

van den Berg, Gerard J.; Pinger, Pia R.; Schoch, Johannes

Working Paper

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

SOEPpapers on Multidisciplinary Panel Data Research, No. 710

Provided in Cooperation with:

German Institute for Economic Research (DIW Berlin)

Suggested Citation: van den Berg, Gerard J.; Pinger, Pia R.; Schoch, Johannes (2014) : Instrumental Variable Estimation of the Causal Effect of Hunger Early in Life on Health Later in Life, SOEPpapers on Multidisciplinary Panel Data Research, No. 710, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

This Version is available at:

<https://hdl.handle.net/10419/106190>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

SOEPpapers

on Multidisciplinary Panel Data Research

SOEP – The German Socio-Economic Panel Study at DIW Berlin

710-2014

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

Gerard J. van den Berg, Pia R. Pinger and Johannes Schoch

SOEPPapers on Multidisciplinary Panel Data Research at DIW Berlin

This series presents research findings based either directly on data from the German Socio-Economic Panel Study (SOEP) or using SOEP data as part of an internationally comparable data set (e.g. CNEF, ECHP, LIS, LWS, CHER/PACO). SOEP is a truly multidisciplinary household panel study covering a wide range of social and behavioral sciences: economics, sociology, psychology, survey methodology, econometrics and applied statistics, educational science, political science, public health, behavioral genetics, demography, geography, and sport science.

The decision to publish a submission in SOEPPapers is made by a board of editors chosen by the DIW Berlin to represent the wide range of disciplines covered by SOEP. There is no external referee process and papers are either accepted or rejected without revision. Papers appear in this series as works in progress and may also appear elsewhere. They often represent preliminary studies and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be requested from the author directly.

Any opinions expressed in this series are those of the author(s) and not those of DIW Berlin. Research disseminated by DIW Berlin may include views on public policy issues, but the institute itself takes no institutional policy positions.

The SOEPPapers are available at
<http://www.diw.de/soeppapers>

Editors:

Jürgen **Schupp** (Sociology)

Gert G. **Wagner** (Social Sciences, Vice Dean DIW Graduate Center)

Conchita **D'Ambrosio** (Public Economics)

Denis **Gerstorff** (Psychology, DIW Research Director)

Elke **Holst** (Gender Studies, DIW Research Director)

Frauke **Kreuter** (Survey Methodology, DIW Research Professor)

Martin **Kroh** (Political Science and Survey Methodology)

Frieder R. **Lang** (Psychology, DIW Research Professor)

Henning **Lohmann** (Sociology, DIW Research Professor)

Jörg-Peter **Schräpler** (Survey Methodology, DIW Research Professor)

Thomas **Siedler** (Empirical Economics)

C. Katharina **Spieß** (Empirical Economics and Educational Science)

ISSN: 1864-6689 (online)

German Socio-Economic Panel Study (SOEP)
DIW Berlin
Mohrenstrasse 58
10117 Berlin, Germany

Contact: Uta Rahmann | soeppapers@diw.de

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

Gerard J. van den Berg*

Pia R. Pinger[†]

Johannes Schoch[‡]

Abstract

We estimate average causal effects of early-life hunger on late-life health by applying instrumental variable estimation, using data with self-reported periods of hunger earlier in life, with famines as instruments. The data contain samples from European countries and include birth cohorts exposed to various famines in the 20th century. We use two-sample IV estimation to deal with imperfect recollection of conditions at very early stages of life. The estimated average causal effects may exceed famine effects by at least a factor three.

Keywords: Nutrition, famine, ageing, developmental origins, height, two-sample IV.

JEL codes: I12, J11, C21, C26.

*Alexander von Humboldt Professor of Econometrics and Empirical Economics, University of Mannheim; IFAU-Uppsala, VU University Amsterdam; IZA. Address: Department of Economics, University of Mannheim, L7, 3–5, 68131 Mannheim, Germany.

[†]University of Bonn, IZA.

[‡]University of Mannheim.

Keywords: Nutrition, famine, ageing, developmental origins, height, two-sample IV.

JEL codes: I12, J11, C21, C26.

Acknowledgements: We thank Alberto Abadie, Bo Honoré, Jim Heckman, participants in a workshop in Munich, Sept 2011, ESPE 2012, ESEM 2012, a seminar in Mannheim, June 2012; a workshop at Georgetown University, Oct 2012, a seminar in Stockholm, Nov 2013, and a workshop in Amsterdam, Nov 2013, for useful comments. This paper 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 www.share-project.org for the full list of funding institutions.

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.¹ 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 birth weight, 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.

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 favourable 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 birth weight and even gene expression.² A few studies

¹For overviews, see e.g. Pollitt *et al.* (2005); Barker (2007); Lawlor (2008) and Almond and Currie (2011).

²Results 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 1 for overviews.

have also reported associations between exposure to a famine at the onset of puberty and late-life health outcomes.³

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.⁴

In this paper 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

³For 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. Van den Berg *et al.* (2014) 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.

⁴If the study outcomes concern health at high ages then an additional problem 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.

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.⁵ This paper 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.

⁵Reduced-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).

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 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.⁶ Their origins are well-established in the literature, and they have been used to study long-run effects (see the literature discussion in Section 1). 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

⁶Barring 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.

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 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 paper proceeds as follows. In Section 1 we review the explanatory frameworks to understand the long-run effects. Section 2 describes the three famines in our observation window and summarizes the evidence obtained so far for those famines. In Section 3 we describe our data. Section 4 formally presents the econometric methods. Here we also examine selectivity issues associated with the famines. Section 5 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 5 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 5 concludes.

1. Explanatory Frameworks for Causal Long-run Effects of Conditions Early in Life

1.1. *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 (Zhang *et al.*, 2006; Miller *et al.*, 2009; Morozink *et al.*, 2010; Miller and Chen, 2010; 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 favourable 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 behaviours 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).

1.2. *Conditions in Later Childhood*

Recently, interest has increased in long-run effects of nutritional conditions beyond infancy. Gluckman *et al.* (2005) 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 behaviour. 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.

1.3. *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, and Van den Berg *et al.*, 2009). For over-all mortality this is echoed by the studies of the Finnish famine

from 1866-1868 (Doblhammer *et al.*, 2013) 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 (Eveleth and Tanner, 1990; Kuh *et al.*, 1991). 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 birth weight, 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).

2. The Famines in European Countries in the 1940s

2.1. *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 1944.⁷

The excess death rate in the first half of 1945 over the rate in 1944 amounts to 269% for men and 173% 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.⁸ 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

⁷For 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).

⁸Recall 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 the introduction.

or late gestational exposure to the famine. Susser and Stein (1994) find that adult stature is susceptible to the postnatal but not the prenatal environment. As noted in the introduction, 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 2.1 above.

2.2. *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.⁹ 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% of the wheat consumption had been imported). In addition, the blockade prohibited fishing at sea.¹⁰ 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

⁹This subsection relies heavily on Hionidou (2006) and Valaoras (1946).

¹⁰In 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’Grada (2009).

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. 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 labour market outcomes. Effects of exposure in utero are very small.

2.3. *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 (2009) 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 equalled 1330, 1083, 1050, and 900 kcal, respectively (see Cornides, 1948, and Echterkamp, 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% 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) analyses 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 (2011) analyses the effect of birth during the German famine on educational attainment and occupation in the labour 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 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”.

3. 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.¹¹

3.1. *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.¹² 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.¹³ Finally, we trim

¹¹For general information on SHARE and SHARELIFE, see Börsch-Supan *et al.* (2008) and Schröder (2011).

¹²We drop complete proxy interviews but keep assisted interview information. This only concerns the interviews yielding the outcome measures and the undernutrition indicator.

