Essays on Skills, Health and Human Inequality

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

Pia Rosina Dovern-Pinger

vorgelegt im Sommer 2013

Abteilungssprecher:	Prof. Dr. Martin Peitz
Referent:	Prof. Gerard van den Berg, Ph.D.
Korreferent:	Prof. James J. Heckman, Ph.D.

Verteidigung: 4. Juli 2013

Für meine Eltern

Acknowledgements

It would not have been possible to write this thesis without the help and support of the kind people around me. Above all, I would like to thank Jonas Dovern for his great personal support, patience and care. I am also highly indebted to my family and friends who supported me continuously.

I am extremely grateful to my supervisor Gerard van den Berg for interesting discussions, helpful comments, continuous support, encouragement and patience. I also want to thank for several amazing and fruitful stays at the University of Chicago and for the support and mentoring of Professor Heckman, who found time to guide and teach me despite his many obligations.

During my PhD studies, I have interacted and collaborated with many exceptionally smart, dedicated and inspiring people. Thank you to my co-authors Gabriella Conti, Friedhelm Pfeiffer, Isabel Ruhmer, Johannes Schoch Rémi Piatek and Arianna Zanolini, to my fellow students Philipp Eisenhauer, Lisandra Flach, Christina Kolerus and Verena Niepel, and in particular to Miriam Gensowski. Finally, I am very grateful to all department members at University of Mannheim and ZEW who accompanied my graduation and research.

Mannheim, Summer 2013 *Pia Rosina Pinger*

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This dissertation analyzes how skills and health as two facets of human capital affect labor market outcomes, education decisions and the intergenerational transmission of inequality. Moreover, it elaborates on how differences in skill and health arise using famines as macroeconomic shocks to the formation of human capital. Throughout, the point of view is an economic one. Skills and health are viewed as a means to generate direct or indirect returns, e.g. in the form of higher wages or lower health care costs.

The thesis covers three aspects of the economics of human inequality: the formation of skills and health, their impact on the transmission of inequality, and the effect of skills on education or labor market outcomes. First, in a joint chapter with Rémi Piatek, we establish that individuals with a more internal locus of control earn higher wages, and that this effect mainly operates through the channel of higher education. In the last chapter of this thesis I show that paternal unemployment causally reduces offspring educational attainment and that a child's subjective probability of school success is an important mechanism through which this effect operates. Second, in two joint chapters with Gabriella Conti, James Heckman and Arianna Zanolini on the one hand, and with Gerard van den Berg on the other hand, we ask how differences in skills and health capital affect next generation outcomes. The first of these chapters takes a closer look at the behavioral channels of maternal smoking and education, while the second one refers to the biological channel of epigenetic imprinting. Third, in a chapter with Gerard van den Berg and Johannes Schoch, we investigate how nutritional shocks during childhood affect health outcomes in adulthood.

Econometrically and and with respect to identification, the dissertation takes two different approaches. In chapters two and three, factor structure models are implemented in a Roy

model setting to address the problem of measurement error that arises if skills and health are measured using imperfect proxies. The estimation and identification relies on a form of matching on unobservables. In chapters four and five, identification comes from an exogenous variation. Famines are used as an instrument that exogenously shifts part of the population into a state of adverse childhood conditions. Chapter six uses a combination of both approaches.

The thesis combines research on two strands of ideas that are novel to the economic literature. First, human capital is increasingly viewed as a broad and multifaceted concept that does not only comprise IQ, schooling or other cognitive measures, but also personality variables, mental and physical health. Second, nature and nurture are inseparable. The genetic make-up, phenotype and character of an individual is influenced by in utero and childhood conditions, and life experiences cause epigenetic imprinting, which in itself is heritable to subsequent generations.

From a policy perspective, economic research that incorporates multifaceted human capital and the interaction of nature and nurture are needed. First, if human capital is multidimensional policy makers need to know its facets which are most important with respect to labor market outcomes and adult health. Second, if nurture determines nature, research needs to identify sensitive and critical periods, as well as favourable and unfavourable conditions for the development of human capital.

The second chapter of this dissertation contributes to the first strand of literature. It investigates how locus of control as a particular facet of human capital influences labor market success and establishes that individuals with an internal locus of control, i.e., who believe that reinforcement in life comes from their own actions instead of being determined by luck or destiny, earn higher wages. However, this positive effect only translates into labor income via the channel of education. Factor structure modeling is implemented on an augmented data set coming from two different samples. By so doing, we are able to correct for potential biases that arise due to reverse causality and spurious correlation, and to investigate the impact of premarket locus of control on later outcomes.

The third chapter combines research on multifaceted human capital and on the interaction between nature and nurture. It focuses on newborn health outcomes as are an important ring in the chain of intergenerational transmission of disadvantage. The chapter contributes to the literature on the determinants of health at birth in two ways. First, we

analyze the role of maternal endowments and investments (education and smoking in pregnancy) in the probability of having a baby who is small for gestational age (SGA). We both estimate the total impact of maternal endowments on birth outcomes, and we also decompose it into a direct, "biological" effect and a "choice" effect, mediated by maternal behaviors. Secondly, we estimate the causal effects of maternal education and smoking in pregnancy, and we investigate whether women endowed with different traits have different returns. We find that cognition affects birth outcomes primarily through education, that personality traits mainly operate by changing smoking behavior, and that the physical fitness of the mother has a direct, "biological" effect on SGA. We also find significant heterogeneity in the effects of education and smoking along the distribution of maternal physical traits, suggesting that women with a less healthy physical constitution should be the primary target of prenatal interventions.

