Banning (1946) mentions that of the potential candidate children examined, 29% had been severely undernourished, while 31% suffered moderate undernourishment. About 27% of the children displayed a weight loss of about 10% of their pre-famine weight.

Studies based on the Dutch famine indicate significant long-run effects on adult morbidity.[9] The overview in Painter et al. (2005) lists long-run effects on the risk of cardiovascular diseases, obesity, breast cancer, cholesterol levels, diabetes, and self-perceived health. See also the survey in Lumey et al. (2011) for effects of prenatal famine in particular. Lumey et al. (2007) find effects on anthropometric measures indicative of the reposition of fat, dermatoglyphic characteristics and a modest relationship with blood pressure. An elevated risk of schizophrenia at adult ages after prenatal exposure has been found for both genders (e.g. Susser and Lin (1992)). Using inpatient population register data, Scholte et al. (2012) conclude that hospitalization rates at ages shortly before retirement are higher after middle or late gestational exposure to the famine. Susser and Stein (1994) find that adult stature

[8]For school children, official rations dropped below 1,200 kcal in early November 1944 and to 1,000 kcal at the end of November 1944 (Dols and van Arcken (1946)). On average, caloric consumption amounted to 1,073 kcal in the last quarter of 1944 (Dols and van Arcken (1946)).

[9]Recall that in this section we only cite studies based on the famines we consider in our empirical analysis. Results based on other famines are discussed in the overview studies listed in Section 4.1.

is susceptible to the postnatal but not the prenatal environment. As noted in Section 4.1, Oppers (1963) finds a negative effect of the Dutch Hunger Winter famine exposure at ages 7-14 on adult height among men. In a landmark study, Heijmans et al. (2008) show that individuals who were exposed to the famine in the early stages of pregnancy had, 60 years later, less DNA methylation of a certain imprinted gene, compared with their unexposed same-sex siblings. The gene is the insulin-like growth factor II, which is a key factor in human growth and development. Their study provides strong empirical evidence for the epigenetic pathway discussed in Subsection 4.2.1 above.

4.3.2 The Greek famine

At the end of April 1941, Greece surrendered to Axis forces and was subsequently divided into zones occupied by Germany, Italy, and Bulgaria.[10] These were isolated from each other and the transfer of goods and individuals was often close to impossible. The famine was triggered by a naval blockade of the Allies which made it impossible to supply foodstuffs to Greece. Despite being an agricultural country, Greece heavily relied on food imports (Hionidou (2006), mentions that in 1939, over 20 percent of the wheat consumption had been imported). In addition, the blockade prohibited fishing at sea.[11] The food situation quickly deteriorated. While the caloric value of the rations allocated by the occupiers in the Athens area had already been below any subsistence level in July 1941 (600 kcal), the rations were cut even further such that they reached 320 kcal in November 1941 (Neelsen and Stratmann (2011)). The regions of Greater Athens and Piraeus were affected most severely. Mortality rates suggest that there was a general positive correlation between the famine's severity and the degree of urbanization. Furthermore, Hionidou (2006) notes that society was very unequally affected by the famine during this period; the low social classes suffered the most. Soon after the rationing system had been superimposed by the occupiers, black market activity flourished with prices beyond the levels an ordinary worker could afford. The winter 1941/42 marked the peak in mortality during the famine period.

The blockade was formally called off in February 1942. As of then, the Red Cross provided assistance targeted mostly at young children in Athens (Neelsen and Stratmann (2011)). At the end of March 1942, shipments by the "Joint Relief Commission" under Swedish command arrived. Subsequently, the situation in most parts of Greece improved.

[10] This subsection relies heavily on Hionidou (2006) and Valaoras (1946).

[11] In the sense that the famine was caused by an external intervention in society, and in the sense that a large fraction of the excess mortality during the famines was due to starvation, the famines discussed in this section are "modern famines" in the terminology of O'Gráda (2009).

Hionidou (2006) reports declines in excess mortality from April 1942 onwards, even though in urban areas mortality remained high longer. We define the Greek famine to run from May 1941 until and including June 1942. The ending date is somewhat later than in Hionidou (2006) and somewhat earlier than in Neelsen and Stratmann (2011).

Valaoras (1946) reports death rates for Athens to have been six times higher than under usual circumstances. Both Valaoras (1946) and Hionidou (2006) state that the largest fraction of the excess mortality during the famine is due to starvation, whereas epidemics and infectious diseases are of minor importance. Of special interest for our present study are the figures cited in Valaoras (1946) on the situation of children and adolescents. The results of surveys conducted during the famine and shortly afterwards found children aged 4 to 14 to be massively underweight, while babies and young children up to age 4 had almost normal weight. Moreover, growth retardation was found among many children 8 to 14 years old.

Neelsen and Stratmann (2011) consider long-run effects of the Greek famine on economic outcomes. They find effects of exposure during infancy on the attained level of education and on prime-aged labor market outcomes. Effects of exposure in utero are very small.

4.3.3 The German famine

After World War II, Germany faced various structural problems that led to the catastrophic situation we label the German Famine. Specifically, (i) many cities and much of the industry and transportation infrastructure were destroyed by Allied bombardments, (ii) millions of citizens and military had died, (iii) factories were dismantled and shipped off to the occupying powers, (iv) the agricultural land in the East which had been a major provider of food had been lost to Poland and the Soviet Union, (v) about 10 million refugees from the lost lands in the East had fled to the new German mainland, and (vi) Germany was divided into 4 occupation zones that were administered separately. Inter-zone trade was difficult because of political and bureaucratic barriers and because of a lack of transportation means. See Klatt (1950), Farquharson (1985), Trittel (1990), Häusser and Maugg (2009) and Reichardt and Zierenberg (2010) for overviews.

Before World War II, estimated daily caloric consumption amounted to about 3000 kcal, while during the war, civilians used foodstuffs of about 2500 kcal per day (Klatt (1950)). At the end of the war (European Theater) in May 1945, nutritional conditions deteriorated to levels far below those before or during the war. For May 1945, Trittel (1990)

reports an average caloric ration of only 1200 kcal. After May 1945, these values decreased further. For example, in the British-administered Ruhrgebiet, only approximately 600 kcal per person was assigned in June 1945. The situation further deteriorated during the so-called "Hungerwinter" of 1946-47, which was an unusually cold winter. From mid-December 1946, the distribution of food collapsed in many areas. Until March 1947, real average daily caloric intake was around 800 kcal per day for some cities in the Ruhrgebiet. The crop of 1947 failed to meet the demand of the starving population, so the period of hunger lasted until the spring of 1948, when a combination of foreign help and political reforms managed to improve nutritional conditions. Conditions improved spectacularly with the currency reform in June 1948. We therefore define the famine to last from June 1945 until June 1948.

As usual, the severity of the famine exhibits regional variation on top of the above-mentioned temporal variation (see Willenborg (1979), and the references above). Notably, the different occupation zones were affected differentially, mostly due to differences in the policies of the occupying powers. For example, the average daily food intake of a standard consumer in 1946 is estimated to equal 1610, 1430, 1535, and 1340 kcal in the US, the Soviet, the British, and the French zone, respectively, while the corresponding official food rations equaled 1330, 1083, 1050, and 900 kcal, respectively (see Cornides (1948) and Echternkamp (2003)).

Apart from excess deaths due to starvation, sources document that the famine exerted a range of additional adverse effects. For instance, Klatt (1950) reports cases of retarded growth of children in the Ruhrgebiet and a weight deficit among children of about 20 percent of the normal weight. The deficit was most pronounced for children above age 10. In a study of undernutrition in the city of Wuppertal, Dean (1951) analyzes birth registers of a local hospital and reports reduced birth weights for the years of the famine with the greatest decline for the year of 1945.

Jürges (2013) analyzes the effect of birth during the German famine on educational attainment and occupation in the labor market in 1970. He finds strong evidence that these outcomes are particularly adverse for those born in the winter of 1945/46. This suggests that an indirect causal pathway from early-life conditions through education and occupation to health later in life may be possible. Some epidemiological studies consider cohort-specific adult health outcomes without focusing a priori on famine cohorts. Notably, Hermanussen et al. (2001) study obesity among male conscripts; they observe that the 1946 cohort stands out in terms of low obesity at age 19. To a small extent, this also applies to the cohorts born during World War II. It is possible that this is because these cohorts faced childhood

conditions that were less out of tune with conditions in utero than the cohorts later in the German famine (recall the discussion in Subsection 4.2.1). In this sense, individuals born in World War II may face less adverse long-run effects than those born in the 1946/47 "Hungerwinter".

4.4 Data

For the empirical analysis we use the Survey of Health, Aging, and Retirement in Europe (SHARE), a pan-European multidisciplinary and longitudinal data set on individuals aged 50 and above. By 2011, three waves of SHARE were available. The first two waves provide information on health, socioeconomic status, family characteristics and well-being. The third wave (called SHARELIFE), collected in 2008 and 2009, provides detailed life-cycle histories of individuals for specific thematic fields, such as employment or accommodation.[12]

Sample construction. We use information on 3533 men and 3823 women born 1920 to 1955, who lived in Germany, the Netherlands or Greece during their youth and who participated in the SHARELIFE survey. This excludes foreign citizens as well as individuals who, due to physical and cognitive limitations, were unable to complete the interview themselves.[13] We also exclude individuals with missing information on the hunger variable, covariates, or the outcome measure adult height. Since height has been collected in waves I and II, individuals who have not participated in either of these surveys are dropped as well.[14] Finally, we trim the upper and lower 2.5 percent tails of each country's gender specific height distribution. Our working sample consists of about 2500 men and 2800 women.

Hunger and famine. One exceptional feature of SHARE is that respondents were asked whether they had ever experienced periods of hunger in their lives and if so, in which year of their life this period began and in which year it ended. The survey question reads "Looking back at your life, was there a distinct period during which you suffered from

[12]For general information on SHARE and SHARELIFE, see Börsch-Supan et al. (2008) and Schröder (2011).

[13]We drop complete proxy interviews but keep assisted interview information. This only concerns the interviews yielding the outcome measures and the undernutrition indicator.

[14]Conversely, there are individuals for whom we observe height but not the hunger variable since they participated in the first two waves of the survey but not in SHARELIFE. Sample attrition is high in SHARE, but there does not seem to be a systematic gender or age attrition bias across SHARE countries (Blom and Schröder (2011)). In most analyses we treat such observations as missing at random, keeping in mind that if attrition were somehow related to health, this would make our estimates more conservative. In addition, we conduct robustness checks with control variables for life expectancy at birth as a measure of general health status.

hunger?". Although this question alludes to an unhappy period of life, nonresponse is very small and amounts to less than 0.5%. However, the information may be prone to recall bias and measurement error, since it is self-reported and dates back many years.[15] Moreover, recall may be systematically biased if a period of severe hunger took place around birth. At the extreme, nutritional shortages that *only* took place in utero can not be reported as episodes of hunger.

In the public health literature, Sidel (1997) and Rose and Oliveira (1997) study the reliability of retrospectively self-reported hunger as a measure of food insufficiency. They conclude that the measure is reliable except for those who experienced food shortages below the age of 5. We now examine the self-reported hunger in our own data as a function of age, and we relate this to famine exposure. Recall that the famines defined in Section 4.3 are:[16]

1. Greece: May 1941 to end of June 1942
2. The Netherlands: November 1944 to end of April 1945, only the Western part[17]
3. Germany: June 1945 to end of June 1948

Figure 4.1 displays the fraction of individuals in the raw data who report hunger during a famine, as a function of age. More specifically, this is the fraction of individuals with an age t during a famine who report a hunger episode that includes this age t, as a function of t. This exploits information on individual's month of birth. As expected, individuals who were aged below 5 in the famine often do not report to have experienced a hunger episode. Individuals aged between 6 and 16 in a famine report to have suffered a hunger episode at a stable rate of just under 20%. Put differently, if an individual lived during a famine, his/her probability to report hunger increases until about age 7, and it remains stable for increasing ages at famine. Interestingly, this also applies to ages beyond 16 (not shown in the figure; note that the number of famine-exposed respondents quickly declines beyond 16). The gender-specific figures corresponding to Figure 4.1 (available upon request) are for all means and purposes identical.

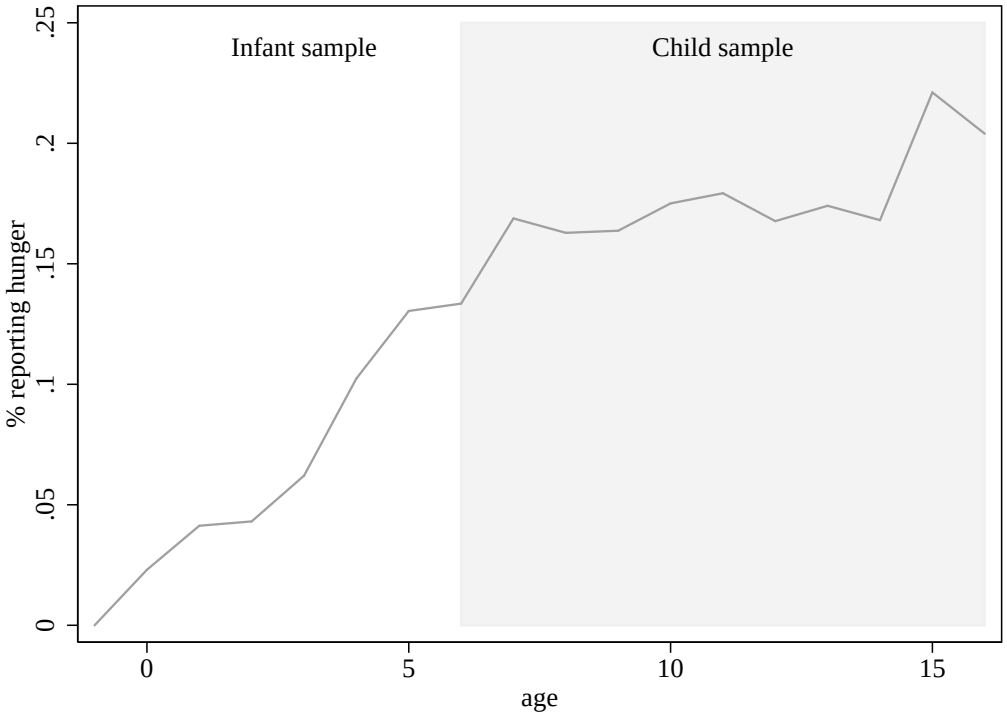
To proceed, we structure our data in two separate ways, depending on the ages at which the hunger episode takes place. With some abuse of language, we call the resulting

[15] Also, only one interval period of hunger can be reported.

[16] Changing the start or end months of the famine only marginally affects the empirical results. Notice that only individuals born at the margin of a specific year and month will switch from being declared as potentially famine exposed to not being exposed and vice versa.

[17] To select the Dutch sample into those exposed to a famine or not, we use information on the accommodation a respondent lived in the year the famine started, i.e. 1944; if a respondent moved house in the very same year, we use the previous accommodation.