¹³Conversely, 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

the upper and lower 2.5% tails of each country’s gender specific height distribution. Our working sample consists of about 2500 men and 2800 women.

3.2. *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 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.¹⁴ 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 2 are:¹⁵

1. Greece: May 1941 to end of June 1942

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.

¹⁴Also, only one interval period of hunger can be reported.

¹⁵Changing 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.

2. The Netherlands: November 1944 to end of April 1945, only the Western part¹⁶
3. Germany: June 1945 to end of June 1948

Figure 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 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 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.

¹⁶To 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.

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 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.¹⁷

Table 1 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 6-16.¹⁸ Moreover, in

¹⁷Figure 5 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.

¹⁸If fertility and mortality are selective, as discussed in Subsection 4.3, 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.

line with the historical evidence described in Section 2, this fraction of hunger-exposed 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%.

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 (SOEP) 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 3000.¹⁹ 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 SOEP 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.

¹⁹The module will become publicly available in December 2014.

If recall errors are a major problem then we expect the SOEP data to be less informative about the effect of the famine on hunger than the SHARE data are.

Figure 4 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 1 in our paper is essentially replicated in the SOEP 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.

Notice that the SOEP module was collected about 5 years later than the SHARELIFE wave, and that the SOEP 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 6 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 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.²⁰ This provides a rationale to exclude the other countries from the baseline analyses. We return to this in Section 5 when we discuss sensitivity analyses.

Figure 7 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 3 shows that a fraction of approximately 15% of those having experienced a famine before age 16, report hunger for the same period. For non-famine periods, this is only 3.5%. This confirms that famines are an instrument for periods of undernutrition.

3.3. *Adult Height*

As an outcome and proxy for late-life health we use adult height, measured in centimetres. 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

²⁰The 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, 2014; 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.

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.²¹ 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 compares relative magnitudes of adult height between those who experienced hunger before age 16 and those who did not.

Table 5 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. 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 likely correlates with unobserved child attributes that also affect stature. (Table 1 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.

²¹Hypertension manifests at age 45-65 (Risemann 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).

3.4. *Covariates*

In the analysis, we include country fixed effects and country-specific time trends.²² As described in Subsection 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 increases with a linear trend (Cole, 2000; Cole, 2003, Deaton, 2007; Staub *et al.*, 2011). For a panel of European countries this can be seen for example from Figure 1 in Deaton (2009) and Figure 1 in Schmidt et al. (1995), although for a somewhat later period of study. For our time period of investigation height data and linear time trends are reported in Niewenweg *et al.* (2003) and Cavelaars *et al.* (2000) for the Netherlands and Germany and in Hatton and Bray (2010) for Greece. Figure 2 displays average height by birth year for all individuals from non-famine and famine countries in the SHARE data respectively. Moreover, Figure 3 displays average height for each famine country separately. We find that the trend in height is linear, country-specific and gender-specific, but average height tends to lie below trend during the famine periods. Table 2 provides the result of a test for nonlinear trends using our estimation sample. We find no evidence of nonlinear trends in our main specification, but nonlinear trends are sometimes significant for individual countries. In these cases, nonlinearities are driven exclusively by observations at the endpoints of our sample period and disappear if we exclude these years.²³

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 provides basic summary statistics for

²²The 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.

²³Our main results are robust to excluding the endpoints of the sample period.

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

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. Empirical Strategy

4.1. Instrumental Variable Methods

Our model framework is as follows,

$$Y_i = \psi(D_i, Z_i, X_i, \epsilon_i) \quad (1)$$

$$D_i = \phi(Z_i, X_i, \varepsilon_i), \quad (2)$$

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 centimetres. The vector of covariates is denoted by X .

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:

$$\begin{aligned} \text{LATE} &= E[Y_{D=1} - Y_{D=0} | D_{Z=1} > D_{Z=0}] \\ &= \int [\psi(1, Z, X, \epsilon) - \psi(0, Z, X, \epsilon)] dF_{X, Z, \epsilon | D_{Z=1} > D_{Z=0}}. \end{aligned} \quad (3)$$

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 $\text{cov}(Y, Z)/\text{cov}(D, Z)$. We use a version that controls for covariates. Exploiting the binary nature of Z , this is the sample equivalent of the following expression,

$$LATE = \frac{\int E[Y|X = x, Z = 1] - E[Y|X = x, Z = 0] f(x)dx}{\int E[D|X = x, Z = 1] - E[D|X = x, Z = 0] f(x)dx}. \quad (4)$$

Equation (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.

4.2. *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 $E[D|X = x, Z = 1] - E[D|X = x, Z = 0]$ in the denominator of (4) by the estimate of $[E[D|X = x, Z = 1] - E[D|X = x, Z = 0]]_{child}$ 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; Valaoras, 1946; Klatt, 1950; De Rooij et al., 2010; 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 2, 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 favoured 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 favour 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 *et al.*, 1982; Behrman, 1997). 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.*, 2012), 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.

4.3. *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 disproportionately 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 2 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 6 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.

5. Estimation Results

5.1. *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.

Tables 7 and 8 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 3.²⁵

²⁵Full lists of coefficient values for both the reduced form models and the IV models are displayed in Table 19.

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 9 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 7 and 8 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-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 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 analysing 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 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 length²⁶ 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 10. The results are qualitatively unchanged.

5.2. *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

²⁶Specifically, 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.

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 11 and 12.²⁷ They show that famine exposure reduces adult male height on average by 0.68 centimetres. 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%, then we expect the causal effect of hunger on adult height equal around -3.4 centimetres 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. The last columns of Tables 11 and 12 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 centimetres 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 the underlying causal effect of hunger on height by a factor exceeding 4 and amounting to more than 2 centimetres. 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.

²⁷Table 20 presents the full list of coefficient estimates for the reduced form models.

²⁸The 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.

Our finding that the height of males is more affected by early nutritious shocks than female stature is supported by studies described in Section 1 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 paper, there does not exist evidence for a drop in the sex ratio except for a small reduction in Greece (Cramer and Lumey, 2010; Jürges, 2011; Neelsen and Stratman, 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.

5.3. *Sensitivity Analyses and Tests*

5.3.1. *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 3. 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 3, 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 13 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.

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 13. 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

5.3.2. *Variable definitions and operationalisations*

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 operationalisation 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

²⁹The 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.

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 14 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 2, 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 $[6, 16)$ and $[-0.75, 4)$, respectively. Table 15 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 16 and 17 give results if we define the age interval to be $[-0.75, 3)$ or $[-0.75, 6)$. 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 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 18 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 centimetres.

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 19 and 20). 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.

6. Conclusions

This paper 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 2). 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 paper are restricted by some notable data limitations. First, the samples are small. Admittedly, the full SHARE data cover many European countries and

³⁰Some 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.

include over 20,000 elderly respondents, but we only use individuals in subsamples from the 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 important avenue for future research is to investigate potential heterogeneity in the effect of hunger as well as variation in famine exposure and intensity. This paper shows a wide dispersion in hunger rates indicating that there may also exist variation in the severity of hunger. If available, such information could be used to investigate differences in the effect of hunger across individuals with similar hunger experiences and non-monotonicities in the effect of hunger for different degrees of undernutrition. Besides, hunger severity might be informative about nonsystematic recall bias among very young individuals, who may report hunger only when it was severe enough

to become part of the family history. Another issue concerns the use of 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.