The fourth chapter stands at the crossroads of economics and human biology. It investigates findings from the recent biological literature according to which lifetime experiences of one generation affect later generations through epigenetic imprinting. Recent studies have found an association between an individual's famine exposure at ages 8-12 and her grandchild's longevity, as well as cardiovascular and diabetes mortality in a single historical dataset. In this chapter, we investigate the validity of these findings, by analyzing the impact of the German famine of 1916-1918 on the children and grandchildren of individuals who were affected by the famine. We find that male second-generation individuals are shorter if their mother has been affected and taller if their father has been affected. Among the third generation, males tend to have higher mental health scores if their paternal grandfather experienced the famine and females tend to have higher mental health scores if their maternal grandmother was affected. We do not find robust effects on schooling as measured by the probability of obtaining a higher secondary school degree.

The fifth chapter again focuses on the idea of human capital formation. It uses famines as exogenous variation, but this time estimates the causal effect of a nutritional shortage during childhood on adult health as opposed to the overall reduced form famine effect. We estimate this average causal effect on adult height as a proxy for late life health outcomes, 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 often exceed famine effects by a factor three.

The last chapter investigates the impact of an important economic shock, namely paternal unemployment, on child education decisions. It uses exogenous variation in the local unemployment rate to identify the unemployment effect and to this end combines German representative household data with labor market information on 97 regions for the years 1998-2009. I find that paternal unemployment decreases the probability of upper secondary school choice by around 18 percentage points. Further, paternal unemployment has negative effects on measures of child self-confidence, locus of control and mental health. My results indicate that the subjective probability of school success is an important mechanism through which paternal unemployment influences a child's educational choices. This finding is consistent with a theoretical framework where paternal unemployment affects the return to education through the subjective probability of successful school completion.

Chapter 2

Maintaining (Locus of) Control? Assessing the impact of locus of control on education decisions and wages[1]

2.1 Introduction

Does it make a difference if you think *you* can make a difference? Will it affect your decision making, or even your productivity? In response to such kinds of questions, the economic literature has recently come to acknowledge the considerable importance of *personality traits* in explaining education choices, as well as a large variety of labor market outcomes. The present chapter focuses on *locus of control*, one dimension of personality that measures the extent to which individuals believe that what happens to them in life is related to their own actions and decisions, or on the contrary to fate and luck. We contribute to the existing literature on personality traits by investigating the impact of locus of control on wages, while making a distinction between the direct—or *productive*—impact of locus of control, and the indirect—or *behavioral*—impact that works through education decisions. We find that locus of control is an important predictor of the decision to obtain higher education. Furthermore, we find that *premarket* locus of control, defined as locus of control

[1]This chapter is joint work with Rémi Piatek. A very similar version of this chapter is published as IZA working paper number 5289.

measured at the time of upper middle school—before the individual enters the labor market—does not significantly affect later wages after controlling for education decisions. In light of the existing literature, which finds mostly positive effects of contemporaneous locus of control measures on wages, this indicates that it is important to distinguish between *premarket* skills and those that are already influenced by labor market experience and age. Last, simulation of our model shows that moving individuals from the first to the last decile of the locus of control distribution significantly shifts the distribution of schooling choices, thus indirectly affecting later wages.

From a methodological point of view, there are two major econometric problems at stake in the economic literature on personality traits: measurement error and endogeneity (Bowles and Gintis, 2002; Borghans et al., 2008). First, measurement error arises because certain traits or characteristics are measured by questions or tests that are imperfect proxies of the true latent ability. Yet, in general, most psychological measures are designed to capture a particular latent trait or skill, such that factor analytical approaches can be used to distinguish true latent abilities from measurement error (Borghans et al., 2008; Heckman et al., 2006; Hansen et al., 2004). Second, endogeneity arises in the study of the impact of locus of control on labor market outcomes for two reasons. On the one hand, the results may be flawed by reverse causality, as (anticipated) labor market outcomes may affect locus of control (e.g., see Trzcinski and Holst, 2010; Gottschalk, 2005). For this reason, locus of control measures may reflect, rather than cause, the outcomes they are supposed to predict (Borghans et al., 2008). In this case, the coefficient on locus of control is biased, because of nonzero covariance between the measures and the error term. On the other hand, both outcomes and measures may be affected by past labor market experiences, which are usually not accounted for. The consequence is, again, an overestimation of the locus of control coefficient due to spurious correlation.

In the literature, four main strategies have been adopted to address this endogeneity issue. First, Duncan and Morgan (1981) and Duncan and Dunifon (1998) using the PSID, extract measures of personality traits as measured 15-25 years prior to earnings. A similar strategy has been adopted by Heckman et al. (2006), who use locus of control measurements in the NLSY taken at age 14-22 to explain later outcomes. Second, Bowles et al. (2001), using the National Longitudinal Survey of Young Women (NLSYW), employ contemporary measurements of locus of control, which they purge of past wage

influences. Third, Osborne (2000) uses past skills to instrument for contemporaneous skill measures. Last, Cunha and Heckman (2008) explicitly model development and accumulation of skills as a technology of skill formation, in which investments in one period affect the productivity of investments in subsequent periods. However, their focus is mainly on early childhood development of skills, and not on the impact of labor market experiences and various life-time shocks on skill development and income.