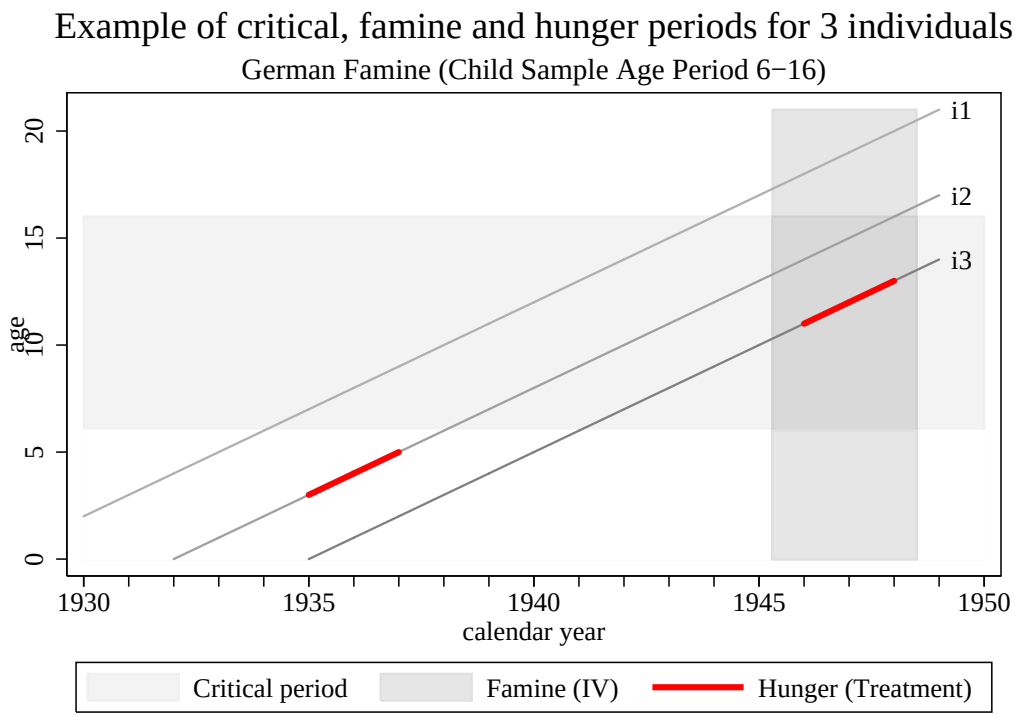
Figure 4.1: Probability to report hunger conditional on famine experience at respective age (males and females)



data structures the *infant sample* and the *child sample*. The infant sample is composed of individuals who potentially experienced hunger or undernutrition in utero or at ages 0-4. They may have been too young to have formed a recollection of the hunger, so if family members or others have not informed them of any actual hunger spell in the earliest years of life, they may not recall it. The individuals in the so-called child sample potentially experienced hunger at ages 6-16.

In the infant sample, we summarize the hunger episode into a binary undernutrition indicator (or "treatment" variable) D which is defined to equal one if the person experienced an episode of hunger in the age interval $[-0.75, 4)$. In the child sample we analogously define an indicator for hunger in the age interval $[6, 16)$. More specifically, a respondent is considered to have suffered undernutrition in a specific age range if (s)he reported that (s)he experienced an episode of severe hunger and if either this episode started before the beginning of the age range and ended after the beginning of the age range, or this episode started within the age range. Notice that we use broad age intervals to characterize the undernutrition indicator, and that we only require *some* overlap between the age interval and the reported hunger episode. This is because of the rather small sample sizes

Figure 4.2: Example treatment definition



of respondents exposed to famines. In extensive sensitivity analyses below we assess the robustness of the results with respect to these operationalizations. In the sequel, for sake of simplicity, we use the terms undernutrition and hunger interchangeably if there is no ambiguity. In addition to the binary "treatment" variables, we also formally define binary instrumental variables Z that take the value one if a famine affected the individual in the age interval $[-0.75, 4)$ or in $[6, 16)$, respectively.[18]

Table 4.4 reports summary statistics of the undernutrition indicators in connection to famine exposure. We find that 15.5% of individuals report to have suffered from hunger during a famine if they lived during a famine in the age interval 6-16.[19] Moreover, in line with the historical evidence described in Section 4.3, this fraction of hunger-exposed

[18] Figure 4.2 provides a graphical illustration of our hunger (treatment) and instrumental variable definitions for the child sample of German individuals. In this example, individual i1 did not report an episode of hunger, whereas i2 reported such a period at ages 3 to 5, that is, it has been affected by undernutrition prior to the relevant age window; the individual, however, did experience a famine in this window. For individual i3, both the treatment status and the famine instrumental variable are set to one since both the period of hunger as well as the famine occurred within the window when i3 was 6 to 16. Notice that the definitions of D and Z do not rule out that $D = Z = 1$ while the hunger episode and the famine do not overlap. This occurs for 3 individuals in the infant sample and for 11 individuals in the child sample. In such cases we set $Z = 0$. If, instead, we set $D = 0$ or if we remove these individuals from the sample then the results are virtually identical to those reported.

[19] If fertility and mortality are selective, as discussed in Subsection 4.5.3, then this number may differ from the percentage of affected individuals at the time of the famine, but is an accurate representation of the percentage among generations alive today.

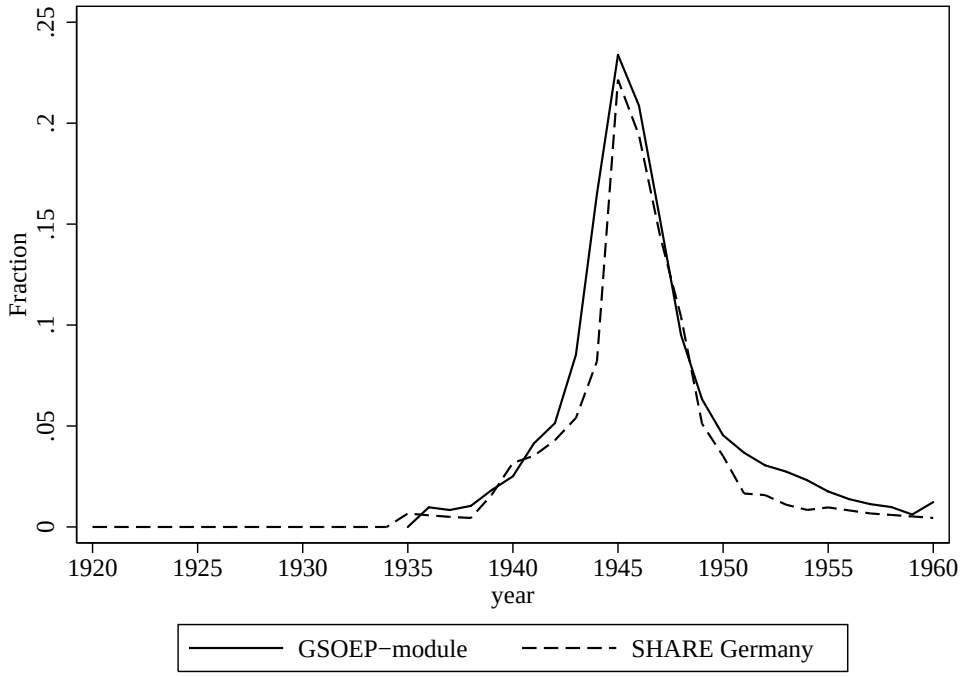
individuals is higher for individuals who lived in urban as compared to rural areas (20% vs. 14%) and in the Western part of the Netherlands as opposed to the Eastern part (25% vs 6%). Importantly, these hunger fractions are in accordance with evidence from van Abeelen et al. (2012) who consider a large sample of adult women who were exposed at ages 0-21 to the famine in the Utrecht area. This area was strongly affected by nutritional shortages during the famine. They find that approximately 16% of the women report severe famine exposure, while 38% report moderate exposure. These numbers are also in line with those found in a large number of developing countries in South Asia and Africa. Alderman (2012) shows that the difference in undernutrition, as measured by stunting, between the highest and lowest income quintiles varies between 12-35 percent.

For Germany, we perform an external validation of our findings concerning the association between hunger and famine, using a completely different dataset. In 2013, the German Socio-Economic Panel (GSOEP) implemented a one-off "early childhood module" that focuses on early-life circumstances. This module was sent by regular mail to individuals born in Germany in the years 1935-1950 and currently living in Germany. The module includes questions that essentially equal the "hunger" questions in the SHARE data that we use in our analysis. In addition, the module asks for the month of birth. The resulting sample size is about 3000.[20] With this we can replicate the analysis of the association between hunger and famine. Specifically, we can compare the fraction of individuals in Germany who report having suffered from hunger in each year in the SHARE data to the corresponding fraction in the GSOEP module, and we can compare the reporting of hunger by age at famine exposure. This is useful because it enables an assessment of the robustness of the findings for Germany with a larger sample collected at a more recent point in time. If recall errors are a major problem then we expect the GSOEP data to be less informative about the effect of the famine on hunger than the SHARE data are.

Figure 4.3 displays the fractions of individuals in Germany who claim to have suffered from hunger during a given calendar year, as a function of calendar time, for each data set separately. Clearly, these fractions show a striking similarity across both data sets. We also find that Figure 4.1 in our study is essentially replicated in the GSOEP module. In the latter data, the fraction reporting hunger at a certain age increases monotonically at ages zero to five, and it remains at exactly the level of age 6 for all higher ages. Thus, the rate of undernutrition during famines agrees to other sources of information, and the evidence confirms that the responses of those exposed at ages 6-16 are informative.

[20] The module will become publicly available in December 2014.

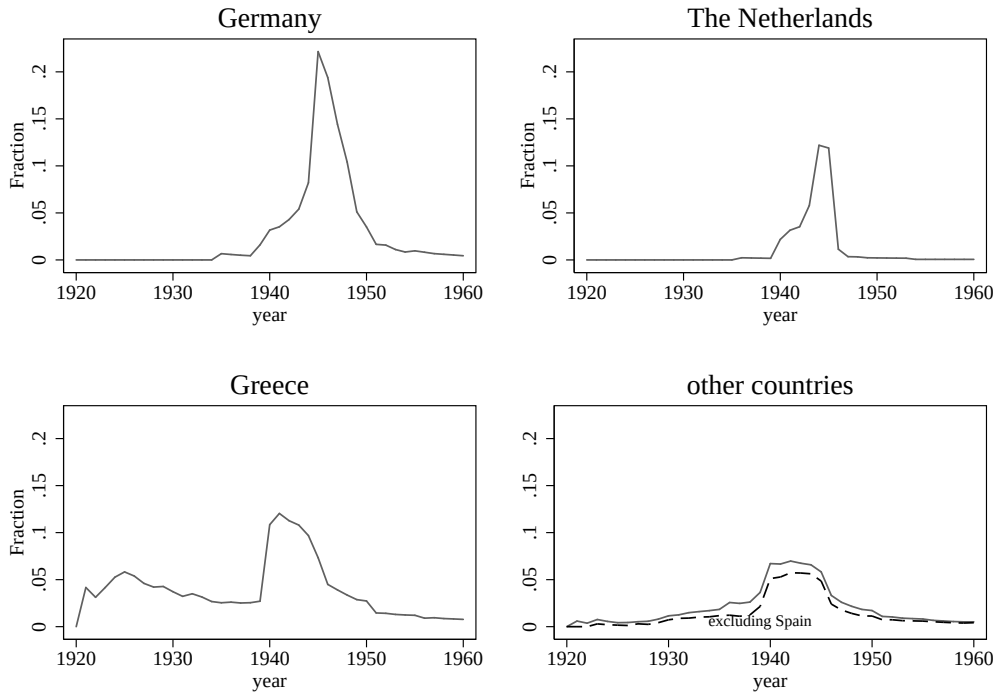
Figure 4.3: Comparison of episodes of hunger in SHARE-Germany and GSOEP-module data



Notice that the GSOEP module was collected about 5 years later than the SHARELIFE wave, and that the GSOEP module does not contain individuals aged 14 or 15 during the German famine (i.e., the composition was slightly tilted towards respondents who were young during the famine). In the light of this, the similarity of the results provides evidence that recall errors are not driving the results.

We now examine the dispersion of hunger episodes, for all countries. Figures 4.4 relates the hunger episodes in the data to the corresponding calendar years, by displaying the fraction of observations reporting to have had hunger in a given year among all individuals alive at that time. For Germany, the graph shows low propensities of hunger for the pre-war years, an increase during the war, and a drastic peak right after the end of the war. The fraction of those reporting hunger stays high until the end of the famine. For the Netherlands, we distinguish one single peak for the famine period while propensities for hunger are basically zero before and afterwards. The fraction of observations across the Netherlands reporting hunger during the famine years is lower than in other countries because only the Western part of the Netherlands was affected by the famine. For Greece we also observe a spike at the famine. The level is also relatively high for the early 1920s, but only a tiny fraction of individuals is born in these early cohorts. The impact of the famines is even more obvious when including the other SHARE countries: there, the fraction reporting

Figure 4.4: Probability for episode of hunger by calendar year (given observations of sample already alive)



episodes of hunger is small. The subsequent exclusion of Spain, which suffered from civil war in the 1930s, reveals that it accounts for a relatively large fraction of those reporting hunger in the pre-war years. The figure confirms that the experience of hunger in the famine countries was different from that in the other countries. In the famine countries, hunger was a severe, transitory shock, while in the other countries hunger can be ascribed to general suffering during the war.[21] This provides a rationale to exclude the other countries from the baseline analyses. We return to this in Section 4.6 when we discuss sensitivity analyses.

Figure 4.5 shows the average duration of a hunger period by year of onset. The graph shows that hunger periods are shorter if they overlap with a famine. Table 4.1 shows that a fraction of approximately 15 percent of those having experienced a famine before age 16, report hunger for the same period. For non-famine periods, this is only 3.5 percent. This confirms that famines are an instrument for periods of undernutrition.

[21]The three countries with famines were also exposed to turbulent episodes around the famine spells. For example, many German citizens were exposed to bombardments and stress during World War II, i.e. just before the German famine (see Akbulut-Yuksel (2009); recall also the literature discussions in previous sections), whereas Greek citizens were exposed to the Greek Civil War shortly after the Greek famine. The non-famine periods should therefore not be seen as tranquil eras of affluence.