Figure 1
Probability to Report Hunger Conditional on Famine Exposure

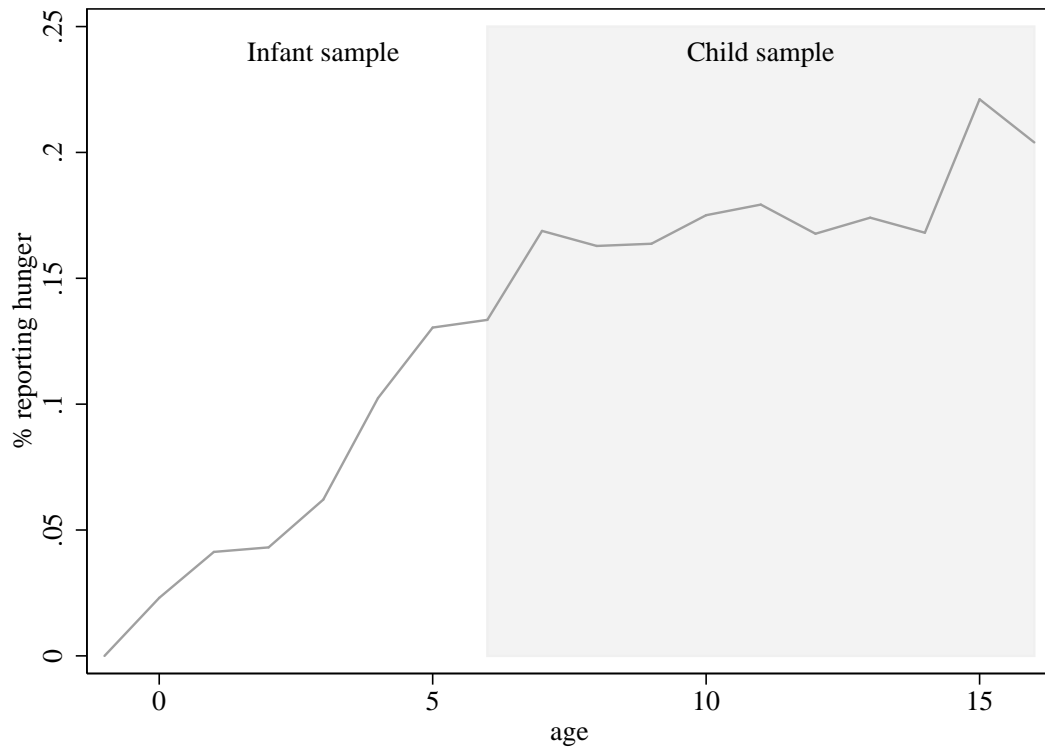


Figure 2
Height Trends

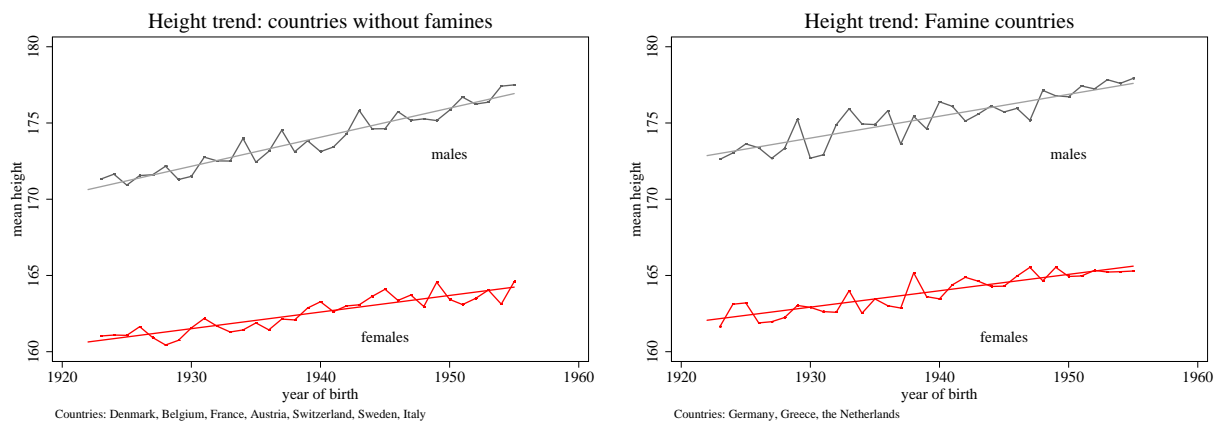


Figure 3
Height Trends (by Country)