Using data from the German Socioeconomic Panel (GSOEP), we address the problem of measurement error by extracting a latent factor reflecting locus of control. In addition, we account for the problem of reverse causality and truncated life-cycle data in that we combine information on both young individuals, who have not yet entered the labor market, and on older, working-age individuals. Our estimation approach follows the work by Heckman et al. (2006); Hansen et al. (2004); Carneiro et al. (2003) in that we use Markov chain Monte Carlo (MCMC) methods to simulate the parameters of the model. Specifically, we use a Gibbs sampler with flat priors that sequentially draws the parameters of interest from their respective conditional distributions. Furthermore, we build on a strategy developed in Cunha et al. (2005), which allows us to retrieve the distribution of locus of control from a sample of young individuals, and to estimate its impact on outcomes in a sample of older individuals.

The contribution of this chapter is twofold. First, we apply novel econometric methods and show that Bayesian factor structure models can be a solution to endogeneity problems if researchers are confronted with truncated life cycle data, as is very often the case in the fields of personality and economics. Second, embedding our empirical results in a simple theoretical framework, we establish that locus of control only affects the psychic cost of education but is not directly rewarded on the labor market of young professionals.

The chapter proceeds as follows. Section 2.2 provides an overview of the existing literature on locus of control. In Section 6.3, a simple framework is introduced to help understand the potential impact of locus of control on education decisions and labor market outcomes. Section 2.4 describes our estimation strategy relying on data set combination to identify the full likelihood. The Bayesian approach used to sample the parameters of interest is outlined, and an overview of the data is provided. Section 6.6 presents the results of our analysis. Section 2.6 concludes.

2.2 Locus of Control

Since the seminal works of Mincer (1958) and Becker (1964), human capital is defined as the stock of knowledge and personal abilities an individual possesses, and is perceived as a factor of production that can be improved through education, training and experience. The focus usually lies on estimating returns to education, training, experience or cognitive skills (Psacharopoulos, 1981; Card, 1999; Heckman et al., 2006).[2] However, this concept mainly refers to the cognitive abilities of an individual, while more recently other facets of human capital have come to the forefront. Bowles and Gintis (1976) were among the first to point out what seems intuitively obvious: economic success is only partly determined by cognitive abilities and knowledge acquired in schools. Personality, incentive-enhancing preferences and socialization are other important components of human capital (Heckman et al., 2006; Heineck and Anger, 2010).[3] Furthermore, a vast literature in experimental economics is currently emerging, which analyzes the economic impact of risk aversion, reciprocity, self-confidence and time preference (Dohmen et al., 2010; Falk et al., 2006; Frey and Meier, 2004).

We decide to focus on locus of control, one of the measures of personality traits that is prominent also in the economic literature (Heckman et al., 2006; Judge and Bono, 2001; Andrisani, 1977, 1981; Osborne, 2000). Originally, locus of control is a psychological concept, generally attributed to Rotter (1966), that measures the attitude regarding the nature of the causal relationship between one's own behavior and its consequences. In this concept, which is related to self-efficacy, people who believe that they have control over their lives are called *internalizers*. People who believe that fate, luck, or other people determine their lives, are termed *externalizers*. Generally, externalizers (in this taxonomy, the low-ability types) do not have much confidence in their ability to influence their environment, and do not see themselves as responsible for their lives. Therefore, these individuals are generally less likely to trust their own abilities or to push themselves through difficult situations. Conversely, internalizers (the high return personality types) perceive themselves as more capable of altering their economic situation.

[2] See Gebel and Pfeiffer (2010); Pischke and Von Wachter (2008); Lauer and Steiner (2000); Flossmann and Pohlmeier (2006) for estimates of returns to education or skills in the German context.

[3] For an overview of the interrelationships between different psychological and economic concepts, see Borghans et al. (2008).

Mostly on empirical grounds, many studies agree that locus of control affects a variety of economic choices individuals make (behavioral impact). This is particularly true for education decisions, which most researchers find to be highly influenced by locus of control.[4] For instance, Coleman and DeLeire (2003) present a model of locus of control and education decisions, where locus of control is viewed as a behavioral trait that affects education decisions, because it has an impact on personal beliefs about the effect of education on expected earnings. Using the National Education Longitudinal Study (NELS), the authors find locus of control to have a high and significant impact on schooling decisions, as well as on ex-ante expected earnings conditional on schooling. Similarly, recent evidence by Caliendo et al. (2010) on German unemployment data shows that locus of control is a behavioral trait that affects the subjective probability of finding a job, which in turn leads to an increased search effort and higher reservations wages. Contrary to this, using the National Longitudinal Survey of Youth (NLSY), Cebi (2007) concludes that locus of control has a productive impact on labor market outcomes and no effect on education choices.