Figure 4.5: Average length of hunger periods by calendar year

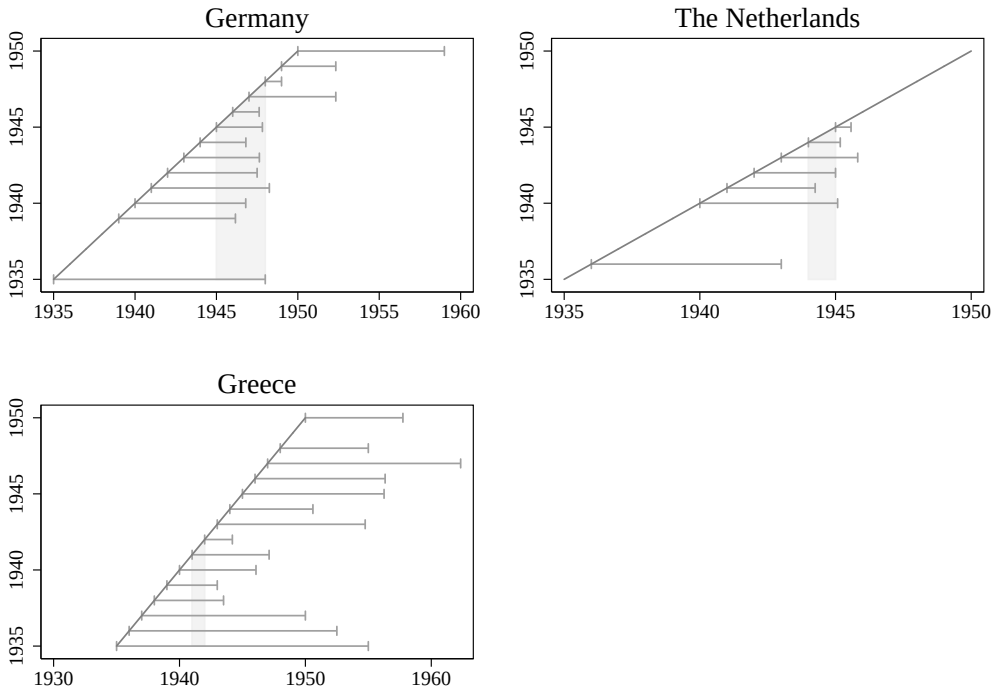


Table 4.1: Reaction of hunger at age 0 – 16 to famine exposure at the same time (cohorts 1920 – 1955)

	Greece		Germany		The Netherlands		all	
	treated	non-treated	treated	non-treated	treated	non-treated	treated	non-treated
Famine-exposed	110	810	119	634	55	172	284	1616
Not famine-exposed	40	1357	41	548	41	1443	122	3348

Table 4.2: Summary statistics for famine countries (cohorts 1920 to 1955)

	all (N= 5370)		control (N=4964)		hunger before age 16 (N=406)	
Variable	mean	(s.d.)	mean	(s.d.)	mean	(s.d.)
Year of birth	1942	(8.70)	1943	(8.70)	1936	(5.80)
Episode of Hunger being 0-16 (1 = yes)	0.076	(0.26)	0.000	(0.00)	1.000	(0.00)
Experienced famine being 0-16 (1 = yes)	0.354	(0.48)	0.326	(0.47)	0.700	(0.46)
Born in rural area (1 = yes)	0.412	(0.49)	0.419	(0.49)	0.335	(0.47)
Female	0.532	(0.50)	0.536	(0.50)	0.490	(0.50)
Male Adult Height in cm	175.756	(6.31)	175.852	(6.30)	174.681	(6.30)
Female Adult Height in cm	164.260	(5.41)	164.316	(5.38)	163.513	(5.78)

Adult height. As an outcome and proxy for late-life health we use adult height, measured in centimeters. Adult height is frequently used in the literature as a marker of early life conditions (Costa (1993)). It is viewed as the best single observable indicator of an individual's dietary history during childhood to the extent that it is of importance for health later in life (Elo and Preston (1992)). Moreover, for men, height is almost linearly related to life expectancy (Steckel (2008)). That adult stature is an uncontroversial indicator of the importance of early-life nutrition is important because the famines we investigate all took place during a period of psychological distress. Many late-life health outcomes such as hypertension, obesity or mental disorders may be affected by child hunger, but arguably they are more sensitive to the psychological burden of the famine. Height has the additional advantage of being less subject to age or period effects. Stature is mostly constant during adult life, while hypertension, heart diseases and obesity only manifest at a certain age.[22] Moreover, observation of the latter outcomes may depend on the propensity to visit to health care workers and may be sensitive to recall errors. Height is measured in the second wave of the survey. Table 4.2 compares relative magnitudes of adult height between those who experienced hunger before age 16 and those who did not.

Table 4.3 displays the results of a descriptive regression of adult height on the prevalence of hunger. Since this regression is not motivated by an instrumental variable setting, we perform it for the full data set and for all ages below 16. Moreover, we have seen that hunger below age 6 is underreported, so that a separate analysis for those below age 6 would not be informative. We obtain a significantly negative association between adult height and hunger before age 16. As argued in the introduction, this "plain vanilla OLS" regression is potentially misspecified because child hunger may be endogenous and

[22] Hypertension manifests at age 45-65 (Riseman and Weiss (1929)). Obesity risk increases linearly until age 60 and falls thereafter (Sanz-de Galdeano (2005)). In a preliminary working paper version of our paper we consider other outcome variables in addition to height (see van den Berg et al. (2011)).

Table 4.3: OLS regression of height on hunger

variables	hunger	
	males	females
Hunger during childhood	-0.653**	-0.499*
	(0.02)	(0.06)
Born in a rural area	-0.784***	-0.304**
	(0.00)	(0.01)
Observations	8177	10195
R^2	0.241	0.163
Cohort FEs	YES	YES
Country FEs	YES	YES
CohortxCountry FEs	YES	YES

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Note: Robust standard errors reported.

likely correlates with unobserved child attributes that also affect stature. (Table 4.4 for example shows that individuals are more likely to report hunger if the father was absent.) Parents who are of a worse physical constitution and shorter may be less able to feed their families. Since adult height is to some extent heritable, this may create a negative association between hunger and adult height of their offspring. Alternatively, there may be genetic predispositions such that tall parents and their potentially tall children suffer more from a lack of nutrition because their body requires a larger intake of food. In general, in the presence of unobserved family determinants, the net bias can be positive or negative.

Covariates. In the analysis, we include country fixed effects and country and gender-specific time trends.[23] As described in Subsection 4.2.3, trends in height for males are stronger than for females and variability in stature over time is lower for females than for males (Kuh et al. (1991)). The literature shows that average adult height increased with a linear trend during the period we study (Staub et al. (2011), Cole (2000)). For a panel of European countries this can be seen for example from Figure 1 in Deaton (2007) and Figure 1 in Schmidt et al. (1995). Figure 4.6 displays average height for all individuals from non-famine countries in the SHARE data confirming that in the absence of famines, the trend in height is linear, country-specific and gender-specific.

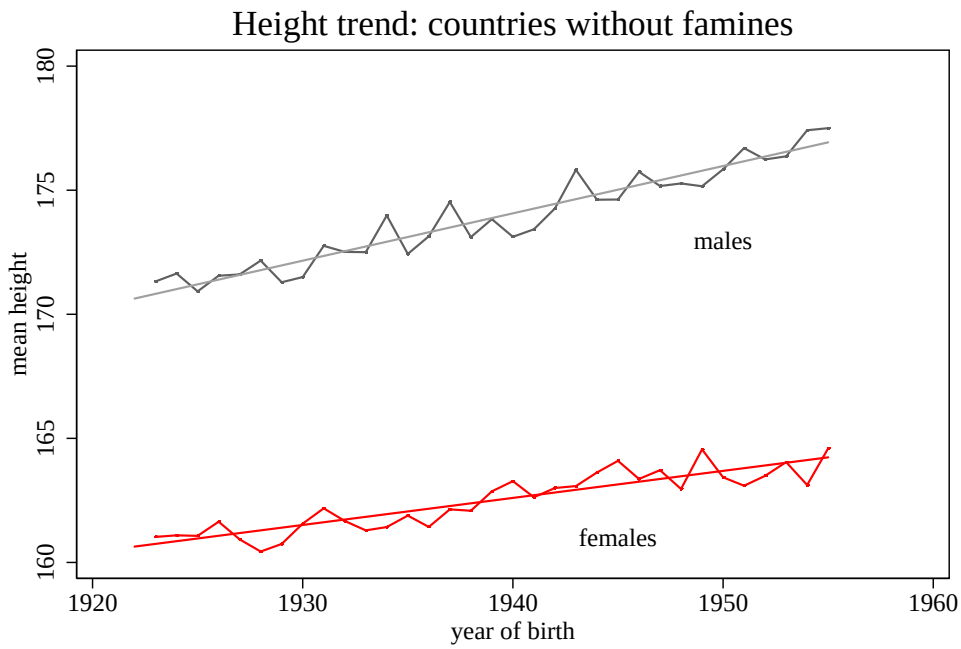
[23]The background household information in the childhood module of SHARELIFE refers to a respondent's living conditions at age 10. Since we are interested in incidence of hunger occurring before age 10, any such measure may be endogenous.

Table 4.4: Prevalence of hunger for different groups of individuals (conditional on famine exposure)

variables	total		region of birth				Netherlands				father occ (age 10)				father absent (age 10)			
	all		**urban**		**rural**		**west**		**east**		**high**		**low**		**yes**		**no**	
hunger measures																		
Episode of hunger 0-16	0.155	(0.36)	0.197	(0.40)	0.138	(0.35)	0.249	(0.43)	0.059	(0.24)	0.162	(0.37)	0.153	(0.36)	0.285	(0.45)	0.138	(0.34)
Episode of hunger 0-4	0.079	(0.27)	0.110	(0.31)	0.066	(0.25)	0.140	(0.35)	0.027	(0.16)	0.063	(0.24)	0.080	(0.27)	0.180	(0.39)	0.063	(0.24)
Episode of hunger 6-16	0.167	(0.37)	0.203	(0.40)	0.153	(0.36)	0.317	(0.47)	0.063	(0.24)	0.174	(0.38)	0.166	(0.37)	0.310	(0.46)	0.151	(0.36)
N	2401		700		1701		241		543		260		2112		267		2104	

Note: Standard errors in brackets. The table comprises summary statistics for both males and females. West/East signifies whether someone was born in the Western (famine affected) or Eastern (non-famine affected part of the Netherlands. High/Low means whether the father's occupation as measured at the age of ten of the child is high skilled (legislator, manager, senior official, professional) or low-skilled.

Figure 4.6: Trend in height



We also include controls for the degree of urbanization of the place of birth or the place where an individual lived at the time of our reference age.[24] We consider this information as a proxy for socio-economic background. Table 4.2 provides basic summary statistics for the overall sample, for the subsample of respondents who reported an incidence of hunger sometime before age 16, and for the subgroup which reported no incidence of hunger.

4.5 Empirical strategy

4.5.1 Instrumental variable methods

Our model framework is as follows,

$$Y_i = \psi(D_i, Z_i, X_i, \epsilon_i) \tag{4.1}$$

$$D_i = \phi(Z_i, X_i, \varepsilon_i), \tag{4.2}$$

where $D_i \in \{0, 1\}$ denotes severe hunger during a respective childhood period of individual i. $Z_i \in \{0, 1\}$ is a binary instrumental variable, where $Z_i = 1$ if an individual has experienced a period of major exogenous food restriction in her area of residence and $Z_i = 0$, otherwise. Y denotes adult height in centimeters. The vector of covariates is denoted by X.

[24] If an individual changed the accommodation in the year he turned six, information on the accommodation inhabited the year before was used.

If effects of nutrition on adult height are heterogeneous, we can identify the so-called local average treatment effect (LATE): the average effect among the "compliers", that is, those whose nutritional status is affected by a famine. The size of the complier group and the estimated effect may depend on the severity of the famine (Angrist et al. (1996)). Formally, the LATE is defined as:

$$\begin{aligned} \text{LATE} &= \mathrm{E}[Y_{D=1} - Y_{D=0} | D_{Z=1} > D_{Z=0}] \\ &= \int \left(\psi(1, Z, X, \epsilon) - \psi(0, Z, X, \epsilon)\right) dF_{X,Z,\epsilon | D_{Z=1} > D_{Z=0}} \end{aligned} \tag{4.3}$$

Identification of the effect is based on the assumption that the famine causes are valid instruments: For example, we assume that the allied food embargo in Greece did not have a direct influence on the health status of individuals 50 years after the event, other than through the effect on individual access to nutrition. Furthermore, we assume that our sample contains a subpopulation of compliers but no defiers, and that the probability of suffering from hunger in a famine is the same for individuals who actually suffered from a famine as for those who did not. Moreover, stratifying on additional covariates requires their supports to be the same for famine and non-famine groups.

We estimate the LATE nonparametrically using a Wald estimator. This estimator is the sample equivalent of $\mathrm{cov}(Y, Z)/\mathrm{cov}(D, Z)$. We use a version that controls for covariates. Exploiting the binary nature of Z, this is the sample equivalent of the following expression,

$$LATE = \frac{\int \mathrm{E}[Y|X = x, Z = 1] - \mathrm{E}[Y|X = x, Z = 0]\, f(x)dx}{\int \mathrm{E}[D|X = x, Z = 1] - \mathrm{E}[D|X = x, Z = 0]\, f(x)dx}. \tag{4.4}$$

Equation (4.4) allows us to compute the difference in mean outcomes for the group of compliers without having to specify a functional form for the effect of hunger on health outcomes. With only discrete covariates, we compute the numerator and denominator cell-wise and then integrate over the distribution of X, where the integral is a sum over various combinations of X-values. As a parametric alternative and robustness check, we also present results from the standard two stage least squares (2SLS) estimator.

4.5.2 Two-sample IV estimation

As explained above, imperfect recall in the “infant sample” may cause the incidence of hunger for that group to be misreported. We tackle this problem by replacing the estimate of $\mathrm{E}[D|X = x, Z = 1] - \mathrm{E}[D|X = x, Z = 0]$ in the denominator of (4.4) by the estimate of $[\mathrm{E}[D|X = x, Z = 1] - \mathrm{E}[D|X = x, Z = 0]]_{child}$ from the child sample (but integrating over the marginal distribution of covariates in the infant sample). Intuitively, the original Wald

estimator for the causal effect of the treatment D on the outcome Y starts with the reduced-form effect of the instrument Z on the outcome and subsequently corrects this for the imperfect compliance, i.e. for the fact that $Z = 1$ is not equivalent to $D = 1$. The correction term is the estimator of the denominator. There is no reason for why the correction term (i.e. the estimator of $\text{cov}(D, Z)$ in the absence of covariates) should be estimated with the same sample as the numerator (i.e. the estimator of $\text{cov}(Y, Z)$ in the absence of covariates). If the older children provide accurate observations of hunger D then the correction term based on the child sample is consistent for older children. And if the effect of a famine on undernutrition is the same among young children as among older children then we can use this correction term for the estimation of the causal effect of hunger for the younger children as well. In effect, this means that we use the two sample IV (2SIV) estimator developed by Angrist and Krueger (1992) and Arellano and Meghir (1992).

We bootstrap standard errors using 500 bootstrap iterations. As a robustness check, we estimate the effects of hunger around birth with information from two samples using *linear* models instead of the nonparametric Wald estimator. For this we use the two-sample two-stage least squares estimator (2S2SLS) proposed by Inoue and Solon (2010), which adjusts the original 2SIV estimator for use in small samples.

The assumption that the effect of a famine on undernutrition is the same among young children as among older children may be incorrect if parents or society at large respond to a famine by redistributing resources towards either the older or the younger children. During the famines, different food support policies were used for different child age classes. Food rations for pregnant mothers and babies were often relatively high and could be supplemented by direct food delivery (see Dols and van Arcken (1946), de Rooij et al. (2010), Klatt (1950), Valaoras (1946) and Neelsen and Stratmann (2011)). For babies, breastfeeding provides an additional protective factor. The fraction of mothers who breastfed their children was not reduced during the Dutch famine (Dean (1951)). At the same time, as we have seen in Section 4.3, special food aid programs were available for children in school-going ages. This was a response to the severity of nutritional shortages among those children. It does not necessarily indicate that society favored food allocation to those children over the allocation to infants.

More in general, parents may redistribute food supply across children with different ages within the household. For parents it may be more efficient to allocate resources towards stronger and older children, as these are more likely to survive a famine. Conversely, they may prefer to invest more into their younger children if they favor more equal outcomes,

or if they prefer all of their children to survive even if this comes at a cost for the older children's development. Whether decisions are driven by equity or efficiency concerns depends on the number of children, their probability of survival, parental preferences, and the parental budget constraint (Becker and Tomes (1976), Behrman (1997), Behrman et al. (1982)). In developing countries, parents' child investment decisions are generally driven by efficiency concerns (see e.g. Rosenzweig and Wolpin (1988)). For the developed economies, most research indicates that parents follow equity considerations (Griliches (1979), Del Bono et al. (2008)), although Datar et al. (2010) point in the opposite direction. The famine-stricken countries in our data would perhaps classify in-between current developed and developing countries, although parents living through the famines may have expected the famines to pass by relatively fast and to be followed by more prosperity.

All in all, it seems plausible that individuals in the infant sample were not more likely to suffer from undernutrition in response to the famine than individuals in the child sample. This suggests that if our two-sample assumption is violated and the resulting estimates are biased then most likely they provide a lower bound of the true effect.