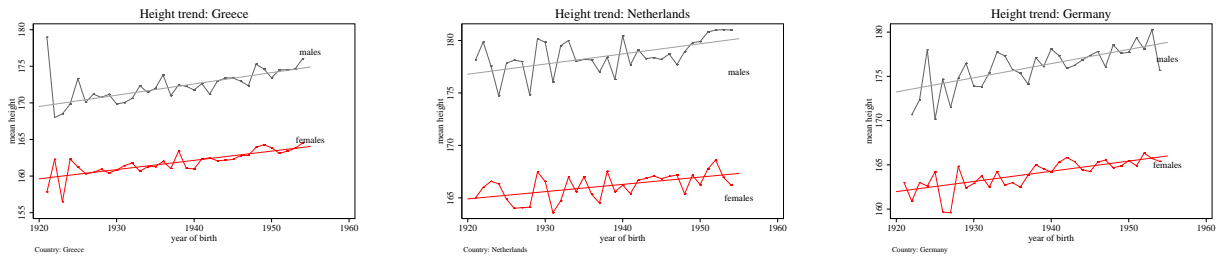


Figure 4
Comparison of Episodes of Hunger in SHARE-Germany and SOEP-Module Data

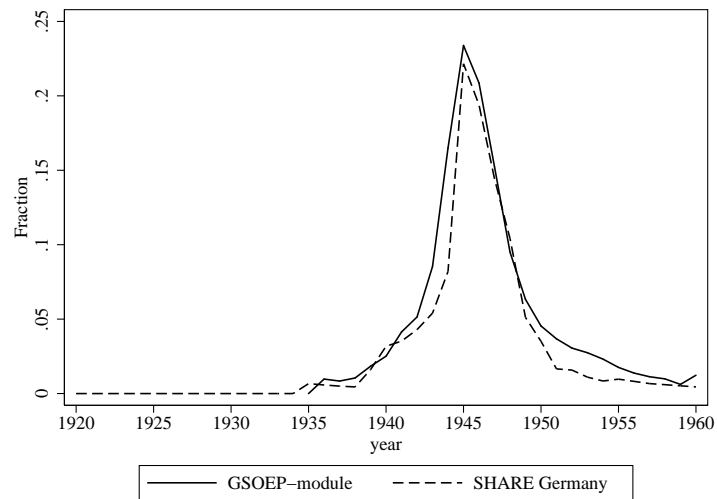


Figure 5
Example Treatment Definition

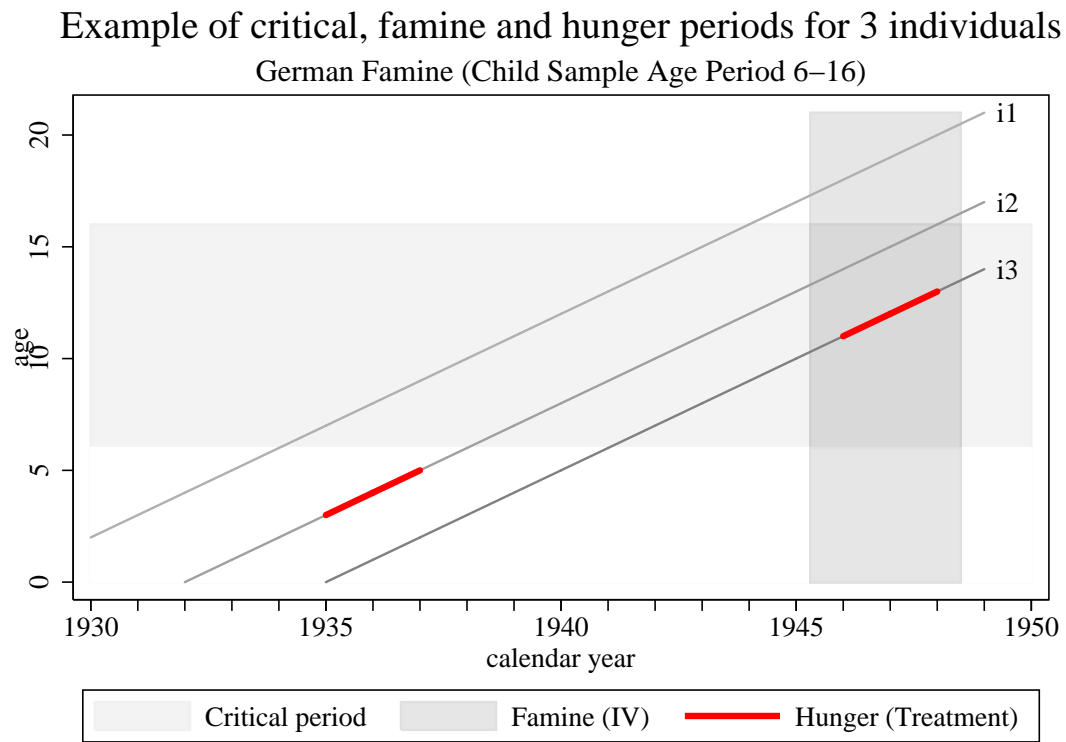


Figure 6
Probability for Episode of Hunger by Calendar Year

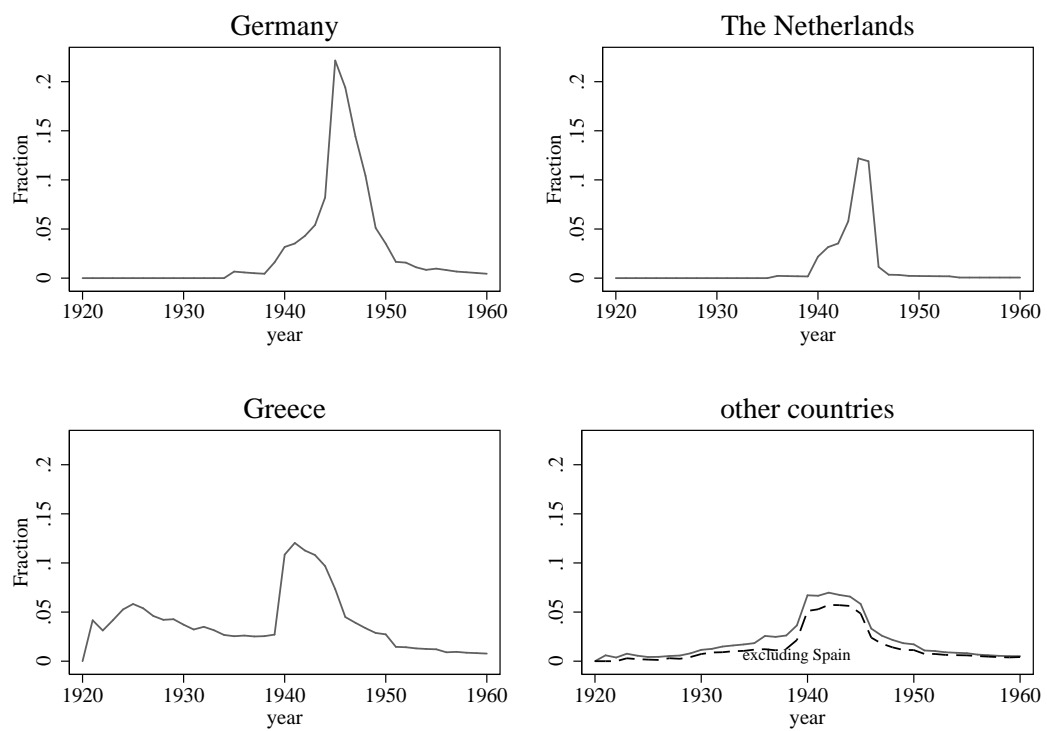


Figure 7
Average Length of Hunger Periods by Calendar Year

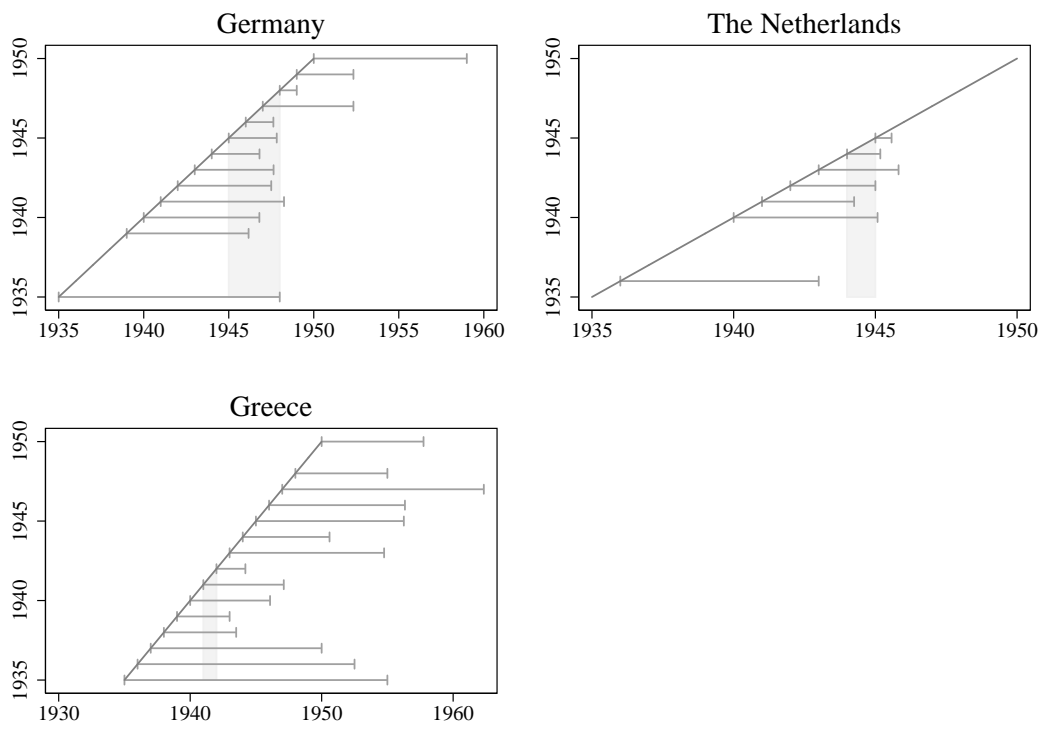


Table 1
Hunger Prevalence

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

Note: Standard errors in parentheses. 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.

Table 2
Testing for Nonlinearities

| variables | Famine countries | |
|--------------------|-------------------|--------------------|
| | males | females |
| Birth year squared | 0.00406 (0.26) | 0.000444 (0.46) |
| Observations | 2511 | 2859 |
| R^2 | 0.209 | 0.134 |
| Country intercepts | YES | YES |
| Country slopes | YES | YES |

Note: *, **, *** indicates significance at the 10%, 5%, and 1% level. Standard errors are robust and clustered at the country level. P -values in parentheses.