Evidence on the effect of locus of control on labor market returns is mixed (productive impact). For example, Andrisani (1977), using the National Longitudinal Study (NLS), finds a positive effect of locus of control on several measures of earnings and occupational attainment of young and middle-aged men. Yet, Duncan and Morgan (1981) find mostly non-significant effects of locus of control on the change in hourly earnings of individuals in the Panel Study of Income Dynamics (PSID). To our knowledge, an analysis of the impact of locus of control on wages using German data has only been conducted by Heineck and Anger (2010), as well as by Flossmann et al. (2007), with both studies finding positive effects.[5] We add to this literature by using factor structure models to account for measurement error and endogeneity issues caused by the use of contemporaneous measurements.

[4] Already 40 years ago, the famous Coleman report (Coleman, 1968) reported that locus of control was not only an important predictor of academic performance, but even a more important determinant of educational achievement than any other factor in a student's background (Coleman and DeLeire, 2003).

[5] Furthermore, Gallo et al. (2003) and Uhlendorff (2004) use German data to investigate the impact of locus of control on transitions from unemployment to employment.

2.3 Empirical Model

Consider a simple model where each individual chooses between obtaining higher education or not. Premarket locus of control, as imperfectly measured by a set of response variables, is captured by a latent factor θ, which influences both schooling decisions and labor market outcomes. The concept of locus of control and its potential impact on education decisions and labor market outcomes is explained in Section 6.3.1, while the empirical setup of the model is detailed in Section 2.3.2.

2.3.1 How locus of control impacts education and labor market outcomes

In this section, we present a theoretical framework for how *premarket* locus of control may affect labor market returns. We assume that the role of locus of control for wages is potentially twofold. First, it may indirectly affect wages through its effect on education decisions, and secondly, it may have a direct influence on labor market returns after the education decision is controlled for.

In our study, locus of control is a latent variable, denoted by θ, that is continuously distributed in the range $(-\infty, +\infty)$, where smaller values represent a more external locus and larger values a more internal locus of control. We assume that an individual's psychic costs of education and wage are both functions of θ. Hence, individuals with $\theta \to -\infty$ are likely to have higher psychic costs of education and earn lower wages, while individuals with $\theta \to +\infty$ incur lower costs of obtaining a degree and earn more.

In a typical model of human capital investment, individuals decide on the level of education based on the expected returns to the respective choice, net of the costs associated with this choice. In this framework, locus of control may affect the perceived psychic costs of education, e.g., because individuals with a more external locus of control believe ex ante that they would need to work harder than internalizers to feel well-prepared for the exams (behavioral impact). Furthermore, locus of control may be viewed as a skill with a direct impact on wages, for example because employers value having employees who exhibit a higher locus of control (productive impact).

Assume that there are two education levels, denoted by $S = 0, 1$, and that agents maximize the latent net present value associated with education to make their decision. Let U^*

denote this latent net present value. The arguments of this function will be specified later. Hence, individuals attend higher education, $S = 1$, if:

$$U^* \geq 0,$$

and $S = 0$ otherwise. The latent net present value from obtaining higher education is a function of discounted future earnings and of education costs. If wages w_t^s in period t conditional on schooling s, as well as the costs of education C, can all be modeled in an additively separable manner, we can specify:

$$\begin{aligned} w_t^0 &= X_{wt}\beta_0 + \theta\alpha_0 + \varepsilon_{0t}, \\ w_t^1 &= X_{wt}\beta_1 + \theta\alpha_1 + \varepsilon_{1t}, \\ C &= X_C\beta_C + \theta\alpha_C + \varepsilon_C, \end{aligned}$$

with $\mathrm{E}[\varepsilon_1|X_{wt},\theta] = \mathrm{E}[\varepsilon_0|X_{wt},\theta] = \mathrm{E}[\varepsilon_C|X_C,\theta] = 0$. Here α_s, β_s (with $s \in \{0,1\}$) and α_C, β_C measure the impact of premarket locus of control θ and observable characteristics (X_{wt}, X_C) on wages and education costs, respectively. Since locus of control is determined before the individual enters the labor market, it does not depend on time t in our model. Moreover, ε_{st} and ε_C are random and independent idiosyncratic shocks. The total net present value from education, accounting for the discounted flow of ex post earnings, is then:

$$\begin{aligned} U^*(X_w, X_C, \theta, \delta, t_1) = &\sum_{t=t_1}^{T} \delta^t \left(X_{wt}\beta_1 + \theta\alpha_1 + \varepsilon_{1t}\right) \\ &- \sum_{t=0}^{T} \delta^t \left(X_{wt}\beta_0 + \theta\alpha_0 + \varepsilon_{0t}\right) \\ &- \left(X_C\beta_C + \theta\alpha_C + \varepsilon_C\right), \end{aligned} \tag{2.1}$$

where $X_w = (X_{w1}, \ldots, X_{wT})$, t_1 represents the time required to achieve higher education, T is the life horizon, and δ denotes the discount rate, which for simplicity is assumed to be constant over time.