4.5.3 Selection issues in famine studies

As in any study of long-run effects of in-utero exposure to famines, the survivors who were exposed to the famine may be systematically different in terms of unobserved characteristics than the survivors who were not exposed to the famine. This selectivity may lead to biased inference.

In our setting, one may distinguish between selective fertility and selective mortality after birth. Concerning fertility, it may be that families living in poor conditions experience a particularly strong fertility reduction during the famine. The same may apply to death in utero, spontaneous abortions, and stillbirths. As a result, the exposed birth cohorts may on average have less frail characteristics than the corresponding control cohorts. It is also conceivable that mortality in between birth and the moment of observation in our data depends on famine exposure, in the sense that such mortality may disproportionally affect the frailer individuals in the cohort. Both types of selectivity would then tend to reduce the observed difference in outcomes between the exposed cohorts and the control cohorts, which would imply that the estimated effect may underestimate the effect of exposure, in absolute value. Notice that any bias in the causal effect of hunger may be smaller than the bias in the reduced-form over-all effect of famine exposure, because less frail individuals may also be less likely to suffer from hunger.

Our data are not suitable to study the extent of selective mortality, since they only contain the exposed cohort members who are still alive at ages around 50 to 55. Selective fertility can only be controlled for to the extent that it is captured by characteristics of the birth region. Some of the studies listed in Section 4.3 that examine long-run effects of the Dutch, Greek and German famines argue that with these famines, selective fertility does not create a major source of bias. Retrospective information on the parents' social class is typically unavailable. However, Stein and Susser (1975), using military conscription data on men, report that the higher the occupational category of the father, the lower the reduction of the birth rate of the cohorts exposed to the Dutch famine, as compared to the non-exposed cohorts. To proceed, we consider the household's main income earner's occupational category when the respondent was 10 years old, provided by SHARELIFE. Famine-related changes in this distribution can yield at most suggestive evidence, since the household situation at age 10 may be endogenous to events that happened before. For example, the famine may have permanently driven individuals out of occupations like fishing. We compare cohorts born within our predefined famine periods including up to nine months thereafter, to those born within five years before and after this interval. See Table 4.5 for the results of this analysis. For Greece, we find that individuals born in the famine are somewhat more likely to be of a higher socio-economic status than those born before. This fraction, however, slightly declines for cohorts born after the famine. For all other countries in our sample, we do not find any statistically significant evidence for selective fertility.

4.6 Estimation results

4.6.1 Average causal long-run effects of hunger in later childhood

For ease of exposition, we start this section with the effects of hunger in later childhood (i.e., at ages 6-16). We focus on causal effects on adult height. First, we relate the outcome measures to the famine indicator while purposely dropping the actual undernutrition indicator. This corresponds to the commonly used approach of exploiting contextual variation or, put differently, not estimating the structural effect of hunger but rather obtaining a reduced-form estimate. We can thus show what the prediction would be in studies that have to remain agnostic about actual individual-specific undernutrition status. After this, we discuss the IV estimation results to see whether these studies yield accurate estimates of the causal effects.

Table 4.5: Occupational categories of main income earner when 10 by birth cohorts

	Greece			Germany			The Netherlands		
occupational status	pre-famine	famine	post-famine	pre-famine	famine	post-famine	pre-famine	famine	post-famine
High	0.030*	0.074	0.067	0.140	0.127	0.168	0.225	0.308	0.271
Low	0.967	0.926	0.925	0.849	0.861	0.824	0.775	0.692	0.701

Note: Pre-famine period refers to 5 years prior to famine; famine period refers to famine period and 9 months thereafter; post-famine period refers to 5 years after the end of the famine excluding the first 9 months. * indicates significant difference to famine-period level at 5 percent level. For the Netherlands, table includes only observations born in the Western part.

Table 4.6: Effects of hunger at age 6 – 16 on adult height in cm (males)

	reduced form	instrumental variables models		
	famine at age 6 – 16	2SLS	cond. Wald	cond. Wald – trend corrected
effect	0.294	1.644	-4.630	0.862
(s.e.)	(0.325)	(1.818)	(1.525)	(1.267)
t-stat.	0.902	0.904	-3.037	0.680

Note: Number of respondents reporting hunger = 189; control group = 2322. Significance computed using standard errors clustered by country-year cells. Significance computed using bootstrap (500 replications) when applying Wald estimators. Control Variables are a dummy for whether the accomodation at age 6 has been in rural area, country fixed effects, and a country specific trend in year of birth.

Table 4.7: Effects of hunger at age 6 – 16 on adult height in cm (females)

	reduced form	instrumental variables models		
	famine at age 6 – 16	2SLS	cond. Wald	cond. Wald – trend corrected
effect	0.227	2.363	-6.730	1.280
(s.e.)	(0.259)	(2.790)	(2.058)	(1.118)
t-stat.	0.877	0.847	-3.270	1.145

Note: Number of respondents reporting hunger = 175; control group = 2684. Significance computed using standard errors clustered by country-year cells. Significance computed using bootstrap (500 replications) when applying Wald estimators. Control Variables are a dummy for whether the accomodation at age 6 has been in rural area, country fixed effects, and a country specific trend in year of birth.

Tables 4.6 and 4.7 show results for men and women, respectively. The set of covariates is listed in the notes underneath the table; it is limited for reasons detailed in Section 4.4.[25] The first column of each table concerns the reduced-form estimates of famine exposure. We find for both males and females that famine exposure in the age window 6 to 16 does not significantly affect adult height.

Next, we verify that the famine instrument is informative. Table 4.8 shows the results of the first stage. linear regression models where the dependent variable is the undernutrition indicator. We show results from both univariate regressions relating hunger to famine exposure only and the full first stage results. In all specifications, the famine indicator is highly significant. For women, the effect of famine exposure on the probability to have suffered hunger is smaller. The models yield F-statistics beyond 20, i.e. beyond the typically recommended value of 10. We thus conclude that the instrumental variable is informative.

We use IV analyses to assess the causal effects of hunger at ages 6 to 16. In Tables 4.6 and 4.7 we provide estimates of two-stage least squares estimations for all outcomes and augment this set of results by estimates of local average treatment effects using a non-

[25] Full lists of coefficient values for both the reduced form models and the IV models are displayed in Table 4.18.

Table 4.8: First-stage coefficients for the probability of experiencing hunger at age 6 – 16

	males		females	
Experienced famine being 6-16 (1 = yes)	0.206***	0.179***	0.135***	0.096***
	(0.020)	(0.024)	(0.018)	(0.025)
German Sample		2.325		4.730*
		(3.992)		(2.668)
Dutch Sample		2.950		2.550
		(2.873)		(2.667)
Lived in rural area at age 6		-0.026**		-0.026***
		(0.010)		(0.009)
Year of birth		-0.002*		-0.002**
		(0.001)		(0.001)
Year of birth × Dutch		-0.002		-0.001
		(0.001)		(0.001)
Year of birth × German		-0.001		-0.002*
		(0.002)		(0.001)
Constant	0.030***	4.197*	0.034***	4.781**
	(0.005)	(2.124)	(0.006)	(1.919)
R^2	0.105	0.116	0.051	0.070
F-Stat.	102.942	27.143	55.270	21.152
N	2511	2511	2859	2859

Note: Standard errors clustered within country-birthyear cells in parentheses. *,**,*** indicates significance at the 10%, 5%, and 1% level, respectively.

parametric Wald estimator to assess what differences in causal effects occur when dropping any functional form or treatment effect homogeneity assumptions inherent in the models. Using a Wald estimator, we cannot control for a linear time trend (the number of strata defined by all possible combinations of covariate values would be too large). This is likely to bias our estimates if height follows a secular trend. We may address the latter to some extent using the following ad-hoc approach. In a first step, height is regressed on a country-specific time trend and country fixed effects. The residuals are then used as the dependent variable in regressions using the Wald estimator. For both men and women, the IV results suggest that

Table 4.9: Effects of hunger at age 6 – 16, instrument is famine *duration* in age interval

	males		females	
	reduced form	2SLS	reduced form	2SLS
effect	0.054	0.690	0.034	1.015
(s.e.)	(0.197)	(2.498)	(0.119)	(3.506)
t-stat.	0.273	0.276	0.284	0.290

Note: Number of male (female) respondents reporting hunger = 189 (175); control group = 2322 (2684). Standard errors clustered by country-year cells.

hunger at ages in [6, 16) does not have an impact on adult height since all estimates reported in the last columns of the table lack statistical significance. The fully nonparametric Wald estimates are large and significant both when analyzing men and women but lose size and precision when first detrending height.

For females we do not reject the hypothesis of zero effects. This may be due to two factors. First, as discussed in Subsection 4.2.3, females are less responsive to early-life conditions. Secondly, a positive effect of adverse conditions just before puberty on height has been found to be present in women exposed to the Leningrad siege (see Koupil et al. (2009)). A biological mechanism that could drive this result also mentioned in this study is that an adverse health shock for women at this time could postpone menarche. Since the linear growth period before menarche may thus be prolonged, females could actually grow taller under such circumstances than in the absence of a shock.

We test the robustness of the above results by replacing the binary famine instrument with the length[26] of famine exposure within the defined age interval in months; we rescale this variable by division by twelve. The length of exposure captures the cumulative severity of the famine. Estimating the same models as before with the new instrument produces effects as depicted in Table 4.9. The results are qualitatively unchanged.

4.6.2 Average causal long-run effects of hunger around birth

The analysis so far has used retrospective information on hunger experiences in childhood. The self-reported nature of this information precludes an analogous approach when analyzing hunger effects in the very first years of life. We thus proceed as detailed above by estimating hunger propensities conditional on background factors for the first years of life by using information on the famines' impact on cohorts born earlier. We report findings for famine exposure in the first four years of life, pooled with exposure in utero. Gestation is defined to begin nine months before birth. We require gestation and the famine period to overlap more than two months in order to preclude additional noise. For Dutch individuals we take the region of the accommodation at birth into account when assigning the value of the famine exposure variable.

We start our discussion by presenting reduced form estimates of the effect of famines at this early age on outcomes. The results are displayed in the first columns of Tables 4.10 and 4.11.[27] They show that famine exposure reduces adult male height on average by 0.68

[26] Specifically, this variable captures just the overlap between the treatment window age [6, 16) and the famine interval defined by the calendar dates given in the data section.

[27] Table 4.19 presents the full list of coefficient estimates for the reduced form models.

Table 4.10: Reduced form estimates and treatment effects of hunger at age 0 – 4 or in utero on adult height in cm (males)

	famine at age 0 – 4	cond. Wald	cond. Wald – trend corrected
effect	-0.622	-1.986	-2.664
(s.e.)	(0.274)	(1.702)	(1.287)
t-stat.	-2.271	-1.167	-2.069

Note: Sample size = 2511; thereof famine-exposed at age 0 – 4: = 401; thereof treated / famine-exposed at age 6 – 16 = 189 / 561. Regressions include control country fixed effects, a dummy for urbanization of birthplace and country specific trends in year of birth (trends cannot be included when using Wald estimator). Significance for Wald estimator computed using a bootstrap (500 repetitions). All standard errors clustered by country-year cells.

Table 4.11: Reduced form estimates and treatment effects of hunger at age 0 – 4 or in utero on adult height in cm (females)

	famine at age 0 – 4	cond. Wald	cond. Wald – trend corrected
effect	0.300	3.859	1.904
(s.e.)	(0.296)	(2.819)	(2.194)
t-stat.	1.014	1.369	0.868

Note: Sample size = 2859; thereof famine-exposed at age 0 – 4: = 420; thereof treated / famine-exposed at age 6 – 16 = 175 / 571. Regressions include control country fixed effects, a dummy for urbanization of birthplace and country specific trends in year of birth (trends cannot be included when using Wald estimator). Significance for Wald estimator computed using a bootstrap (500 repetitions). All standard errors clustered by country-year cells.

centimeters. The estimate for females is insignificant. Returning to our findings for men, if we now take a crude estimate for the effect of famine exposure on the propensity to report hunger to amount to 20 percent, then we expect the causal effect of hunger on adult height equal around -3.4 centimeters on average (applying the Wald estimator formula).

In what follows, the prevalence of hunger for ages 6 to 16 is taken as the reference point for the hunger prevalence in early years of life. To obtain estimates of the average causal effect of undernutrition around birth we use the 2-sample IV methods from Section 4.5. The last columns of Tables 4.10 and 4.11 show the results of both the fully nonparametric Wald Estimator and the Wald estimator when applied to a detrended version of adult height as the dependent variable. For men we find that a famine-caused hunger experience early in life has a negative impact of almost 3 centimeters when using the Wald estimator both with and without correcting for a secular trend in height.[28] For women we find insignificant effects. The results show that the reduced form estimate for men presented above underestimates

[28]The fact that a trend correction does not produce different results as it did for the age window 6 to 16 may be due to the fact that cohorts exposed to famine early in life may now be on average as old as those not exposed while for later childhood, the non-exposed group may be on average younger.

the underlying causal effect of hunger on height by a factor exceeding 4 and amounting to more than 2 centimeters. We also used a 2S2SLS estimator for this exercise. This restricts both the relationship between height and hunger and the first stage to be linear. Using this alternative estimation technique produces causal effects in the same order of magnitude and significance.

Our finding that the height of males is more affected by early nutritious shocks than female stature is supported by studies described in Section 4.2 showing that boys tend to be biologically more sensitive to early health shocks than girls. Because frail males are more likely to die than frail females, dynamic selection may even lead some studies of severe famines to find larger adverse famine effects on females than males (Mu and Zhang (2011)). For the famines we investigate in this study, there does not exist evidence for a drop in the sex ratio except for a small reduction in Greece (Cramer and Lumey (2010), Jürges (2013), Neelsen and Stratmann (2011)). The male disadvantage in coping with nutritious shocks early in life is thus likely reflected in worse later life health outcomes for males than for females.

4.6.3 Sensitivity analyses and tests

Data from other countries. Recall that the famines we consider took place around the time of World War II and its aftermath. To examine whether the estimated effects reflect the effect of the turbulences around the war instead of the famine, we perform a placebo test and a sensitivity analysis. In the placebo test, we select Belgium, France, and Italy as "clones" for the Netherlands, Germany, and Greece, respectively, and pretend that each of these countries was exposed to a famine in the same period as its companion country. We select SHARE samples for these countries analogous to the description in Section 4.4. We first estimate a first stage, i.e. the effect of pseudo famine exposure on the probability to report hunger for ages 6 to 16. These turn out to be statistically significant for men and women; however, in both cases, the size of the effects is less than half the size of what we find in the original analysis. As noted in Section 4.4, this may reflect the effect of World War II and its aftermath in Belgium, France and Italy. Next, we estimate placebo reduced-form relationships of pseudo famine exposure on adult height for both ages 6 to 16 and in utero to age 4. These results are displayed in Table 4.12 and provide no significant evidence along the lines of the patterns we found in the main analyses. This confirms that we capture causal effects of hunger in the main analyses in this section.

Table 4.12: Robustness of effects at ages 6 – 16 (top panel) and ages 0 – 4 or in utero (bottom panel): placebo famine exposure and sensitivity and inclusion of other countries

	ages 6 – 16	
	males	
	placebo famine exposure	inclusion of other countries
Effect	0.155	0.288
(S.E.)	(0.331)	(0.325)
t-stat.	0.467	0.886
	females	
	placebo famine exposure	inclusion of other countries
Effect	0.019	0.224
(S.E.)	(0.263)	(0.258)
t-stat.	0.073	0.870

	ages 0 – 4 or in utero	
	males	
	placebo famine exposure	inclusion of other countries
Effect	0.050	-0.625
(S.E.)	(0.294)	(0.273)
t-stat.	0.170	-2.287
	females	
	placebo famine exposure	inclusion of other countries
Effect	0.321	0.299
(S.E.)	(0.261)	(0.295)
t-stat.	1.231	1.014

Note: For column 1, original sample has been replaced by individuals from France (pseudo famine exposure as in Germany), Italy (pseudo famine exposure as in Greece), and Belgium (pseudo famine exposure as in the Netherlands). Sample Size = 2864 males (3417 females); thereof pseudo-famine exposed = 836 males (929 females). For column 2, original sample has been extended to include France, Belgium, and Italy. Sample Size = 5375 males (6276 females); thereof famine exposed = 555 males (573 females). Significance computed using standard errors clustered by country-year cells. Control variables are a dummy for whether the accomodation at birth has been in rural area, country fixed effects, and country specific trend in year of birth.