Table 3
Reaction of Hunger at Age 0 – 16 to Famine Exposure

| | 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
Summary Statistics for Famine Countries

| | all (N= 5370) | | control (N=4964) | | hunger <age 16 (N=406) | |
|--------------------------------|------------------|--------|---------------------|--------|---------------------------|--------|
| | mean | (s.d.) | mean | (s.d.) | mean | (s.d.) |
| Year of birth | 1942 | (8.70) | 1943 | (8.70) | 1936 | (5.80) |
| Hunger at age 0 – 16 (1 = yes) | 0.076 | (0.26) | 0.000 | (0.00) | 1.000 | (0.00) |
| Famine at age 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) |

Table 5
OLS Regression of Height on Hunger

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

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

Table 6
Occupational Categories of Main Income Earner at Age 10

| occupational status | Greece | | | Germany | | | The Netherlands | | |
|---------------------|------------|--------|-------------|------------|--------|-------------|-----------------|--------|-------------|
| | 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 % level. For the Netherlands, table includes only observations born in the Western part.

Table 7
Effects of Hunger at Age 6 – 16 on Adult Height in cm (Males)

| | reduced form famine at age 6 – 16 | 2SLS | instrumental variables models | |
|---------|--------------------------------------|---------|-------------------------------|--------------------------|
| | | | cond. Wald | cond. Wald – trend corr. |
| 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 8
Effects of Hunger at Age 6 – 16 on Adult Height in cm (Females)

| | reduced form famine at age 6 – 16 | 2SLS | instrumental variables models | |
|---------|--------------------------------------|---------|-------------------------------|--------------------------|
| | | | cond. Wald | cond. Wald – trend corr. |
| 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.

Table 9
First-stage Coefficients (Age 6 – 16)

| | males | | females | |
|------------------------------|---------------------|---------------------|---------------------|----------------------|
| Famine age 6 – 16 (1 = yes) | 0.206*** (0.020) | 0.179*** (0.024) | 0.135*** (0.018) | 0.096*** (0.025) |
| German Sample | | 2.325 (3.992) | | 4.730* (2.668) |
| Dutch Sample | | 2.950 (2.873) | | 2.550 (2.667) |
| Lived in rural area at age 6 | | −0.026** (0.010) | | −0.026*** (0.009) |
| Year of birth | | −0.002* (0.001) | | −0.002** (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*** (0.005) | 4.197* (2.124) | 0.034*** (0.006) | 4.781** (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.

Table 10
*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.

Table 11
Reduced Form Estimates (Males)

| | famine at age 0 – 4 | cond. Wald | cond. Wald – trend corr. |
|---------|---------------------|------------|--------------------------|
| 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 12
Reduced Form Estimates (Females)

| | famine at age 0 – 4 | cond. Wald | cond. Wald – trend corr. |
|---------|---------------------|------------|--------------------------|
| 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.

Table 13
Placebo Famine Exposure and Inclusion of Other Countries

| age 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 |
| age 6 – 16 (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 |
| age 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 |
| age 0 – 4 or in utero (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.

Table 14
Reduced Form Estimates

| | 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 15
(Urbanization \times Famine Exposure) as Additional Instrument

| | | 2SLS | | 2-sample 2SLS | |
|-----------------------|---------|-------------|-------------|---------------|-------------|
| | | males | | females | |
| | | outcome eq. | first stage | outcome eq. | first stage |
| Hunger | effect | 0.299 | | 3.646 | |
| | (s.e.) | (1.558) | | (2.460) | |
| | t-stat. | 0.192 | | 1.482 | |
| Famine | effect | | 0.235 | | 0.145 |
| | (s.e.) | | (0.018) | | (0.016) |
| | t-stat. | | 12.687 | | 8.882 |
| Famine \times rural | effect | | –0.125 | | –0.111 |
| | (s.e.) | | (0.024) | | (0.022) |
| | t-stat. | | –5.190 | | –5.120 |
| N | | 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 16
Different Age Intervals (Males)

| | hunger at age 0 – 3 or in utero | | |
|---------|---------------------------------|------------|--------------------------|
| | reduced form | cond. Wald | cond. Wald – trend corr. |
| Effect | –0.554 | –1.330 | –2.383 |
| (s.e.) | (0.285) | (1.861) | (1.337) |
| t-stat. | –1.948 | –0.715 | –1.782 |
| | hunger at age 0 – 6 or in utero | | |
| | reduced form | cond. Wald | cond. Wald – trend corr. |
| Effect | –0.402 | –2.118 | –1.597 |
| (s.e.) | (0.301) | (1.659) | (1.390) |
| t-stat. | –1.333 | –1.277 | –1.149 |

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 17
Different Age Intervals (Females)

| | hunger at age 0 – 3 or in utero | | |
|---------|---------------------------------|------------|--------------------------|
| | reduced form | cond. Wald | cond. Wald – trend corr. |
| Effect | 0.073 | 2.910 | 0.404 |
| (s.e.) | (0.287) | (3.077) | (2.028) |
| t-stat. | 0.254 | 0.946 | 0.199 |
| | hunger at age 0 – 6 or in utero | | |
| | reduced form | cond. Wald | cond. Wald – trend corr. |
| Effect | 0.191 | 1.080 | 1.156 |
| (s.e.) | (0.248) | (2.334) | (1.736) |
| t-stat. | 0.769 | 0.463 | 0.666 |

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 18
Reference Age Window 8 – 16

| males | | | |
|---------|---------------------|------------|--------------------------|
| | famine at age 0 – 4 | cond. Wald | cond. Wald – trend corr. |
| Effect | –0.622 | –1.960 | –2.584 |
| (s.e.) | (0.274) | (1.721) | (1.242) |
| t-stat. | –2.271 | –1.138 | –2.081 |
| females | | | |
| | famine at age 0 – 4 | cond. Wald | cond. Wald – trend corr. |
| Effect | 0.300 | 3.754 | 1.798 |
| (s.e.) | (0.296) | (2.642) | (2.028) |
| t-stat. | 1.014 | 1.421 | 0.887 |

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 19
Coefficient Estimates, Age 6 – 16

| | males | | | | females | | | |
|--------------------------------|-------------------------|-------------------------|---------------------|---------------------|------------------------|------------------------|------------------------|-----------------------|
| | reduced form | IV model | reduced form | IV model | reduced form | IV model | reduced form | IV model |
| Famine at age 6 – 16 (1 = yes) | 0.294 (0.325) | | 0.201 (0.337) | | 0.227 (0.259) | | 0.242 (0.269) | |
| Lived in rural area at age 6 | –0.380 (0.257) | –0.338 (0.259) | –0.387 (0.257) | –0.357 (0.260) | 0.018 (0.178) | 0.080 (0.195) | 0.018 (0.178) | 0.086 (0.197) |
| Year of birth | 0.161*** (0.026) | 0.165*** (0.028) | 0.118*** (0.039) | 0.124*** (0.045) | 0.133*** (0.018) | 0.139*** (0.024) | 0.141*** (0.027) | 0.152*** (0.039) |
| Year of birth × Dutch | –0.063* (0.035) | –0.061* (0.034) | –0.037 (0.040) | –0.038 (0.041) | –0.062** (0.025) | –0.059** (0.023) | –0.066** (0.029) | –0.065** (0.029) |
| Year of birth × German | 0.007 (0.036) | 0.009 (0.035) | 0.032 (0.040) | 0.031 (0.040) | –0.013 (0.019) | –0.007 (0.019) | –0.016 (0.021) | –0.013 (0.020) |
| German sample | –9.864 (69.402) | –13.686 (68.733) | –58.401 (77.570) | –56.847 (78.008) | 26.883 (36.313) | 15.707 (35.992) | 34.143 (40.678) | 26.489 (39.761) |
| Dutch sample | 128.485* (68.944) | 123.636* (66.912) | 78.245 (78.849) | 79.218 (79.167) | 124.252** (48.295) | 118.226*** (45.567) | 132.432** (56.126) | 131.070** (56.636) |
| Hunger at age 6 – 16 (1 = yes) | | 1.644 (1.818) | | 1.170 (1.954) | | 2.363 (2.790) | | 2.613 (3.035) |
| Life expectancy at birth | | | 0.055* (0.030) | 0.050 (0.033) | | | –0.009 (0.022) | –0.015 (0.027) |
| Constant | –140.031*** (50.766) | –146.931*** (55.212) | –59.497 (73.883) | –71.497 (86.227) | –96.449*** (34.539) | –107.747** (46.919) | –110.094** (52.108) | –131.424* (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 20
Coefficient Estimates, Age 0 – 4