By differentiating Equation (6.3) with respect to θ, it appears that a ceteris paribus change in locus of control affects education decisions as follows:

$$\frac{\partial U^*(X_w, X_C, \theta, t_1)}{\partial \theta} = \alpha_1 \sum_{t=t_1}^{T} \delta^t - \alpha_0 \sum_{t=0}^{T} \delta^t - \alpha_C.$$

Given that α_1 and α_0 are independent of t, and making use of revealed education choices, our goal is to identify α_1, α_0 and α_C. More precisely, we are investigating whether locus of control enters the education decision and outcomes both directly as a skill, in which case we would have $\alpha_1 > 0$ and $\alpha_0 > 0$, or only indirectly via the costs of education, in which case $\alpha_C < 0$. We cannot identify α_C directly, because we do not observe education costs. However, we can make inference on the overall impact of locus of control on education choices, and given the identification of α_1 and α_0, we can retrieve α_C. More specifically, if we find that $\alpha_1 = \alpha_0 = 0$, we know that any impact of locus of control on education choices must work through α_C.

The empirical model we specify in the next section is an approximation to this very simple theoretical framework. By combining different subsamples and using revealed schooling decisions, we are able to identify the impact of premarket locus of control on wages, and thus to make inferences about its productive or behavioral impact, respectively.

2.3.2 Specification of the Model

To investigate the impact of premarket locus of control on schooling decisions and later outcomes, we use a factor structure model in the spirit of Heckman et al. (2006), where a single latent factor is assumed to capture the latent trait of interest. The overall simultaneous equation model consists of different sets of equations using continuous, dichotomous and ordered response variables. The latent factor is common across all equations, and therefore represents the only source of dependence between the outcomes, conditional on the observed covariates.

Education decision

Each agent is assumed to choose the level of schooling that maximizes her utility. The utility derived from higher education $S^\star$, where higher education is defined as staying in school beyond compulsory education, is supposed to linearly depend on a vector of

personal characteristics X_S and on the latent factor θ:

$$\begin{aligned} S &= \mathbb{1}[S^\star > 0], \\ S^\star &= X_S\beta_S + \theta\alpha_S + \varepsilon_S, \qquad\qquad \varepsilon_S \sim \mathcal{N}(0;\, 1)\,, \end{aligned} \tag{2.2}$$

where β_S denotes the vector of parameters related to personal characteristics, α_S represents the factor loading associated with θ, and ε_S is an idiosyncratic error term assumed to be independent of the covariates and of the latent factor. The indicator function $\mathbb{1}[\cdot]$ is equal to 1 if the corresponding condition is verified, and to 0 otherwise. Conditional on θ, this model is a standard probit when the distribution of the error term is assumed to be standard normal.

Labor market outcomes

Individuals with different levels of schooling become active on different segments of the labor market, where their personal characteristics, as well as their level of locus of control, may be valued differently. Labor market outcomes are modeled as a two-stage process: people first select into the labor market, and then a wage equation is estimated for those actually working. Observed characteristics and locus of control are allowed to play a role in both stages. Estimating the two equations simultaneously makes it possible to correct for potential sample selection bias that might affect the parameters if only the wage equation for working people were estimated (Heckman, 1979).

The labor market participation decision is assumed to be a threshold-crossing model for each level of education $s \in \{0, 1\}$, where the latent utility of working ($E_s^\star$) linearly depends on a set of covariates X_E through a vector of parameters $\beta_{E,s}$, and on the latent factor θ with its associated factor loading $\alpha_{E,s}$:

$$\begin{aligned} E_s &= \mathbb{1}[E_s^\star > 0], \\ E_s^\star &= X_E\beta_{E,s} + \theta\alpha_{E,s} + \varepsilon_{E,s}, \qquad\qquad \varepsilon_{E,s} \sim \mathcal{N}(0;\, 1)\,, \end{aligned} \tag{2.3}$$

The idiosyncratic error term $\varepsilon_{E,s}$ is assumed to be standard normal and independent of X_E and θ for identification purposes. Nevertheless, this equation should not be regarded as a usual employment equation, but rather considered in a broader sense. People participating in the labor market ($E = 1$) are those who are actually active and declare a positive wage, while the group of non-participating people encompasses unemployed people, but

also adult individuals who are not on the market. Therefore, this equation should be interpreted with care,[6] and serves more as a technical means to tackle the selection problem into the sample of people declaring a positive wage. For wages, a log-linear specification with education group specific parameters is assumed:

$$Y_s = X_Y \beta_{Y,s} + \theta \alpha_{Y,s} + \varepsilon_{Y,s} \qquad \text{for } s = 0, 1, \tag{2.4}$$

where Y_s represents the log hourly wage ($\ln w_s$), X_Y is a set of observed covariates with the associated vector of returns $\beta_{Y,s}$, $\alpha_{Y,s}$ denotes the return to locus of control, and $\varepsilon_{Y,s}$ is an idiosyncratic error term such that $\varepsilon_{Y,s} \perp\!\!\!\perp (\theta, X_Y)$. For the specification of the error term, we relax the usual normality assumption by specifying a mixture of h normal distributions with zero mean:

$$\varepsilon_{Y,s} \sim \sum_{j=1}^{h} \pi_{s,j} \, \mathcal{N}\left(\mu_{s,j};\, \omega_{s,j}^2\right), \qquad \mathrm{E}[\varepsilon_{Y,s}] = \sum_{j=1}^{h} \pi_{s,j} \, \mu_{s,j} = 0, \tag{2.5}$$

for $s = 0, 1$, where $\pi_{s,j}$, $\mu_{s,j}$ and $\omega_{s,j}^2$ denote, respectively, the weight, mean and variance of mixture component j. Mixtures of normals are widely used as a flexible semiparametric approach for density estimation (Ferguson, 1983a; Escobar and West, 1995). In our empirical application, we find that a three-component mixture ($h = 3$) for the error term of the wage equation is crucial to achieve a good fit to our data. It allows us to capture unobserved heterogeneity that arises because individuals work in different areas or sectors of modern complex labor markets.[7]