To proceed, we add Belgium, Italy and France as famine-free countries to the sample from the Netherlands, Greece and Germany, and we perform reduced-form estimations of actual famine exposure on this extended sample. Since these countries also experienced turbulent conditions around World War II, we would expect the reduced-form impacts of the famine to lose size and precision compared to the baseline results, if the baseline results were driven by these turbulences rather than by the famine. The reduced form estimate for this analysis is displayed in the second column of the Table 4.12. It turns out that for all means and purposes, the results are the same as in the original analysis. If anything, the estimates slightly gain in precision, which can be attributed to the increase in the sample size. This confirms that the results are not driven by exposure early in life to non-famine adverse conditions.[29]

Variable definitions and operationalizations. In addition to the sensitivity analyses discussed so far, we perform a range of estimations to assess the robustness of the results with respect to a number of assumptions and decisions concerning the operationalization of the key variables.

First, note that individuals who are not undernourished within a given age interval may be undernourished within another age interval. This may contaminate the control groups, which may in turn lead to an underestimation of the effects we are after. We examine this issue in a number of ways. First, as we have seen above, the addition of samples from non-famine countries, consisting mostly of control group members, does not lead to a change in the results. Secondly, we re-estimate the baseline reduced-form models and account for famine exposure in the two relevant age intervals simultaneously. Table 4.13 presents the results. (Notice that the binary measures of famine exposure are not mutually exclusive.) Our main findings are robust to this alternative. We conclude that the results are not sensitive to this potential contamination of the control groups.

Next, we allow the severity of the famine to vary with the degree of urbanization of the region where the respondent lived at the onset of the famine. As suggested in Section 4.3, rural areas may have easier access to food. To proceed, we include the interaction of famine and urbanization as an additional instrumental variable. Using two instruments entails that we use the 2SLS and 2S2SLS estimators to obtain causal effect estimates for the age intervals

[29]The recent working paper by Havari and Peracchi (2012) adopts our IV approach, using war exposure as an instrumental variable for a number of adverse childhood conditions, including hunger. They use data from a larger number of countries. Our results of the effect of hunger on height are conservative compared to their results which for a number of countries exceed 10 cm.

$[6, 16)$ and $[-0.75, 4)$, respectively. Table 4.14 presents the estimates. Clearly, our findings on the causal effect of undernutrition are unaffected. The interaction term has the expected first stage coefficient: having lived in a rural area at the time of the famine significantly lowers the probability for having suffered from undernutrition.

Concerning our key variables, it is important to assess how their definitions affect the results. This concerns, first of all, the precise starting and ending dates of each of the famines. For example, for Germany, one may argue that May 1945 should be added to the famine period. It turns out that such one-month changes have no effect on the results. Next, we may vary the age intervals within which early-life conditions are assumed to exert their influence. For example, we may exclude "in utero" from the age interval $[-0.75, 4)$, or we may exclude age 3 from these intervals, or consider a wider interval. Tables 4.15 and 4.16 give results if we define the age interval to be $[-0.75, 3)$ or $[-0.75, 5)$. With the extended interval, our estimates become smaller and are measured less precisely; this confirms that conditions at the onset of life are the most important. With the age interval $[-0.75, 3)$, the estimates are similar to those presented in the main analysis, although the precision is somewhat lower. We conclude that our original choice of ages 0 to 4 or in utero exposure as the relevant age interval provides a good compromise between effect relevance and statistical precision.

Changing the child age exposure interval in an analogous way does not change the results either. In particular, it is interesting to consider the alternative interval $[8, 16)$ since Figure 4.1 shows that there is still a slight increase in the propensity to report hunger between age 6 to 8 and this might lead to an over-estimate of the effect of hunger in absolute terms. Table 4.17 shows the results for men and women, respectively. For men, we indeed find that using this alternative specification slightly reduces the local average treatment effect. However, the difference between the new and the original estimate is negligible, at less than 0.1 centimeters.

Next, we drop our measure of urbanization at the reference age from the covariates in order to see whether results are sensitive to this change. We also include life expectancy at birth as an additional covariate, to account for the high rates of attrition in SHARE (see Tables 4.18 and 4.19). It turns out that the results are insensitive with respect to all this. Other sensitivity analyses suffer from a lack of informative observations. As is clear from the above, the "in utero" age interval for exposure (using the recording of hunger in the birth year) does not give a sufficient number of exposed individuals for a meaningful analysis. For separate estimation by country, the samples by country are too small as well. Furthermore, we suspect that results by country are more sensitive to effects of cohort-specific events.

4.7 Conclusions

This chapter investigates the causal effect of undernutrition in infancy and childhood on adult height as a proxy for late life health, using instrumental variable estimation. We deal with the problem of selective recall and systematic underreporting of hunger periods in utero or infancy by applying two-sample IV estimation. Specifically, we estimate the probability to report hunger when exposed to a famine around birth by using the observed association between hunger and famine at higher childhood ages. This serves as input in the adjusted nonparametric Wald estimator for data from two different samples. We bootstrap the standard errors.

For males undernutrition in the age interval from in utero until age 4 results in a reduction of adult height of around 3 cm. This effect remains after controlling for trends in height over time. For women, we do not find evidence of an effect of adverse nutritional conditions on adult height.

What distinguishes our results from the literature on contextual famine effects is that we are the first to use the occurrence of famines to obtain an estimate of the average *causal* effect of *nutritional shortages* during childhood on health later in life. We can thus compare the reduced-form famine effects that are usually reported in the literature with the causal effects of undernutrition. Our estimated reduced-form famine effects are in accordance to those in the reduced-form studies of the long run effects of exposure to the famines that we consider (those studies were discussed in Section 3). However, our estimated causal effects are about four times larger than our reduced-form famine effects. This emphasizes the importance of nutrition in early childhood - over and above the findings and statements in the famine literature.

The difference between contextual and causal effects is related to the rate of "compliance" to the famine. In our study, the fraction of "compliers" is close to 20%. To the extent that reduced-form studies implicitly assume perfect compliance, this assumption seems untenable even for shocks as severe as the famines used in this study.[30]

The analyses in this chapter are restricted by some notable data limitations. First, the samples are small. Admittedly, the full SHARE data cover many European countries and include over 20,000 elderly respondents, but we only use men in subsamples from the

[30]Some of the reduced-form studies do explicitly acknowledge the difference between famines or episodes with reduced food availability on the one hand and nutritional shortages early in life on the other, and they carefully define the parameter of interest to be the over-all "intention to treat" effect. See for example Almond and Mazumder (2011) and Lumey et al. (2011) for excellent expositions.

three countries with famines. Within these, the subsets of individuals who were exposed to famines during childhood are of even smaller size, because the famines had a duration of at most a few years. This implies, among other things, that we can only examine exposures within age intervals of at least 3 years, since otherwise the numbers of "treated" and "controls" are insufficient for reliable inference. Another data limitation is that the individual spells of severe nutritional shortages are only measured in full calendar years. Inevitably, then, the assignment rules for the actual treatment status are open to debate. Fortunately, the results are not sensitive to moderate changes in these rules, and in general, the results are in agreement to the findings in the reduced-form literature where famine exposure is directly related to adult height.

Nevertheless, the data limitations rule out the analysis of a number of interesting issues. This concerns, first of all, the connection between the timing of the spell of nutritional shortage and the age of the child. The famine literature distinguishes between different reduced-form effects at different stages of pregnancy, for different outcomes. More in general, the developmental origins literature finds that long-run reduced-form effects of conditions in utero and right after birth are larger than those at subsequent ages. The size of the effects may be non-monotonous during the pre-puberty and puberty ages. Our data do not enable us to go into such details. A second issue concerns the temporal and spatial variation in the strength of the famine instrument. Contextual information can be used to create an indicator of famine severity. For example, one may use official food rations per region and month. However, the meaning of a ration size depends on the country, since in some countries charities and the informal sector are more important than in others. As we have seen, the sample sizes are too small to allow for meaningful estimation by country, and on top of that, the temporal connection between a famine and the observed spell of nutritional shortage is insufficiently tight to exploit fine temporal contextual variations. The present study may therefore motivate the construction or usage of data sets that are larger and/or more focused on specific spatial areas, where ideally, such data sets would have more elaborate retrospective information on childhood episodes of hunger.

4.A Additional tables

Table 4.13: Reduced form estimates of famine exposure in infancy and childhood

	males		females	
	famine 0–4	famine 6–16	famine 0–4	famine 6–16
effect	-0.571	0.196	0.392	0.351
(s.e.)	(0.297)	(0.361)	(0.310)	(0.289)
t-stat.	-1.921	0.542	1.266	1.215
N	2511		2859	

Note: Further control variables are urbanization of the place of birth, country fixed effects and country specific trends in year of birth. All standard errors clustered by country-year cells.

Table 4.14: 2SLS results for hunger at age 6 – 16 on height using the interaction between urbanization of accomodation and famine exposure as additional instrument

		2SLS				2-sample 2SLS	
		males		females		males	females
		outcome eq.	first stage	outcome eq.	first stage	outcome eq.	outcome eq.
hunger	effect	0.299		3.646		-3.035	0.637
	(s.e.)	(1.558)		(2.460)		(1.739)	(2.168)
	t-stat.	0.192		1.482		-1.745	0.294
famine	effect		0.235		0.145		
	(s.e.)		(0.018)		(0.016)		
	t-stat.		12.687		8.882		
famine × rural	effect		-0.125		-0.111		
	(s.e.)		(0.024)		(0.022)		
	t-stat.		-5.190		-5.120		
N		2511		2859		2511	2859

Note: Standard errors clustered by country-year cells. Control Variables for outcome equation are a dummy for urbanization of accomodation, country fixed effects, and a country specific trend in year of birth. Standard errors for 2-sample 2SLS have been computed using a bootstrap with 500 repititions. First stage for 2-sample 2SLS is the first stage for 2SLS analysing effects of hunger at age 6–16.

Table 4.15: Reduced form estimates and treatment effects of hunger in different age intervals on adult height in cm (males)

	hunger at age 0–3 or in utero		
	reduced form	cond. Wald	cond. Wald – trend corrected
effect	-0.625	-1.611	-2.708
(s.e.)	(0.283)	(1.862)	(1.320)
t-stat.	-2.209	-0.865	-2.052
	hunger at age 0–6 or in utero		
	reduced form	cond. Wald	cond. Wald – trend corrected
Effect	-0.434	-2.244	-1.747
(s.e.)	(0.299)	(1.641)	(1.358)
t-stat.	-1.452	-1.367	-1.286

Note: Sample size = 2511. Regressions include control country fixed effects, a dummy for urbanization of birthplace and country specific trends in year of birth (trends cannot be included when using Wald estimator). Significance for Wald estimator computed using a bootstrap (500 repetitions). All standard errors clustered by country-year cells.

Table 4.16: Reduced form estimates and treatment effects of hunger in different age intervals on adult height in cm (females)

	hunger at age 0–3 or in utero		
	reduced form	cond. Wald	cond. Wald – trend corrected
Effect	0.026	2.269	0.088
(S.E.)	(0.281)	(2.530)	(1.552)
t-stat.	0.093	0.897	0.057
	hunger at age 0–6 or in utero		
	reduced form	cond. Wald	cond. Wald – trend corrected
Effect	0.155	0.806	0.800
(S.E.)	(0.245)	(2.001)	(1.417)
t-stat.	0.633	0.403	0.564

Note: Sample size = 2859. Regressions include control country fixed effects, a dummy for urbanization of birthplace and country specific trends in year of birth (trends cannot be included when using Wald estimator). Significance for Wald estimator computed using a bootstrap (500 repetitions). All standard errors clustered by country-year cells.

Table 4.17: Reduced form estimates and treatment effects of hunger at age 0 – 4 or in utero on adult height in cm, setting reference age window to 8–16

	males		
	famine at age 0 – 4	cond. Wald	cond. Wald – trend corrected
effect	-0.683	-2.154	-2.808
(s.e.)	(0.268)	(1.713)	(1.216)
t-stat.	-2.550	-1.258	-2.310
	females		
	famine at age 0 – 4	cond. Wald	cond. Wald – trend corrected
effect	0.259	3.147	1.282
(s.e.)	(0.288)	(2.241)	(1.605)
t-stat.	0.899	1.405	0.799

Note: Sample size = 2511 males (2859 females); thereof famine-exposed at age 0 – 4: = 398 males (417 females); thereof treated / famine-exposed at age 8 – 16 = 170 / 441 males (151 / 456 females). Regressions include control country fixed effects, a dummy for urbanization of birthplace and country specific trends in year of birth (trends cannot be included when using Wald estimator). Significance for Wald estimator computed using a bootstrap (500 repetitions). All standard errors clustered by country-year cells.

Table 4.18: Coefficient estimates for models explaining adult height

	males				females			
	reduced form	IV model	reduced form	IV model	reduced form	IV model	reduced form	IV model
Experienced famine being 6-16 (1 = yes)	0.294		0.201		0.227		0.242	
	(0.325)		(0.337)		(0.259)		(0.269)	
Lived in rural area at age 6	-0.380	-0.338	-0.387	-0.357	0.018	0.080	0.018	0.086
	(0.257)	(0.259)	(0.257)	(0.260)	(0.178)	(0.195)	(0.178)	(0.197)
Year of birth	0.161***	0.165***	0.118***	0.124***	0.133***	0.139***	0.141***	0.152***
	(0.026)	(0.028)	(0.039)	(0.045)	(0.018)	(0.024)	(0.027)	(0.039)
Year of birth × Dutch	-0.063*	-0.061*	-0.037	-0.038	-0.062**	-0.059**	-0.066**	-0.065**
	(0.035)	(0.034)	(0.040)	(0.041)	(0.025)	(0.023)	(0.029)	(0.029)
Year of birth × German	0.007	0.009	0.032	0.031	-0.013	-0.007	-0.016	-0.013
	(0.036)	(0.035)	(0.040)	(0.040)	(0.019)	(0.019)	(0.021)	(0.020)
German Sample	-9.864	-13.686	-58.401	-56.847	26.883	15.707	34.143	26.489
	(69.402)	(68.733)	(77.570)	(78.008)	(36.313)	(35.992)	(40.678)	(39.761)
Dutch Sample	128.485*	123.636*	78.245	79.218	124.252**	118.226***	132.432**	131.070**
	(68.944)	(66.912)	(78.849)	(79.167)	(48.295)	(45.567)	(56.126)	(56.636)
Episode of Hunger being 6-16 (1 = yes)		1.644		1.170		2.363		2.613
		(1.818)		(1.954)		(2.790)		(3.035)
Life expectancy at birth			0.055*	0.050			-0.009	-0.015
			(0.030)	(0.033)			(0.022)	(0.027)
Constant	-140.031***	-146.931***	-59.497	-71.497	-96.449***	-107.747**	-110.094**	-131.424*
	(50.766)	(55.212)	(73.883)	(86.227)	(34.539)	(46.919)	(52.108)	(74.597)
R^2	0.208	0.201	0.208	0.205	0.135	0.123	0.135	0.121
N	2511	2511	2511	2511	2859	2859	2859	2859

Note: Standard errors clustered within country-birthyear cells in parentheses. *,**,*** indicates significance at the 10%, 5%, and 1% level, respectively.