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

References

- Akbulut-Yuksel, M. (2014). 'Children of war: the long-Run effects of large-scale physical destruction and warfare on children', *Journal of Human Resources*, vol. 49(3), pp. 634–662.
- Alderman, H. (2012). 'The response of child nutrition to changes in income: linking biology with economics', *CESifo Economic Studies*, vol. 58(2), pp. 256–273.
- Almond, D. and Currie, J. (2011). 'Killing me softly: The fetal origins hypothesis' *Journal of Economic Perspectives*, vol. 25(3), pp. 153–172.
- Almond, D. and Mazumder, B. (2011). 'Health capital and the prenatal environment: the effect of Ramadan observance during pregnancy', *American Economic Journal: Applied Economics*, vol. 3(4), pp. 56–85.
- Angrist, J.D., Krueger, A.B. (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*, vol. 87, pp. 328–336.
- Angrist, J.D., Imbens, G.W., and Rubin, D.B. (1996). 'Identification of causal effects using instrumental variables', *Journal of the American Statistical Association*, vol. 91(434), pp. 444–455.
- Arellano, M., Meghir, C. (1992). 'Female labour supply and on-the-job-search: An empirical model estimated using complementary data sets', *Review of Economic Studies*, vol. 59, pp. 537–557.
- Banning, C. (1946). 'Food shortage and public health, first half of 1945', *Annals of the American Academy of Political and Social Science*, vol. 245, pp. 93–110.
- Barker, D.J.P. (1994). *Mothers, Babies, and Disease in Later Life*, London: British Medical Journal Publishing Group.
- Barker, D. (2007). 'The origins of the developmental origins theory', *Journal of Internal Medicine*, vol. 261, pp. 412–417.
- Bateson, P. (2001). 'Fetal experience and good adult design', *International Journal of Epidemiology*, vol. 30:, pp. 928–934.
- Becker, G.S. and Tomes, N. (1976). 'Child endowments and the quantity and quality of children', *Journal of Political Economy*, vol. 84(4), pp. 143–162.
- Behrman, J.R. and Pollak, R.A. and Taubman, P. (1982). 'Parental preferences and provision for progeny', *Journal of Political Economy*, vol. 90(1), pp. 52–73.
- Behrman, J.R. (1997). 'Intrahousehold distribution and the family', *Handbook of Population and Family Economics*, vol. 1, pp. 125–187.
- Blom, A. G., and Schröder, M. (2011). '7 Sample Composition 4 Years on: Retention in SHARE Wave 3', in (M. Schröder, ed.), *SHARELIFE Methodology*, pp. 55–61, Mannheim: Mannheim Research Institute for the Economics of Ageing (MEA).
- Börsch-Supan, A., Brügiavini, A., Jürges, H., Kapteyn, A., Mackenbach, J., Siegrist, J., et al. (2008). *Health, ageing and retirement in Europe (2004-2007). Starting the longitudinal dimension*, Mannheim: Mannheim Research Institute for the Economics of Aging (MEA).
- Cameron, N. and Demerath, E. W. (2002). 'Critical periods in human growth and their relationship to diseases of aging', *Yearbook of Physical Anthropology*, vol. 45, pp. 159–184.

- Case, A. and Paxson, C. (2005). ‘Sex Differences in Morbidity & Mortality’, *Demography*, 42, pp. 189–214.
- Cavelaars, A. E. et al. (2000). ‘Persistent variations in average height between countries and between socio-economic groups: an overview of 10 European countries’, *Annals of Human Biology*, vol. 27(4), pp. 407–421.
- Chen, Y. and Zhou, L. (2007). ‘The long-term health and economic consequences of the 1959-1961 famine in China’, *Journal of Health Economics*, vol. 26(4), pp. 659–681.
- Cole, S.W., Hawkey, L.C., Arevalo, J.M.G., Cacioppo, J.T. (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 of the USA*, vol. 108, pp. 3080–3085.
- Cole, T.J. (2000). ‘Secular trends in growth’, *Proceedings of the Nutrition Society*, vol. 59(2), pp. 317–324.
- Cole, T.J. (2003). ‘The secular trend in human physical growth: a biological view’, *Economics & Human Biology*, vol. 1, pp. 161–168.
- Cornides, W. (1948). *Wirtschaftsstatistik der deutschen Besatzungszonen 1945-1948 in Verbindung mit der deutschen Produktionsstatistik der Vorkriegszeit*. Oberursel: Verlag Europa-Archiv.
- Costa, D. L. (1993). ‘Height, Wealth, and Disease among the Native-Born in the Rural, Antebellum North’, *Social Science History*, vol. 17(3), pp. 355–383.
- Cramer, J.S. and Lumey, L.H. (2010). ‘Maternal preconception diet and the sex ratio’, *Human biology*, vol. 82(1), pp. 103–107.
- Datar, A., Kilburn, M.R. and Loughran, D.S. (2010). ‘Endowments and parental investments in infancy and early childhood’, *Demography*, vol. 47(1), pp. 145–162.
- Dean, R. F.A. (1951). ‘The size of the baby at birth and the yield of breast milk’, *Medical Research Council Special Report Series*, vol. 275, pp. 346–378.
- Deaton, A. (2007). ‘Height, health, and development’, *Proceedings of the National Academy of Sciences*, vol. 104(33), pp. 13232–13237.
- Del Bono, E. and Ermisch, J. and Francesconi, M. (2012). ‘Intrafamily Resource Allocations: A Dynamic Structural Model of Birth Weight’, *Journal of Labor Economics*, vol. 30(3), pp. 657–706.
- De Rooij, S.R., Wouters, H., Yonker, J.E., Painter, R.C. and Roseboom, T.J. (2010). ‘Prenatal undernutrition and cognitive function in late adulthood’, *PNAS*, vol. 107(39), pp. 16881–16886.
- Doblhammer, G., Van den Berg, G.J., Lumey, L.H. (2013). ‘A re-analysis of the long-term effects on life expectancy of the Great Finnish Famine of 1866-68’, *Population Studies*, vol. 67, pp. 309–322.
- Dols, M.J.L. and Van Arcken, D.J.A.M. (1946). ‘Food supply and nutrition in the Netherlands during and immediately after World War II’, *The Milbank Memorial Fund Quarterly*, vol. 24(4), pp. 319–358.
- Echternkamp, J. (2003). *Nach dem Krieg – Alltagsnot, Neuorientierung und die Last der Vergangenheit 1945–1949*. Zürich: Pendo Verlag.