Within this specification, premarket locus of control can affect labor market outcomes both directly and indirectly. The direct effect is measured by the factor loadings $\alpha_{E,s}$ and $\alpha_{Y,s}$, for $s = 0, 1$, while the indirect effect operates through the schooling decision. Two different models are considered. First, we estimate the employment and wage equations without conditioning on education, to capture the total effect of locus of control on wages. To achieve this, individuals from both schooling groups are pooled, and the subscript s is therefore dropped from Equations (2.3) to (2.5). In a second stage, both direct and indirect effects are separately accounted for by specifying the model as stated above. Comparing

[6] Especially for the people who achieve higher education, since in this subsample some individuals who do not participate in the labor market are still enrolled in the education system.

[7] In a frequentist approach, Dagsvik et al. (2011) also find that Gaussian mixtures improve the fit of heavy-tailed log earnings distributions compared to normal distributions.

the results from these two approaches turns out to be instructive to understand through which channels premarket locus of control affects labor market outcomes.

A measurement system for locus of control

In our data, as in most empirical applications, variables measuring latent locus of control come from a psychometric test using Likert scales with a small number of categories. Although techniques to deal with ordinal variables in a multivariate context have a long history in statistics and are now well-documented (see Jöreskog and Moustaki, 2001, for a survey of different approaches), a widespread approach in empirical research consists of ignoring ordinality and treating the manifest items as continuous. This can however distort the results in several ways, especially when the number of categories is limited, and/or the distributions of the answers show high kurtosis.

In this chapter, the ordinal nature of the K measurements is explicitly accounted for by specifying that each individual has a latent level of agreement $M_k^\star$ with the corresponding statement k of the corresponding test, for $k = 1, ..., K$. This latent level of agreement is assumed to linearly depend on some covariates X_M and on the factor θ, and is discretized by a set of cut-points $\{\gamma_k\}$ to produce the observed measurement, with C different alternative ordered answers as follows:

$$\begin{aligned} M_k &= c \quad \text{if } \gamma_{k,c-1} \leq M_k^\star < \gamma_{k,c}, && c = 1, ..., C, \\ M_k^\star &= X_M \beta_{M,k} + \theta \alpha_{M,k} + \varepsilon_{M,k}, && \text{for } k = 1, ..., K, \end{aligned} \tag{2.6}$$

where $\beta_{M,k}$ denotes the vector of parameters associated with X_M, $\alpha_{M,k}$ represents the factor loading, and the idiosyncratic error term $\varepsilon_{M,k}$ is assumed to be standard normal and independent of θ and X_M. Assuming standard normality for the error term is the usual solution adopted to guarantee invariance of the latent response variable to scale transformation. As for the cut-points, they are such that $\gamma_{k,0} = -\infty < \gamma_{k,1} = 0 < ... < \gamma_{k,C-1} < +\infty = \gamma_{k,C}$.

Latent factor for locus of control

To complete the specification of the model, one last distributional assumption is required for the latent factor θ. In a similar framework, Carneiro et al. (2003); Hansen et al. (2004)

achieve nonparametric identification of the latent factors thanks to some independence and support assumptions. When the measurement system consists of a combination of discrete and continuous outcomes, they first nonparametrically identify the joint distribution of the observed and latent measurements, before turning to the identification of the latent factors and error terms using a theorem proposed by Kotlarski (1967) and extended in Cunha et al. (2010). In our case, this identification strategy cannot be applied, insofar as the measurements are all discrete. Nonparametric identification of the latent factor distribution, as well as of the error term distributions, would only be possible if we first managed to nonparametrically identify the joint distribution of the latent measurements. However, the lack of variability and of exclusion restrictions for each measurement make nonparametric identification and the use of more flexible distributional assumptions such as mixtures impossible. For these reasons, and for the sake of simplicity, we specify a normal distribution and make the following independence assumption:

$$\theta \sim \mathcal{N}\left(0; \sigma_\theta^2\right), \qquad \theta \perp\!\!\!\perp (X, \varepsilon),$$

where $X = (X_S, X_E, X_Y, X_M)$ and $\varepsilon = (\varepsilon_S, \{\varepsilon_{E,s}\}, \{\varepsilon_{Y,s}\}, \{\varepsilon_{M,k}\})$.
Since the variance of the latent factor is not constrained, we need to impose one restriction to set the scale of θ. For this purpose of identification, we fix one of the factor loadings to a given value in the measurement system.

2.4 Estimation Strategy

In this section, we present the identification strategy that relies on data set combination in Section 2.4.1, as well as our estimation method and data in Section 6.5. The parameters of interest are simulated through the implementation of Bayesian Markov chain Monte Carlo techniques.