Table 4.19: Coefficient estimates for models explaining adult height

	males		females	
Experienced famine being 0-4 (1 = yes)	-0.622**	-0.533	0.300	0.393
	(0.274)	(0.348)	(0.296)	(0.351)
Lived in rural area at age 0	-0.361	-0.362	0.081	0.081
	(0.272)	(0.272)	(0.180)	(0.181)
Year of birth	0.148***	0.133***	0.128***	0.112***
	(0.021)	(0.034)	(0.012)	(0.023)
Year of birth × Dutch	-0.052	-0.043	-0.059**	-0.050*
	(0.033)	(0.037)	(0.022)	(0.026)
Year of birth × German	0.020	0.029	-0.015	-0.008
	(0.036)	(0.039)	(0.018)	(0.020)
German Sample	-35.793	-51.798	31.932	17.405
	(70.588)	(75.642)	(35.000)	(38.422)
Dutch Sample	106.100*	88.782	117.952***	100.976**
	(63.414)	(71.422)	(43.636)	(50.883)
Life expectancy at birth		0.022		0.022
		(0.036)		(0.026)
Constant	-115.205***	-85.999	-86.575***	-57.155
	(41.286)	(64.178)	(24.205)	(43.713)
R^2	0.209	0.209	0.135	0.135
N	2511	2511	2859	2859

Note: Standard errors clustered within country-birthyear cells in parentheses. *,**,*** indicates significance at the 10%, 5%, and 1% level, respectively.

Bibliography

AIZER, A., L. STROUD, AND S. BUKA (2012): "Maternal Stress and Child Outcomes: Evidence from Siblings," *NBER Working Paper 18422*.

AKBULUT-YUKSEL, M. (2009): "Children of war: the long-run effects of large-scale physical destruction and warfare on children," Institute for the Study of Labor (IZA) Discussion Paper No. 4407.

AKBULUT-YUKSEL, M., G. J. VAN DEN BERG, AND K. WALLICZEK (2013): "Exploiting spatial and temporal variation in nutrition at birth, using data on infant mortality and famines: a study of long-run health effects," *mimeo*.

ALDERMAN, H. (2012): "The response of child nutrition to changes in income: linking biology with economics," *CESifo Economic Studies*, 58, 256–273.

ALMOND, D. (2006): "Is the 1918 Influenza Pandemic Over? Long-term Effects of In Utero Influenza Exposure in the Post-1940 U.S. Population," *Journal of Political Economy*, 114, 672–712.

ALMOND, D. AND J. CURRIE (2011a): "Human Capital Development before Age Five," in *Handbook of Labor Economics Vol. 4b*, Elsevier, 1315–1486.

——— (2011b): "Killing me softly: The fetal origins hypothesis," *The Journal of Economic Perspectives*, 25, 153–172.

ALMOND, D. AND B. MAZUMDER (2011): "Health Capital and the Prenatal Environment: The Effect of Ramdan Observance During Pregnancy," *American Economic Journal: Applied Economics*, 3, 56–85.

AMERICAN PSYCHIATRIC ASSOCIATION (2000): *Diagnostic and statistical manual of mental disorders: DSM-IV-TR.*, Diagnostic and Statistical Manual of Mental Disorders, American Psychiatric Association.

ANGRIST, J. AND A. KRUEGER (1992): "The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables with Moments from Two Samples," *Journal of the American Statistical Association*, 87, 328–336.

ANGRIST, J. D., G. W. IMBENS, AND D. B. RUBIN (1996): "Identification of Causal Effects Using Instrumental Variables," *Journal of the American Statistical Association*, 91, 444–455.

ARELLANO, M. AND C. MEGHIR (1992): "Female labour supply and on-the-job search: an empirical model estimated using complementary data sets," *The Review of Economic Studies*, 59, 537.

BANKS, J., C. O'DEA, AND Z. OLDFIELD (2010): "Cognitive Function, Numeracy and Retirement Saving Trajectories," *The Economic Journal*, 120, F381–F410.

BANKS, J. AND Z. OLDFIELD (2007): "Understanding pensions: cognitive function, numerical ability and retirement saving," *Fiscal Studies*, 28, 143–170.

BANNING, C. (1946): "Food Shortage and Public Health, First Half of 1945," *Annals of the American Academy of Political and Social Science*, 245, 93–110.

BARKER, D. (1990): "The fetal and infant origins of adult disease." *BMJ: British Medical Journal*, 301, 1111.

——— (1992): *Foetal and infant origins of adult diseases*, BMJ Publishing Group, London.

——— (1994): *Mothers, babies, and disease in later life*, BMJ Publishing Group London.

——— (2007): "The origins of the developmental origins theory," *Journal of Internal Medicine*, 261, 412–417.

BARKER, D., C. OSMOND, AND J. GOLDING (1990): "Height and mortality in the counties of England and Wales," *Annals of Human Biology*, 17, 1–6.

BATESON, P. (2001): "Fetal experience and good adult design," *International Journal of Epidemiology*, 30, 928–934.

BECKER, G. AND N. TOMES (1976): "Child Endowments and the Quantity and Quality of Children," *Journal of Political Economy*, 84, 143–162.

BEHRMAN, J. R. (1997): "Intrahousehold Distribution and the Family," in *Handbook of Population and Family Economics*, ed. by M. R. Rosenzweig and O. Stark, New York: Elsevier Science, vol. 1A, 125–187.

BEHRMAN, J. R., R. A. POLLAK, AND P. TAUBMAN (1982): "Parental Preferences and Provision for Progeny," *Journal of Political Economy*, 90, 52–73.

BERGENDAHL, M., M. L. VANCE, A. IRANMANESH, M. O. THORNER, AND J. D. VELDHUIS (1996): "Fasting as a Metabolic Stress Paradigm Selectively Amplifies Cortisol Secretory Burst Mass and Delays the Time of Maximal Nyctohemeral Cortisol Concentrations in Healthy Men," *Journal of Clinical Endocrinology and Metabolism*, 81(2), 692–699.

BIGGS, Q. M., J. M. GUIMOND, C. S. FULLERTON, R. J. URSANO, C. GRAY, M. GOLDENBERG, D. REISSMAN, E. MCCARROLL, JAMES, P. SANTIAGO, AND M. P. TYLER (2012): "The Epidemiology of Acute Stress Disorder and Other Early Responses to Trauma in Adults," in *The Oxford Handbook of Traumatic Stress Disorders*, Oxford Library of Psychology, 69–83.

BLACK, S. E., P. J. DEVEREUX, AND K. G. SALVANES (2014): "Does Grief Transfer across Generations? In-Utero Deaths and Child Outcomes," Institute for the Study of Labor (IZA) Discussion Paper No. 8043.

BLOM, A. AND M. SCHRÖDER (2011): "Sample composition 4 years on: Retention in SHARE Wave 3," in *SHARELIFE Methodology*, 55–61.

BOLDORF, M. (1998): *Sozialfürsorge in der SBZ, DDR 1945 - 1953: Ursachen, Ausmass und Bewältigung der Nachkriegsarmut*, Stuttgart: Steiner.

BOLIN, K., B. LINDGREN, AND P. LUNDBORG (2008): "Your next of kin or your own career? Caring and working among the 50+ of Europe," *Journal of Health Economics*, 27, 718–738.

BOLTEN, M. I., H. WURMSER, A. BUSKE-KIRSCHBAUM, M. PAPOUSEK, K.-M. PIRKE, AND D. HELLHAMMER (2011): "Cortisol levels in pregnancy as a psychobiological predictor for birth weight," *Archives of Women's Mental Health*, 14, 33–41.

BÖRSCH-SUPAN, A., A. BRUGIAVINI, H. JÜRGES, A. KAPTEYN, J. MACKENBACH, J. SIEGRIST, AND G. WEBER (2008): *Health, ageing and retirement in Europe (2004-2007). Starting the longitudinal dimension*, Mannheim Research Institute for the Economics of Aging (MEA).

BÖRSCH-SUPAN, A. AND H. JÜRGES, eds. (2005): *The Survey of Health, Aging, and Retirement in Europe – Methodology*, Mannheim Research Institute for the Economics of Aging (MEA).

CAMERON, N. AND E. W. DEMERATH (2002): "Critical Periods in Human Growth and Their Relationship to Diseases of Aging," *Yearbook of Physical Anthropology*, 45, 159–184.

CASE, A. AND C. PAXSON (2005): "Sex differences in morbidity and mortality," *Demography*, 42, 189–214.

——— (2008): "Stature and Status: Height, Ability, and Labor Market Outcomes," *Journal of Political Economy*, 116, pp. 499–532.

——— (2009): "Early Life Health and Cognitive Function in Old Age," *American Economic Review: Papers & Proceedings*, 99, 104–109.

——— (2010): "Causes and consequences of early-life health," *Demography*, 47, S65–S85.

CHEN, Y. AND L.-A. ZHOU (2007): "The long-term health and economic consequences of the 1959-1961 famine in China." *Journal of Health Economics*, 26, 659–81.

CHRISTELIS, D., T. JAPPELLI, AND M. PADULA (2010): "Cognitive abilities and portfolio choice," *European Economic Review*, 54, 18–38.

COLE, M. G. AND N. DENDUKURI (2003): "Risk factors for depression among elderly community subjects: a systematic review and meta-analysis," *American Journal of Psychiatry*, 160, 1147–1156.

COLE, S., L. HAWKLEY, J. AREVALO, AND J. CACIOPPO (2011): "Transcript origin analysis identifies antigen-presenting cells as primary targets of socially regulated gene expression in leukocytes," *Proceedings of the National Academy of Sciences*, 108, 3080–3085.

COLE, T. (2000): "Secular trends in growth," *Proceedings of the Nutrition Society*, 59, 317–324.

CORNIDES, W. (1948): *Wirtschaftsstatistik der deutschen Besatzungszonen, 1945-1948, in Verbidung mit der deutschen Produktionsstatistik der Vorkriegszeit*, vol. 3, Oberursel: Verlag Europa-Archiv.

COSTA, D. L. (1993): "Height, wealth, and disease among the native-born in the rural, antebellum North," *Social Science History*, 17, 355–383.

CRAMER, J. AND L. LUMEY (2010): "Maternal preconception diet and the sex ratio," *Human Biology*, 82, 103–107.

DATAR, A., M. KILBURN, AND D. LOUGHRAN (2010): "Endowments and parental investments in infancy and early childhood," *Demography*, 47, 145–162.

DAVIS, D. R. (1951): "Emotional disturbances and behavioural reactions," in *Medical Research Council Special Report No 275: Studies of Undernutrition, Wuppertal 1946-9*, His Majesty's Stationary Office, chap. VIII, 147–164.

DAVIS, R. G. (2006): *Bombing the European Axis Powers. A Historical Digest of the Combined Bomber Offensive 1939—1945*, Air University Press.

DE ROOIJ, S., H. WOUTERS, J. E. YONKER, R. C. PAINTER, AND T. J. ROSEBOOM (2010): "Prenatal undernutrition and cognitive function in late adulthood," *Proceedings of the National Academy of Sciences*, 107, 16881–16886.

DEAN, R. F. A. (1951): "The Size of the Baby at Birth and the Yield of Breast Milk," in *Medical Research Council Special Report Series No. 275: Studies of Undernutrition, Wuppertal 1946-9, by Members of the Department of Experimental Medicine, Cambridge, and Associated Workers*, His Majesty's Stationary Office, chap. XXVIII, 346–378.

DEATON, A. (2007): "Height, health, and development," *Proceedings of the National Academy of Sciences*, 104, 13232–13237.

DEL BONO, E., J. ERMISCH, AND M. FRANCESCONI (2008): "Intrafamily resource allocations: a dynamic model of birth weight," Institute for the Study of Labor (IZA) Discussion Paper No. 3704.

DOBLHAMMER, G., G. J. VAN DEN BERG, AND T. FRITZE (2013a): "Economic Conditions at the Time of Birth and Cognitive Abilities Late in Life: Evidence from Ten European Countries," *PloS one*, 8, e74915.

DOBLHAMMER, G., G. J. VAN DEN BERG, AND L. LUMEY (2013b): "Long-term Effects of Famine on Life Expectancy: A Re-analysis of the Great Finnish Famine of 1866-1868," *Population Studies*, 67, 309–322.

DOBLHAMMER, G. AND J. W. VAUPEL (2001): "Lifespan depends on month of birth," *Proceedings of the National Academy of Sciences*, 98, 2934–2939.

DOLS, M. J. L. AND D. J. A. M. VAN ARCKEN (1946): "Food Supply and Nutrition in the Netherlands during and Immediately after World War II," *The Milbank Memorial Fund Quarterly*, 24, 319–358.

DROESE, W. AND E. ROMINGER (1950): "Die Auswirkungen der Mangelernährung auf Körpergewicht und Körpergröße schleswigholsteinischer Kinder in der Nachkriegszeit," *Zeitschrift für Kinderheilkunde*, 67, 615–638.

ECHTERNKAMP, J. (2003): *Nach dem Krieg – Alltagsnot, Neuorientierung und die Last der Vergangenheit 1945–1949*, Pendo Verlag Zürich.

ELDER, G. (1999): *Children of the Great Depression: Social change in life experience.*, vol. 28, Westview Press.

ELO, I. AND S. PRESTON (1992): "Effects of early-life conditions on adult mortality: a review," *Population Index*, 58, 186–212.

EPEL, E. S., E. H. BLACKBURN, J. LIN, F. S. DHABHAR, N. E. ADLER, J. D. MORROW, AND R. M. CAWTHON (2004): "Accelerated telomere shortening in response to life stress," *Proceedings of the National Academy of Science*, 101, 17312–17315.

ERIKSSON, J. G., E. KAJANTIE, C. OSMOND, K. THORNBURG, AND D. J. BARKER (2010): "Boys live dangerously in the womb," *American Journal of Human Biology*, 22, 330–335.

EVELETH, P. AND J. TANNER (1990): *Worldwide variation in human growth*, Cambridge University Press.

FARQUHARSON, J. E. (1985): *The Western allies and the politics of food: agrarian management in postwar Germany*, Berg Publishers.

FICHTER, M. M., K.-M. PIRKE, AND F. HOLSBOER (1986): "Weight Loss Causes Neuroendocrine Disturbances: Experimental Study in Healthy Starving Subjects," *Psychiatry Research*, 17, 61–72.

FISKE, A., M. GATZ, AND N. L. PEDERSEN (2003): "Depressive Symptoms and Aging: The Effects of Illness and Non-Health-Related Events," *Journal of Gerontology Series B: Psychological Sciences*, 58, P320–P328.

FRIEDRICH, J. (2002): *Der Brand – Deutschland im Bombenkrieg 1940–1945*, Propyläen.

GALAZKA, A. AND S. DITTMANN (2000): "The Changing Epidemiology of Diphtheria in the Vaccine Era," *Journal of Infectious Diseases*, 181, S2–S9.

GASSER, T., P. ZIEGLER, R. H. LARGO, L. MOLINARI, AND A. PRADER (1994): "A longitudinal study of lean and fat areas at the arm." *Annals of Human Biology*, 21, 303–14.

GLAESMER, H. (2013): "Die Schatten des Zweiten Weltkrieges Folgen der traumatischen Erfahrungen in der älteren deutschen Bevölkerung," *Wissenschaft und Frieden*, 35–38.