- Elder, G. (1999). *Children of the Great Depression: Social Change in Life Experience*. Boulder: Westview Press.
- Elo, I.T., and Preston, S.H. (1992). 'Effects of early-life conditions on adult mortality: a review', *Population Index*, vol. 58, pp. 186–212.
- Eriksson, J.G., E. Kajantie, C. Osmond, K. Thornburg and Barker, D.J.P. (2010). 'Boys live dangerously in the womb', *American Journal of Human Biology*, vol. 22, pp. 330–335.
- Eveleth P.B. and Tanner J.M. (1990). *Worldwide Variation in Human Growth*. Cambridge: Cambridge University Press.
- Farquharson, J. (1985). *The Western Allies and the Politics of Food – Agrarian Management in Postwar Germany*. Oxford: Berg Publishers Ltd.
- Gasser, T., Ziegler, P., Largo, R.H., Molinari, L., and Prader, A. (1994). 'A longitudinal study of lean and fat areas at the arm', *Annals of Human Biology*, vol. 21(4), pp. 303–14.
- Gluckman, P.D., and Hanson, M.A. (2004). 'Living with the past: evolution, development, and patterns of disease', *Science*, vol. 305, pp. 1733–1736.
- Gluckman, P.D., Hanson, M.A., and Pinal, C. (2005). 'The developmental origins of adult disease', *Maternal and Child Nutrition*, vol. 1, pp. 130–141.
- Griliches, Z. (1979). 'Sibling models and data in economics: beginnings of a survey', *Journal of Political Economy*, vol. 87(5), pp. 37–64.
- Hadley, C. and Patil, C. (2006). 'Food insecurity in rural Tanzania is associated with maternal anxiety and depression', *American Journal of Human Biology*, vol. 18(3), pp. 359–368.
- Hales, C.N., and Barker, D.J. (1992). 'Type 2 (non-insulin-dependent) diabetes mellitus: the thrifty phenotype hypothesis', *Diabetologia*, vol. 35(7), pp. 595–601.
- Hamelin, A., Habicht, J., and Beaudry, M. (1999). 'Food insecurity: consequences for the household and broader social implications', *The Journal of Nutrition*, vol. 129(2), pp. 525S–528S.
- Hatton, T. J. and Bray, B. E. (2010). 'Long run trends in the heights of European men, 19th/20th centuries', *Economics & Human Biology*, vol. 8(3), pp. 405–413.
- Häusser, A. and Maugg, G. (2009). *Hungerwinter. Deutschlands humanitäre Katastrophe 1946/47*, Berlin: Propyläen.
- Havari, E. and Peracchi, F. (2012). 'Childhood circumstances and adult outcomes: Evidence from SHARELIFE', Working Paper, Einaudi Institute for Economic and Finance (EIEF).
- Heijmans, B. T., Tobi, E. W., Stein, A. D., Putter, H., Blauw, G. J., Susser, E. S., et al. (2008). 'Persistent epigenetic differences associated with prenatal exposure to famine in humans', *Proceedings of the National Academy of Sciences*, vol. 105(44), pp. 17046–17049.
- Hermanussen, M., Danker-Hopfe, H., Weber, G.W. (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*, vol. 25, pp. 1550–1553.
- Hionidou, V. (2006). *Famine and Death in Occupied Greece, 1941-1944*. Cambridge: Cambridge University Press.

- Huxley, R., Owen, C., Whincup, P., Cook, D., Rich-Edwards, J., Smith, G., and Collins, R. (2007). 'Is birth weight a risk factor for ischemic heart disease in later life?', *The American journal of clinical nutrition*, vol. 85(5), pp. 1244–1250.
- Imbens, G.W., and Angrist, J.D. (1994). 'Identification and estimation of local average treatment effects', *Econometrica*, vol. 62(2), pp. 467–475.
- Inoue, A., and Solon, G. (2010). 'Two-sample instrumental variables estimators', *Review of Economics and Statistics*, vol. 92(3), pp. 557–561.
- Jürges, H. (2011). 'Collateral damage: Educational attainment and labor market outcomes among German war and post-war cohorts', Working Paper, University of Wuppertal.
- Jyoti, D., Frongillo, E., and Jones, S. (2005). 'Food insecurity affects school children's academic performance, weight gain, and social skills', *The Journal of Nutrition*, vol. 135(12), pp. 2831–2839.
- Klatt, W. (1950). 'Food and farming in Germany: I. food and nutrition', *International Affairs*, vol. 26(1), pp. 45–58.
- Koupil I., Plavinskaja S., Parfenova N., Shestov D.B., Danziger P.D., Vågerö D. (2009). 'Cancer mortality in women and men who survived the siege of Leningrad (1941-1944)', *International Journal of Cancer*, vol. 124(6), pp. 1416–1421.
- Kraemer, S. (2000). 'The fragile male', *BMJ: British Medical Journal*, vol. 321(7276), pp. 1609–1612.
- Kuh, D.L., Power C., and Rodgers, B. (1991). 'Secular trends in social class and sex differences in adult height', *International Journal of Epidemiology*, vol. 20(4), pp. 1001–1009.
- Kuzawa, C., and Quinn, E.A. (2009). 'Developmental origins of adult function and health: evolutionary hypotheses', *Annual Review of Anthropology*, vol. 38, pp. 131–147.
- Lawlor D.A., Ben-Shlomo Y., Leon D.A. (2004). 'Pre-adult influences on cardiovascular disease', in (Kuh D., Ben-Shlomo Y., eds.), *A Life Course Approach to Chronic Disease Epidemiology*, pp. 41–76, Oxford: Oxford University Press.
- Lawlor, D.A. (2008). 'The developmental origins of health and disease: where do we go from here?', *Epidemiology*, vol. 19, pp. 206–208.
- Leigh, J. (1983). 'Direct and indirect effects of education on health', *Social Science & Medicine*, vol. 17(4), pp. 227–234.
- Lindeboom, M., Portrait, F., and Van den Berg, G. (2010). 'Long-run effects on longevity of a nutritional shock early in life: The Dutch potato famine of 1846-1847', *Journal of health economics*, vol. 29(5), pp. 617–629.
- Low, B.S. (2001). *Why sex matters: A Darwinian look at human behavior*. Princeton: Princeton University Press.
- Lumey, L.H., Stein, A.D., Kahn, H.S., Van der Pal-de Bruin, K.M., Blauw, G.J., Zybert, P.A. et al. (2007). 'Cohort profile: the Dutch Hunger Winter Families Study', *International Journal of Epidemiology*, vol. 36, pp. 1196–1204.
- Lumey, L.H., Stein, A.D., and Susser, E. (2011). 'Prenatal famine and adult health', *Annual Review of Public Health*, vol. 32, pp. 237–262.