2.4.1 Combining data sets to identify the model likelihood

Ideally, we would have access to a data set where individuals are observed at different periods of their life cycle. The likelihood of the model for such an hypothetical sample

can be expressed as

$$L(\psi|S,E,Y,M,X) = \int_{\Theta} \prod_{s=0}^{1} \left[\Pr(S=s|X_S,\theta,\psi)\, f(E_s|X_E,\theta,\psi) f(Y_s|X_Y,\theta,\psi)\right]^{\mathbb{1}[S=s]} \times \prod_{k=1}^{K} f(M_k|X_M,\theta,\psi)\ \mathrm{d}F_\theta(\theta), \tag{2.7}$$

where ψ represents the vector containing all model parameters, $f(\cdot)$ invariantly denotes a density function, and $F_\theta(\cdot)$ is the cumulative distribution function (cdf) of the latent factor θ on the support Θ. In our case, this would require information on people's labor market outcomes and personal background, as well as on their premarket locus of control. Estimation based on the likelihood (2.7) would be straightforward.

Unfortunately, the structure of the GSOEP only offers this opportunity for a subsample of the population, which turns out to be too small to conduct any relevant analysis. Although the GSOEP is a longitudinal study, youth are surveyed since 2000 only, and many of them still have not entered the labor market in 2008. We therefore have to face a major dilemma: on the one hand, we have a large data set of working-age people (adult sample), but without any information on their locus of control at the time of schooling. On the other hand, a sample of 17-year-olds is available (youth sample), including premarket locus of control measurements, but labor market outcomes only for a very small group of mostly low-educated individuals. The adult and the youth samples can nevertheless be combined to overcome this problem. We rely on an idea implemented in Cunha et al. (2005), which consists of identifying one part of the likelihood in each subsample, getting rid of the unobserved response variables by integrating them out of the likelihood.

To understand the mechanisms of the data set combination, consider the following sketch of proof. First, derive the contribution to the likelihood of a person with higher education. Since her future labor market participation and wage cannot be observed, they are integrated out to provide

$$\begin{aligned}
&\int_{\Theta} \Pr(S=1|X_S,\theta,\psi) \left\{ \iint f(E_1|X_E,\theta,\psi) f(Y_1|X_Y,\theta,\psi)\, \mathrm{d}F_{E_1}(E_1)\, \mathrm{d}F_{Y_1}(Y_1) \right\} \\
&\qquad \times \prod_{k=1}^{K} f(M_k|X_M,\theta,\psi)\, \mathrm{d}F_\theta(\theta) \\
&= \int_{\Theta} \Pr(S=1|X_S,\theta,\psi) \prod_{k=1}^{K} f(M_k|X_M,\theta,\psi)\ \mathrm{d}F_\theta(\theta),
\end{aligned}$$

where $F_W(\cdot)$ represents the cdf of the corresponding random variable W. As a consequence, the parameters of the measurement system and of the schooling equation can be identified from the youth sample. However, due to the small sample size of youth who already earn a wage on the labor market, identification and estimation of the parameters of the labor market participation and wage equations from this sample is impossible.

In a similar fashion, consider a person without higher education from the adult sample, whose measurements for premarket locus of control are not observed. Her contribution to the likelihood is

$$\begin{aligned}
&\int_\Theta \Pr(S=0|X_S,\theta,\psi)\, f(E_0|X_E,\theta,\psi) f(Y_0|X_Y,\theta,\psi) \\
&\qquad\qquad \times \left\{ \prod_{k=1}^{K} \int f(M_k|X_M,\theta,\psi)\, \mathrm{d}F_{M_k}(M_k) \right\} \mathrm{d}F_\theta(\theta) \\
&= \int_\Theta \Pr(S=0|X_S,\theta,\psi)\, f(E_0|X_E,\theta,\psi) f(Y_0|X_Y,\theta,\psi)\ \mathrm{d}F_\theta(\theta),
\end{aligned}$$

and is obtained by integrating out the locus of control measures. Full identification of the model is clearly infeasible in this subsample, since no observations on premarket locus of control are available for the adults. However, since we are combining the two data sets and estimating the overall model simultaneously, the distribution of the latent factor is already identified from the youth sample.

Full identification of the model rests on the education equation, which is the only source of common information for most of the sample, and therefore the bridge between the two samples. Although our model can in theory be identified from two non-overlapping samples of youth and adults, in practice we found it helpful to use all available information—i.e., measurement, schooling and labor market information—for the small sample of individuals for whom both labor market outcomes and locus of control measurements are available.

2.4.2 Estimation

A fully Bayesian approach is used for the estimation of our model. Since the equations are independent once θ is conditioned on, the estimation can be divided into several pieces, and MCMC methods are particularly suited for this kind of problem. In the wake of Cunha et al. (2005); Carneiro et al. (2003); Hansen et al. (2004), we use a Gibbs sampler

that sequentially draws the parameters of interest from their respective conditional distributions, using flat priors to remain as general as possible.[8]

Data augmentation procedures (Tanner and Wong, 1987) make it possible to simulate the latent outcomes of the measurement system, of the schooling and labor market participation equations, as well as the latent factor θ.[9] Besides the practical convenience of the approach, augmenting the observed data with the latent variables has another major advantage in our case: the simulated latent factors and outcomes can be saved during the sampling process, and used for post-processing analyses, such as simulations.[10] In Section 2.5.2 for instance, these simulated variables are used to assess the fit of the model, and to conduct some formal tests.