GLAESMER, H., E. BÄHLER, H. GÜNDEL, AND S. RIEDEL-HELLER (2011): "The Association of Traumatic Experiences and Posttraumatic Stress Disorder With Physical Morbidity in Old Age: A German Population-Based Study," *Psychosomatic Medicine*, 73, 401–406.

GLUCKMAN, P. AND M. HANSON (2004): "Living with the past: evolution, development, and patterns of disease," *Science*, 305, 1733–1736.

GLUCKMAN, P., M. HANSON, AND C. PINAL (2005a): "The developmental origins of adult disease," *Maternal & Child Nutrition*, 1, 130–141.

GLUCKMAN, P. D., M. A. HANSON, C. COOPER, AND K. L. THORNBURG (2008): "Effect of in utero and early-life conditions on adult health and disease," *New England Journal of Medicine*, 359, 61–73.

GLUCKMAN, P. D., M. A. HANSON, AND H. G. SPENCER (2005b): "Predictive adaptive responses and human evolution," *Trends in Ecology & Evolution*, 20, 527–533.

GRILICHES, Z. (1979): "Sibling Models and Data in Economics: Beginnings of a Survey," *Journal of Political Economy*, 87, 37–64.

HADLEY, C. AND C. PATIL (2006): "Food insecurity in rural Tanzania is associated with maternal anxiety and depression," *American Journal of Human Biology*, 18, 359–368.

HALES, C. AND D. BARKER (1992): "Type 2 (non-insulin-dependent) diabetes mellitus: the thrifty phenotype hypothesis," *Diabetologia*, 35, 595–601.

HAMELIN, A., J. HABICHT, AND M. BEAUDRY (1999): "Food insecurity: consequences for the household and broader social implications," *The Journal of Nutrition*, 129, 525.

HARRIS, S. E., I. J. DEARY, A. MACINTYRE, K. J. LAMBB, K. RADHAKRISHNAN, J. M. STARR, L. J. WHALLEY, AND P. G. SHIELS (2006): "The association between telomere length, physical health, cognitive ageing, and mortality in non-demented older people," *Neuroscience Letters*, 406, 260–264.

HARVEY, A. G. AND R. A. BRYANT (1999): "Acute Stress Disorder across Trauma Populations," *The Journal of Nervous and Mental Disease*, 187(7), 443–446.

HARVILLE, E., X. XIONG, AND P. BUEKENS (2010): "Disasters and perinatal health: a systematic review," *Obstetrical & Gynecological Survey*, 65, 713.

HAUPTAMT FÜR STATISTIK BERLIN-WILMERSDORF (1947): *Berlin in Zahlen 1947*, Berliner Kulturbuch-Verlag G.m.b.H.

HÄUSSER, A. AND G. MAUGG (2009): *Hungerwinter. Deutschlands humanitäre Katastrophe 1946/47*, Propyläen.

HAVARI, E. AND F. PERACCHI (2012): "Childhood circumstances and adult outcomes: Evidence from SHARELIFE," EIEF Working Paper 1115.

HEIJMANS, B. T., E. W. TOBI, A. D. STEIN, H. PUTTER, G. J. BLAUW, E. S. SUSSER, P. E. SLAGBOOM, AND L. H. LUMEY (2008): "Persistent epigenetic differences associated with prenatal exposure to famine in humans," *Proceedings of the National Academy of Sciences*, 105, 17046–17049.

HEINEMAN, E. (1996): "The Hour of the Woman: Memories of Germany's "Crisis Years" and West German National Identity," *The American Historical Review*, 101, 354–395.

HERMANUSSEN, M., H. DANKER-HOPFE, AND G. WEBER (2001): "Body weight and the shape of the natural distribution of weight, in very large samples of German, Austrian and Norwegian conscripts," *International Journal of Obesity*, 25, 1550–1553.

HEWITT, K. (1983): "Place Annihilation: Area Bombing and the Fate of Urban Places," *Annals of the Association of American Geographers*, 73, 257–284.

HIONIDOU, V. (2006): *Famine and Death in Occupied Greece, 1941-1944*, Cambridge University Press.

HODRICK, R. J. AND E. C. PRESCOTT (1997): "Postwar U.S. Business Cycles: An Empirical Investigation," *Journal of Money, Credit and Banking*, 29, 1–16.

HUXLEY, R., C. OWEN, P. WHINCUP, D. COOK, J. RICH-EDWARDS, G. SMITH, AND R. COLLINS (2007): "Is birth weight a risk factor for ischemic heart disease in later life?" *The American Journal of Clinical Nutrition*, 85, 1244.

ICHINO, A. AND R. WINTER-EBMER (2004): "The long-run educational cost of World War II," *Journal of Labor Economics*, 22, 57–87.

IMBENS, G. AND J. ANGRIST (1994): "Identification and estimation of local average treatment effects," *Econometrica*, 62, 467–475.

INOUE, A. AND G. SOLON (2010): "Two-Sample Instrumental Variables Estimators," *The Review of Economics and Statistics*, 92, 557–561.

ISSERLIN, L., G. ZERACH, AND Z. SOLOMON (2008): "Acute Stress Responses: A Review and Synthesis of ASD, ASR, and CSR," *American Journal of Orthopsychiatry*, 78(4), 423–429.

IVANOVIC, D. M., B. P. LEIVA, H. T. PEREZ, M. G. OLIVARES, N. S. DIAZ, M. S. C. URRUTIA, A. F. ALMAGIÀ, T. D. TORO, P. T. MILLER, E. O. BOSCH, AND C. G. LARRAIN (2004): "Head size and intelligence, learning, nutritional status and brain development – Head, IQ, learning, nutrition and brain," *Neuropsychologia*, 42, 1118–1131.

JÜRGES, H. (2013): "Collateral damage: the German food crisis, educational attainment and labor market outcomes of German post-war cohorts," *Journal of Health Economics*, 32, 286 – 303.

JYOTI, D., E. FRONGILLO, AND S. JONES (2005): "Food insecurity affects school children's academic performance, weight gain, and social skills," *The Journal of Nutrition*, 135, 2831.

KAATI, G., L. O. BYGREN, AND S. EDVINSSON (2002): "Cardiovascular and diabetes mortality determined by nutrition during parents' and grandparents' slow growth period." *European Journal of Human Genetics*, 10.

KANANEN, L., I. SURAKKA, S. PIRKOLA, J. SUVISAARI, J. LÖNNQVIST, L. PELTONEN, S. RIPATTI, AND I. HOVATTA (2010): "Childhood Adversities Are Associated with Shorter Telomere Length at Adult Age both in Individuals with an Anxiety Disorder and Controls," *PLoS ONE*, 5, 1–7.

KAREHOLT, I. (2001): "The long shadow of socioeconomic conditions in childhood: do they affect class inequalities in mortality?" in *Cradle to Grave: Life-course Change in Modern Sweden*, ed. by J. O. Jonsson and C. Mills, Sociology Press, Durham.

KESTERNICH, I., B. SIFLINGER, J. P. SMITH, AND J. K. WINTER (2013): "Individual Behavior as a Pathway between Early-Life Shocks and Adult Health: Evidence from Hunger Episodes in Post-War Germany," Institute for the Study of Labor (IZA) Discussion Paper No. 7713.

——— (2014): "The effects of World War II on economic and health outcomes across Europe," *Review of Economics and Statistics*, 96, 103–118.

KHASHAN, A. S., K. M. ABEL, R. MCNAMEE, M. G. PEDERSEN, R. T. WEBB, P. N. BAKER, L. C. KENNY, AND P. B. MORTENSEN (2008): "Higher risk of offspring schizophrenia following antenatal maternal exposure to severe adverse life events," *Archives of general psychiatry*, 65, 146–152.

KIRBY, M. AND R. CAPEY (1997): "The area bombing of Germany in World War II: An operational research perspective," *Journal of the Operational Research Society*, 48, 661–677.

KLATT, W. (1950): "Food and Farming in Germany: I. Food and Nutrition," *International Affairs*, 26, 45–58.

KNUDSEN, E. I., J. J. HECKMAN, J. L. CAMERON, AND J. P. SHONKOFF (2006): "Economic, neurobiological, and behavioral perspectives on building America's future workforce," *Proceedings of the National Academy of Science*, 103, 10155–10162.

KOCK, G. (1997): *Der Führer sorgt für unsere Kinder–: die Kinderlandverschickung im Zweiten Weltkrieg*, F. Schöningh.

KOUPIL, I., S. PLAVINSKAJA, N. PARFENOVA, D. SHESTOV, P. DANZIGER, AND D. VAGERÖ (2009): "Cancer mortality in women and men who survived the siege of Leningrad," *International Journal of Cancer*, 124, 1416–1421.

KRABBENDAM, L., L. SMITS, R. DE BIE, J. BASTIAANSSEN, F. STELMA, AND J. VAN OS (2005): "The impact of maternal stress on pregnancy outcome in a well-educated Caucasian population," *Paediatric and Perinatal Epidemiology*, 19, 421–425.

KRAEMER, S. (2000): "The fragile male," *Clinical Medicine NetPrints*, 1.

KROH, M. (2013): "Documentation of sample sizes and panel attrition in the German Socio Economic Panel (SOEP)(1984 until 2012)," Tech. rep., Data Documentation, DIW.

KUH, D., C. POWER, AND B. RODGERS (1991): "Secular trends in social class and sex differences in adult height," *International Journal of Epidemiology*, 20, 1001–1009.

KUZAWA, C. AND E. QUINN (2009): "Developmental origins of adult function and health: evolutionary hypotheses," *Annual Review of Anthropology*, 38, 131–147.

KUZAWA, C. W. (2012): "Early Environments, Developmental Plasticity and Chronic Degenerative Disease," in *Human Growth and Development Second Edition*, Elsevier, 325–341.

LAWLOR, D. A. (2008): "The developmental origins of health and disease: where do we go from here?" *Epidemiology*, 19, 206–208.

LAWLOR, D. A., Y. BEN-SHLOMO, AND D. A. LEON (2004): "Pre-adult influences on cardiovascular disease," in *A life course approach to chronic disease epidemiology*, Oxford University Press.

LEIGH, J. P. (1983): "Direct and Indirect Effects of Education on Health," *Social Science and Medicine*, 17, 227–234.

LI, Y., Y. HE, L. QI, V. W. JADDOE, E. J. FESKENS, X. YANG, G. MA, AND F. B. HU (2010): "Exposure to the Chinese famine in early life and the risk of hyperglycemia and type 2 diabetes in adulthood," *Diabetes*, 59, 2400–2406.

LIN, N. AND W. M. ENSEL (1989): "Life Stress and Health: Stressors and Resources," *American Sociological Review*, 54, 382–399.

LINDEBOOM, M., F. PORTRAIT, AND G. J. VAN DEN BERG (2002): "An Econometric Analysis of the Mental-Health Effects of Major Events in the life of Older Individuals," *Health Economics*, 11, 505–520.

——— (2010): "Long-run effects on longevity of a nutritional shock early in life: The Dutch Potato famine of 1846-1847," *Journal of Health Economics*, 29, 617–629.

LIPPS, O. (2009): "Attrition of Households and Individuals in Panel Surveys," Tech. rep.

LOW, B. S. (2001): *Why sex matters: A Darwinian look at human behavior*, Princeton University Press.

LUMEY, L., A. STEIN, H. KAHN, K. VAN DER PAL-DE BRUIN, ET AL. (2007): "Cohort profile: the Dutch Hunger Winter families study," *International Journal of Epidemiology*, 36, 1196.

LUMEY, L., A. STEIN, AND E. SUSSER (2011): "Prenatal famine and adult health," *Annual Review of Public Health*, 32, 237–262.

LUPIEN, S. J., B. S. MCEWEN, M. R. GUNNAR, AND C. HEIM (2009): "Effects of stress throughout the lifespan on the brain, behaviour and cognition," *Nature*, 10, 434–445.

LYONS, D. M., K. J. PARKER, AND A. F. SCHATZBERG (2010): "Animal models of early life stress: implications for understanding resilience," *Developmental psychobiology*, 52, 616–624.

MADDISON, A. (2010): "Historical Statistics of the World Economy: 1–2008 AD," University of Groningen, available at www.ggdc.net/MADDISON/Historical_Statistics/vertical-file_02-2010.xls.

MALASPINA, D., C. CORCORAN, K. KLEINHAUS, M. PERRIN, S. FENNIG, D. NAHON, Y. FRIEDLANDER, AND S. HARLAP (2008): "Acute maternal stress in pregnancy and schizophrenia in offspring: a cohort prospective study," *BMC psychiatry*, 8, 71.

MANZ, M. (1985): "Stagnation und Aufschwung in der französischen Besatzungszone 1945-1948," Dissertation, Universität Mannheim.

MARIC, N. P., B. DUNJIC, D. J. STOJILJKOVIC, D. BRITVIC, AND M. JASOVIC-GASIC (2010): "Prenatal stress during the 1999 bombing associated with lower birth weight—a study of 3,815 births from Belgrade," *Archives of Women's Mental Health*, 13, 83–89.

MARSHALL, W. AND J. TANNER (1986): "Puberty," in *Human growth. A comprehensive treatise. Part 2: Postnatal growth Neurobiology*, ed. by F. Falkner and J. Tanner, Plenum Press.

MARTIN-RUIZ, C., H. O. DICKINSON, B. KEYS, E. ROWAN, R. A. KENNY, AND T. VON ZGLINICKI (2006): "Telomere Length Predicts Poststroke Mortality, Dementia, and Cognitive Decline," *Annals of Neurology*, 60, 174–180.

MARTINI, P. (1947): "Über die gesundheitliche Lage in Deutschland," in *Die soziale und gesundheitliche Lage in Deutschland - Vorträge und Entschliessung aus der Tagung des Hauptausschusses des Deutschen Städtetages zu Goslar am 18. Juli 1947*, München: Urban und Schwarzenberg, vol. 8, 28–47.

MERRITT, A. J. AND R. L. MERRITT (1970): *Public Opinion in Occupied Germany*, Urbana, University of Illinois Press.

MIDDLEBROOK, M., ed. (1985): *The Bomber Command war diaries : an operational reference book, 1939 - 1945*, Viking.

MILLER, G. AND E. CHEN (2007): "Unfavorable Socioeconomic Conditions in Early Life Presage Expression of Proinflammatory Phenotype in Adolescence," *Psychosomatic Medicine*, 69, 402–409.

——— (2010): "Harsh family climate in early life presages the emergence of a proinflammatory phenotype in adolescence," *Psychological Science*, 21, 848.

MILLER, G., E. CHEN, A. FOK, H. WALKER, A. LIM, E. NICHOLLS, S. COLE, AND M. KOBOR (2009): "Low early-life social class leaves a biological residue manifested by decreased glucocorticoid and increased proinflammatory signaling," *Proceedings of the National Academy of Sciences*, 106, 14716–14721.

MOEBUS, S., J. HANISCH, P. BRAMLAGE, C. LÖSCH, H. HAUNER, J. WASEM, AND K.-H. JÖCKEL (2008): "Regional unterschiedliche Prävalenz des metabolischen Syndroms," *Deutsches Ärzteblatt*, 105, 207–13.

MOROZINK, J., E. FRIEDMAN, C. COE, AND C. RYFF (2010): "Socioeconomic and psychosocial predictors of interleukin-6 in the MIDUS national sample." *Health Psychology*, 29, 626.

MU, R. AND X. ZHANG (2011): "Why does the Great Chinese Famine affect the male and female survivors differently? Mortality selection versus son preference," *Economics & Human Biology*, 9, 92–105.

NEELSEN, S. AND T. STRATMANN (2011): "Effects of prenatal and early life malnutrition: Evidence from the Greek famine," *Journal of Health Economics*, 30, 479–488.