- Marshall, W., and Tanner, J.M. (1986). 'Puberty', in (F. Falkner and J.M. Tanner, eds.), *Human Growth. A Comprehensive Treatise. Part 2. Postnatal Growth Neurobiology*, pp. 171–209, New York: Plenum Press.
- Miller, G.E., Chen, E., Fok, A.K., Walker, H., Lim, A., Nicholls, E.F., et al. (2009). 'Low early-life social class leaves a biological residue manifest by decreased glucocorticoid and increased proinflammatory signaling', *Proceedings of the National Academy of Sciences*, vol. 106, pp. 14716–14721.
- Miller, G.E., and Chen, E. (2010). 'Harsh family climate in early life presages the emergence of a proinflammatory phenotype in adolescence', *Psychological Science*, vol. 21, pp. 848–856.
- Morozink, J.A., Friedman, E.M., Coe, C.L., and Ryff, C.D. (2010). 'Socioeconomic and psychosocial predictors of interleukin-6 in the MIDUS national sample', *Health Psychology*, vol. 29(6), pp. 626–635.
- Mu, R. and Zhang, X. (2011). 'Why does the Great Chinese Famine affect the male and female survivors differently? Mortality selection versus son preference', *Economics & Human Biology*, vol. 9(1), pp. 92–105.
- Neelsen, S. and Stratmann, T. (2011). 'Effects of prenatal and early life malnutrition: Evidence from the Greek famine', *Journal of Health Economics*, vol. 30(3), pp. 479–488.
- Niewenweg, R., Smit, M. L., Walenkamp, M. J. E. and Wit, J. M. (2003). 'Adult height corrected for shrinking and secular trend', *Annals of human biology*, vol. 30(5), pp. 563–569.
- O'Grada, C. (2009). *Famine: A Short History*. Princeton: Princeton University Press.
- Oppers, V.M. (1963). 'Analyse van de acceleratie van de menselijke lengtegroei door bepaling van het tijdstip van de groeifasen', PhD thesis, University of Amsterdam.
- Painter, R.C., Roseboom, T.J., and Bleker, O.P. (2005). 'Prenatal exposure to the Dutch Famine and disease in later life: an overview', *Reproductive Toxicology*, vol. 20, pp. 345–352.
- Pitkäinen, K.J. (1993). *Mortality during the Great Finnish Famine of the 1860s*. Helsinki: Finnish Demographic Society.
- Pollitt, R.A., K.M. Rose and J.S. Kaufman (2005). 'Evaluating the evidence for models of life course socioeconomic factors and cardiovascular outcomes: A systematic review', *BMC Public Health*, vol. 5(7), pp. 1–13.
- Poulter, N.R., C.L. Chang, A.J. MacGregor, H. Snieder and T.D. Spector (1999). 'Association between birth weight and adult blood pressure in twins: historical cohort study', *British Medical Journal*, vol. 319(7221), pp. 1330–1333.
- Rasmussen, K. (2001). 'The fetal origins hypothesis: challenges and opportunities for maternal and child nutrition', *Annual Review of Nutrition*, vol. 21(1), pp. 73–95.
- Reichardt, S. and Zierenberg, M. (2009). *Damals nach dem Krieg. Eine Geschichte Deutschlands 1945 bis 1949*. München: Goldmann.
- Ridder, G. and Moffitt, R. (2007). 'The econometrics of data combination', in (J.J. Heckman and E.E. Leamer, eds.), *Handbook of Econometrics, Volume 6*, pp. 5469–5547, Amsterdam: Elsevier.
- Riseman, J.E.F., and Weiss, S. (1929). 'The age and sex incidence of arterial hypertension', *American Heart Journal*, vol. 5(2), pp. 172–190.

- Rose, D. and Oliveira, V. (1997). 'Nutrient intakes of individuals from food-insufficient households in the United States', *American Journal of Public Health*, vol. 87(12), pp. 1956–1961.
- Rosenzweig, M. R. and Wolpin, K.I. (1988). 'Heterogeneity, intrafamily distribution, and child health', *Journal of Human Resources*, vol. 23(4), pp. 437–461.
- Sanz-de-Galdeano, A. (2005). 'The obesity epidemic in Europe', Working Paper, IZA Bonn.
- Schmidt, I.M., Jørgensen, M.H., and Michaelsen, K.F. (1995). 'Height of conscripts in Europe: is postneonatal mortality a predictor?', *Annals of Human Biology*, vol. 22(1), pp. 57–67.
- Scholte, R.S., van den Berg, G.J., and Lindeboom, M. (2012). 'Long-run effects of gestation during the Dutch Hunger Winter famine on labor market and hospitalization outcomes', Working Paper, IZA Bonn.
- 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.C. (2010). 'The Dutch Hunger Winter and the developmental origins of health and disease', *Proceedings of the National Academy of Sciences*, vol. 107, pp. 16757–16758.
- Sidel, V.W. (1997). 'The public health impact of hunger', *American Journal of Public Health*, vol. 87(12), pp. 1921–1922.
- Sparén, P., Vågerö, D., Shestov, D.B., Plavinskaja, S., Parfenova, N., Hoptiar, V., et al. (2004). 'Long term mortality after severe starvation during the siege of Leningrad: prospective cohort study', *British Medical Journal*, vol. 328, pp. 11–15.
- Staub, K., Rühli, F.J., Woitek, U. and Pfister, C. (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: official journal of the Swiss Society of Infectious Diseases, the Swiss Society of Internal Medicine, the Swiss Society of Pneumology*, vol. 141, pp. w13238–w13238.
- Steckel, R. H. (2008). 'Biological measures of the standard of living', *Journal of Economic Perspectives*, vol. 22(1), pp. 129–152.
- Stein, Z. and Susser, M., T.P. (1975). 'Fertility, fecundity, famine: food rations in the Dutch famine 1944/5 have a causal relation to fertility, and probably to fecundity', *Human Biology*, vol. 47(1), pp. 131–154.
- Susser, E.S., and Lin, S.P. (1992). 'Schizophrenia after prenatal exposure to the Dutch Hunger Winter of 1944-45', *Archives of General Psychiatry*, vol. 49, pp. 983–988.
- Susser, M., and Stein, Z. (1994). 'Timing in prenatal nutrition: A reprise of the Dutch famine study', *Nutrition Reviews*, vol. 52, pp. 84–94.
- Trittel, G. J. (1990). *Hunger und Politik – Die Ernährungskrise in der Bizone (1945–1949)*. Frankfurt: Campus Verlag.
- Valaoras, V.G. (1946). 'Effects of famine on the population of Greece', *The Milbank Memorial Fund Quarterly*, vol. 24, pp. 215–235.
- Van Abeelen, A.F., Elias, S.G., Bossuyt, P.M., Grobbee, D.E., van der Schouw, Y.T., Roseboom, T.J., et al. (2012). 'Famine exposure in the young and the risk of type 2 diabetes in adulthood', *Diabetes*, vol. 61(9), pp. 2255–2260.

- Van den Berg, G.J., Lundborg, P., Nystedt, P., and Rooth, D. (2014). ‘Critical periods during childhood and adolescence’, *Journal of the European Economic Association*, forthcoming.
- Van den Berg, G.J., Doblhammer G., and Christensen, K. (2009). ‘Exogenous determinants of early-life conditions, and mortality later in life’, *Social Science & Medicine*, vol. 68, pp. 1591–1598.
- Van den Berg, G.J., Pinger, P.R. and Schoch, J. (2011). ‘Instrumental variable estimation of the causal effect of hunger early in life on health later in life’, Working Paper, IZA Bonn.
- Van den Berg, G.J. and Pinger, P.R. (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’, Working Paper, IZA Bonn.
- Van den Berg, G.J. and Lindeboom, M. (2014). ‘Aging: conditions early in life and health and mortality at advanced ages’, in: A. Culyer, ed., *Encyclopedia of Health Economics*, Elsevier, Amsterdam.
- Waldron, I. (1983). ‘The role of genetic and biological factors in sex differences in mortality’, in (A.D. Lopez and L.T. Ruzicka, eds.), : *Sex Differentials in Mortality: Trends, Determinants and Consequences*, Canberra: Australian National University.
- Whitaker, R., Phillips, S., and Orzol, S. (2006). ‘Food insecurity and the risks of depression and anxiety in mothers and behavior problems in their preschool-aged children’, *Pediatrics*, vol. 118(3), pp. 859–868.
- Willenborg, K.-H. (1979). ‘Überleben nach dem Zusammenbruch – die Not der Besiegten’, in (J. Weber, ed.), *Geschichte der Bundesrepublik Deutschland, Band 1: Auf dem Wege zur Republik 1945–1947*, pp. 115–144, Paderborn: Verlag Ferdinand Schöningh.
- Zemel, B. (2002). ‘Body composition during growth and development’, in (N. Cameron, ed.), *Human Growth and Development*, pp. 271–294, St. Louis: Academic Press.
- Zhang, T.Y., Bagot, R., Parent, C., Nesbitt, C., Bredy, T.W., Caldji, C., et al. (2006). ‘Maternal programming of defensive responses through sustained effects on gene expression’, *Biological Psychology*, vol. 73(1), pp. 72–89.