NEUGEBAUER, R., H. W. HOEK, AND E. SUSSER (1999): "Prenatal Exposure to Wartime Famine and Development of Antisocial Personality Disorder in Early Adulthood," *Journal of the American Medical Association*, 281, 455–462.

O'GRÁDA, C. (2009): *Famine*, Princeton University Press.

OJEDA, N. B., D. GRIGORE, AND B. T. ALEXANDER (2008): "Developmental programming of hypertension insight from animal models of nutritional manipulation," *Hypertension*, 52, 44–50.

OPPERS, V. M. (1963): "Analyse van de acceleratie van de menselijke lengtegroei door bepaling van het tijdstip van de groeifasen," Ph.D. thesis.

PAINTER, R., T. ROSEBOOM, AND O. BLEKER (2005): "Prenatal exposure to the Dutch famine and disease in later life: an overview," *Reproductive Toxicology*, 20, 345–352.

PARKES, C. M. (1992): "Bereavement and mental health in the elderly," *Reviews in Clinical Gerontology*, 2, 5.

PECHTEL, P. AND D. A. PIZZAGALLI (2011): "Effects of early life stress on cognitive and affective function: an integrated review of human literature," *Psychopharmacology*, 214, 55–70.

PINGER, P. R. (2013): *Essays on Skills, Health and Human Inequality*, Dissertation, University of Mannheim.

POLLITT, R., K. ROSE, AND J. KAUFMAN (2005): "Evaluating the evidence for models of life course socioeconomic factors and cardiovascular outcomes: a systematic review," *BMC public health*, 5, 7.

POULTER, N., C. CHANG, A. MACGREGOR, H. SNIEDER, AND T. SPECTOR (1999): "Association between birth weight and adult blood pressure in twins: historical cohort study," *British Medical Journal*, 319, 1330.

PRINCE, M., F. REISCHIES, A. BEEKMAN, R. FUHRER, C. JONKER, S. KIVELA, B. LAWLOR, A. LOBO, H. MAGNUSSON, M. FICHTER, H. VAN OYEN, M. ROELANDS, I. SKOOG, C. TURRINA, AND J. COPELAND (1999): "Development of the EURO-D scale–a European, Union initiative to compare symptoms of depression in 14 European centres," *The British Journal of Psychiatry*, 174, 330–338.

RASMUSSEN, K. (2001): "The fetal origins hypothesis: challenges and opportunities for maternal and child nutrition," *Annual Review of Nutrition*, 21, 73–95.

REICHARDT, S. AND M. ZIERENBERG (2010): *Damals nach dem Krieg: eine Geschichte Deutschlands 1945 bis 1949*, München: Goldmann.

REPORT OF THE MILITARY GOVERNOR NO. 39 (1948): *Statistical Annex, Issue No. XIX - Report of the Military Governor*, Office of Military Government for Germany.

REPORT OF THE MILITARY GOVERNOR NO. 49 (1949): *Statistical Annex, Issue No. XXIX*, Office of Military Government for Germany.

RIDDER, G. AND R. MOFFITT (2007): "The Econometrics of Data Combination," in *Handbook of Econometrics*, ed. by J. Heckman and E. Leamer, Amsterdam: Elsevier, vol. 6B, 5469–5547.

RISEMAN, J. E. AND S. WEISS (1929): "The age and sex incidence of arterial hypertension," *American Heart Journal*, 5, 172–190.

ROSE, D. AND V. OLIVEIRA (1997): "Nutrient intakes of individuals from food-insufficient households in the United States." *American Journal of Public Health*, 87, 1956–1961.

ROSEBOOM, T. J., J. H. VAN DER MEULEN, A. C. RAVELLI, C. OSMOND, D. J. BARKER, AND O. P. 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.

ROSENZWEIG, M. AND K. WOLPIN (1988): "Heterogeneity, intrafamily distribution, and child health," *Journal of Human Resources*, 23, 437–461.

ROSNICK, C. B., B. J. SMALL, C. L. MCEVOY, A. R. BORENSTEIN, AND J. A. MORTIMER (2007): "Negative Life Events and Cognitive Performance in a Population of Older Adults," *Journal of Aging and Health*, 19, 612–629.

ROTHENBERGER, K.-H. (1980): *Die Hungerjahre nach dem Zweiten Weltkrieg: Ernährungs- und Landwirtschaft in Rheinland-Pfalz 1945-1950*, Harald Boldt Verlag, Boppard am Rhein.

ROYER, H. AND A. WITMAN (2013): "Intergenerational Effects on Health – In Utero and Early Life," *Working Paper*.

SANZ-DE GALDEANO, A. (2005): "The obesity epidemic in Europe," Institute for the Study of Labor (IZA) Discussion Paper No. 1814.

SCHAFFER, R. (1980): "American Military Ethics in World War II: The Bombing of German Civilians," *The Journal of American History*, 67, 318–334.

SCHLANGE-SCHÖNINGEN, H. (1955): *Im Schatten des Hungers - Dokumentarisches zur Ernährungspolitik und Ernährungswirtschaft in den Jahren 1945-1949*, Verlag Paul Parey, Hamburg und Berlin.

SCHMIDT, I. M., M. JØRGENSEN, AND K. F. MICHAELSEN (1995): "Height of conscripts in Europe: is postneonatal mortality a predictor?" *Annals of Human Biology*, 22, 57–67.

SCHMITZ, H. (1956): *Die Bewirtschaftung der Nahrungsmittel und Verbrauchsgüter 1939-1950: Dargestellt an dem Beispiel der Stadt Essen*, Stadtverwaltung Essen.

SCHOLTE, R., G. J. VAN DEN BERG, M. LINDEBOOM, AND D. J. DEEG (2014): "Does the Size of the Effect of Adverse Events at High Ages on Daily-Life Physical Functioning Depend on the Economic Conditions Around Birth?" Institute for the Study of Labor (IZA) Discussion Paper No. 8075.

SCHOLTE, R. S., G. J. VAN DEN BERG, AND M. LINDEBOOM (2012): "Long-Run Effects of Gestation During the Dutch Hunger Winter Famine on Labor Market and Hospitalization Outcomes," Institute for the Study of Labor (IZA) Discussion Paper No. 6307.

SCHRÖDER, M. (2011): *Retrospective Data Collection in the Survey of Health, Ageing and Retirement in Europe – SHARELIFE Methodology*, Mannheim Research Institute for the Economics of Ageing (MEA).

SCHULZ, L. (2010): "The Dutch Hunger Winter and the developmental origins of health and disease," *Proceedings of the National Academy of Sciences*, 107, 16757–16758.

SEEL, G. (1950): "Ernährung und Ernährungszustand der Bevölkerung der britischen Besatzungszone Deutschlands in den Jahren 1947-1948," Ph.D. thesis.

SIDEL, V. W. (1997): "The public health impact of hunger." *American Journal of Public Health*, 87, 1921–1922.

SPARÉN, P., D. VÅGERÖ, D. SHESTOV, S. PLAVINSKAJA, N. PARFENOVA, V. HOPTIAR, D. PATUROT, AND M. GALANTI (2004): "Long term mortality after severe starvation during the siege of Leningrad: prospective cohort study," *British Medical Journal*, 328, 11.

STAUB, K., F. RÜHLI, U. WOITEK, AND C. PFISTER (2011): "The average height of 18-and 19-year-old conscripts (N= 458,322) in Switzerland from 1992 to 2009, and the secular height trend since 1878," *Swiss Medical Weekly*, 144.

STECKEL, R. H. (2008): "Biological measures of the standard of living." *The Journal of Economic Perspectives*, 22, 129–152.

STEIN, A., F. PIERIK, G. VERRIPS, E. SUSSER, AND L. LUMEY (2009): "Maternal exposure to the Dutch famine before conception and during pregnancy: quality of life and depressive symptoms in adult offspring," *Epidemiology*, 20, 909–915.

STEIN, Z. AND M. SUSSER (1975): "Fertility, fecundity, famine: food rations in the Dutch famine 1944/5 have a causal relation to fertility, and probably to fecundity," *Human Biology*, 131–154.

STERN, Y. (2002): "What is cognitive reserve? Theory and research application of the reserve concept," *Journal of the International Neuropsychological Society*, 8, 448–460.

——— (2009): "Cognitive reserve," *Neuropsychologia*, 47, 2015–2028.

SÜSS, W. (2003): *Der "Volkskörper" im Krieg - Gesundheitspolitik, Gesundheitsverhältnisse und Krankenmord im national-sozialistischen Deutschland 1939-1945*, Institut für Zeitgeschichte, Oldenbourg Verlag, Munich.

SUSSER, E. AND S. LIN (1992): "Schizophrenia after prenatal exposure to the Dutch Hunger Winter of 1944-1945," *Archives of General Psychiatry*, 49, 983.

SUSSER, E., R. NEUGEBAUER, H. HOEK, A. BROWN, S. LIN, D. LABOVITZ, AND J. GORMAN (1996): "Schizophrenia after prenatal famine. Further evidence." *Archives of General Psychiatry*, 53, 25–31.

SUSSER, E. AND D. ST CLAIR (2013): "Prenatal famine and adult mental illness: Interpreting concordant and discordant results from the Dutch and Chinese Famines." *Social Science & Medicine*.

SUSSER, M. AND Z. STEIN (1994): "Timing in prenatal nutrition: a reprise of the Dutch Famine Study," *Nutrition Reviews*, 52, 84–94.

TALGE, N. M., C. NEAL, AND V. GLOVER (2007): "Antenatal maternal stress and long-term effects on child neurodevelopment: how and why?" *Journal of Child Psychology and Psychiatry*, 48, 245–261.

THE UNITED STATES STRATEGIC BOMBING SURVEY (1947): *The effect of bombing on health and medical care in Germany*.

TONKISS, J., M. TRZCIŃSKA, J. R. GALLER, N. RUIZ-OPAZO, AND V. L. HERRERA (1998): "Prenatal malnutrition-induced changes in blood pressure dissociation of stress and nonstress responses using radiotelemetry," *Hypertension*, 32, 108–114.

TRITTEL, G. J. (1990): *Hunger und Politik – Die Ernährungskrise in der Bizone (1945–1949)*, Campus Verlag.

TURVEY, C. L., C. CARNEY, S. ARNDT, R. B. WALLACE, AND R. HERZOG (1999): "Conjugal loss and syndromal depression in a sample of elders aged 70 years or older," *American Journal of Psychiatry*, 156, 1596–1601.

TYRKA, A. R., L. H. PRICE, H.-T. KAO, B. PORTON, S. A. MARSELLA, AND L. L. CARPENTER (2010): "Childhood Maltreatment and Telomere Shortening: Preliminary Support for an Effect of Early Stress on Cellular Aging," *Biological Psychiatry*, 67, 531–534.

VALAORAS, V. (1946): "Some effects of famine on the population of Greece," *The Milbank Memorial Fund Quarterly*, 24, 215–234.

VAN ABEELEN, A. F., S. G. ELIAS, P. M. BOSSUYT, D. E. GROBBEE, Y. T. VAN DER SCHOUW, T. J. ROSEBOOM, AND C. S. UITERWAAL (2012): "Famine exposure in the young and the risk of type 2 diabetes in adulthood," *Diabetes*, 61, 2255–2260.

VAN DEN BERG, G. J., D. J. H. DEEG, M. LINDEBOOM, AND F. PORTRAIT (2010): "The Role of Early-Life Conditions in the Cognitive Decline due to Adverse Events Later in Life," *The Economic Journal*, 120, F411–F428.

VAN DEN BERG, G. J., G. DOBLHAMMER, AND K. CHRISTENSEN (2009a): "Exogenous determinants of early-life conditions, and mortality later in life," *Social Science & Medicine*, 68, 1591–1598.

VAN DEN BERG, G. J. AND M. LINDEBOOM (2014): "Aging: conditions early in life and health and mortality at advanced ages," in *Encyclopedia of Health Economics*, ed. by A. Culyer, Elsevier.

VAN DEN BERG, G. J., M. LINDEBOOM, AND M. LOPEZ (2009b): "Inequality in individual mortality and economic conditions earlier in life," *Social Science & Medicine*, 69, 1360–1367.

VAN DEN BERG, G. J., M. LINDEBOOM, AND F. PORTRAIT (2006): "Economic Conditions Early in Life and Individual Mortality," *American Economic Review*, 96, 290–302.

VAN DEN BERG, G. J., P. LUNDBORG, B. NYSTEDT, AND D.-O. ROOTH (forthcoming): "Critical periods during childhood and adolescence: a study of adult height among immigrant siblings," *Journal of the European Economic Association*.

VAN DEN BERG, G. J. AND P. PINGER (2014): "A validation study of transgenerational effects of childhood conditions on the third generation offspring's economic and health outcomes potentially driven by epigenetic imprinting," Institute for the Study of Labor (IZA) Discussion Paper No. 7999.

VAN DEN BERG, G. J., P. PINGER, AND J. SCHOCH (2011): "Instrumental variable estimation of the causal effect of hunger early in life on health later in life," Institute for the Study of Labor (IZA) Discussion Paper No. 6110.

VAN OS, J. AND J.-P. SELTEN (1998): "Prenatal exposure to maternal stress and subsequent schizophrenia. The May 1940 invasion of The Netherlands." *The British Journal of Psychiatry*, 172, 324–326.

VONYÓ, T. (2012): "The bombing of Germany: the economic geography of war-induced dislocation in West German industry," *European Review of Economic History*, 16, 97–118.

WALDRON, I. (1983): "The role of genetic and biological factors in sex differences in mortality," in *Sex differentials in mortality: trends, determinants and consequences*, ed. by A. Lopez and L. Ruzicka, Australian National University, Canberra.

WARD, L., J. L. MATHIAS, AND S. E. HITCHINGS (2007): "Relationships between bereavement and cognitive functioning in older adults," *Gerontology*, 53, 362–372.

WELBERG, L. A. M. AND J. R. SECKL (2001): "Prenatal Stress, Glucocorticoids and the Programming of the Brain," *Journal of Neuroendocrinology*, 13, 113–128.

WERRELL, K. P. (1986): "The strategic Bombing of Germany in World War II: Costs and Accomplishments," *The Journal of American History*, 73, 702–713.

WHITAKER, R., S. PHILLIPS, AND S. ORZOL (2006): "Food insecurity and the risks of depression and anxiety in mothers and behavior problems in their preschool-aged children," *Pediatrics*, 118, 859.

WILLENBORG, K.-H. (1979): "Überleben nach dem Zusammenbruch – Die Not der Besiegten," in *Geschichte der Bundesrepublik Deutschland, Band1: Auf dem Wege zur Republik 1945–1947*, ed. by J. Weber, Schöningh, chap. 6, 187–210.

ZEITLER, R. AND B. MEWES, eds. (1941): *Statistisches Jahrbuch Deutscher Gemeinden*, vol. 36, Gustav Fischer Jena.

ZEMEL, B. (2002): "Body Composition During Growth and Development," in *Body Composition During Growth and Development*, 271–293.

ZHANG, T., R. BAGOT, C. PARENT, C. NESBITT, T. BREDY, C. CALDJI, E. FISH, H. ANISMAN, M. SZYF, AND M. MEANEY (2006): "Maternal programming of defensive responses through sustained effects on gene expression," *Biological Psychology*, 73, 72–89.

Curriculum Vitae

2011 – 2014	Doctoral Studies in Economics, Graduate School for Economic and Social Sciences, University of Mannheim, Germany
2009 – 2011	Master of Science in Economic Research, University of Mannheim, Germany
2006 – 2009	Bachelor of Science in Economics, University of Mannheim, Germany
2006	Abitur, Isolde-Kurz-Gymnasium, Reutlingen, Germany