Bayesian inference for ordinal variable models can be challenging. Slow convergence and high autocorrelation of the parameter chains are typical symptoms of the algorithm failing to cover the entire posterior distribution of the parameters. As noted by Cowles (1996), the high correlation between the cut-points and the latent response variable results in a poor mixing of the Markov chain for the parameters of Equation (2.6). In the end, this can lead to overinflated standard errors of the parameters, or even worse, to wrong estimates (in terms of bias) if the chain is not long enough to provide a representative sample of the conditional distribution. To remedy this problem, several technical improvements have been proposed.[11] We opt for the group transformation approach introduced by Liu and Sabatti (2000), which speeds up convergence and enhances the mixing of the chain, while being less computationally burdensome than other methods. We run a chain of 1,010,000 iterations for each gender. After a burn-in period of 10,000 iterations, 10,000 iterations are saved every 100^{th} sweep of the Gibbs sampler for post-processing inference. We observe a fast convergence to the stationary distribution, and a good mixing of the chain thanks to the implementation of the group transformation.

[8] For technical details on the Gibbs sampler in this framework, see BLINDED, 2010 where all posterior distributions are derived.

[9] Data augmentation procedures are increasingly used in applied labor market and education research (for recent examples see Horny et al., 2009; Koop and Tobias, 2004; Li, 2006).

[10] See van Dyk and Meng (2001) for a review of data augmentation.

[11] Cowles (1996) introduces a Hastings-within-Gibbs step in the algorithm to draw the cut-points and the latent response variable simultaneously, while Nandram and Chen (1996) propose a simple reparameterization that proves to be particularly effective, especially in the three-category case.

Sample construction

We draw a combined sample of 1,534 youth (age 17-24) and 1,192 'young adults' (age 26-35) from recent waves of the GSOEP. The special feature of the youth sample is that for these youth, premarket measures of locus of control were administered when they were 17 years of age. In the German education system, individuals decide at around the age of 17 whether to finish their studies with a vocational high school certificate, or to continue their schooling with academic high school credentials. Only the latter entitles agents to attend higher education. Hence, our binary education variable reflects this choice of obtaining a vocational or an academic high school degree. Summary statistics of the education variable in the two samples are presented in Table 2.5. For a small part of our youth sample (about 280 individuals), also wage and employment information is available. However, because these individuals can be at most 24 years of age, most of them did not achieve higher education. Furthermore, separate estimations by gender and schooling considerably reduce the available sample size. Hence, as explained in the previous section, we augment the youth sample with a second sample of young adults, whose education and labor market outcomes can be assumed to be generated by the same data generating process. Summary statistics on wages and employment participation of the combined sample can be found in Table 2.6. The table displays that males earn higher wages than females, and that the observed wage gap between high and low educated individuals is higher for males than for females. The low levels of labor market participation arise because many individuals still participate in education or training. To fully account for gender differences in the impact of locus of control on education decisions and outcomes, all estimates are obtained separately for males and females.

In order to be able to identify different parts of the likelihood from different samples, we make the assumption that both samples are generated by the same underlying data generating process (DGP). Specifically, we assume that if premarket locus of control and labor market outcomes were available for both youths and adults, we would expect to obtain the same estimated coefficients. This assumption is restrictive in the sense that Table 2.5 shows that among the youth sample, there is a slightly higher fraction of highly educated individuals. In order to deal with this problem, we include age and cohort dummies as covariates in the education, employment and wage equations, so as to capture possible time trends or cohort effects.

Locus of control measurements

Table 2.1: Locus of control questions and descriptive statistics

		Males		Females	
	Variables	**Mean**	**SD**	**Mean**	**SD**
Q1	My life's course depends on me	3.55	0.63	3.51	0.59
Q2	I have not achieved what I deserve	2.05	0.85	1.92	0.79
Q3	Success is a matter of fate or luck	2.22	0.81	2.29	0.77
Q4	Others decide about my life	2.18	0.83	2.12	0.83
Q5	Success is a matter of hard work	3.48	0.62	3.51	0.57
Q6	In case of difficulties, doubt about own abilities	2.08	0.81	2.31	0.85
Q7	Possibilities in life depend on social conditions	2.69	0.78	2.72	0.75
Q8	Abilities are more important than effort	3.02	0.71	3.05	0.69
Q9	Little control over what happens to me	1.92	0.75	1.95	0.76
Q10	Social involvement can help influence social conditions	2.48	0.87	2.51	0.77
	# Observations	760		774	

In the GSOEP youth questionnaire, locus of control is measured by a 10-item questionnaire. Each question is answered on a Likert scale ranging from 1 ("disagree completely") to 4 ("agree completely"). Table 2.1 gives an overview of the questions and items we use. We check whether, given these measurements, locus of control can indeed be represented by a single factor. Conducting a principal component analysis, and calculating the eigenvalues of the correlation matrix, we find two eigenvalues larger than 1. Hence, the Kaiser criterion (eigenvalue$<$1) is violated. However, the scree plot analysis displayed in Figure 2.6 reveals an early flattening of the curve, suggesting no more than one or two underlying factors. Furthermore, locus of control is usually conceptualized as referring to a unidimensional continuum, ranging from external to internal. Hence, we think that we are making a reasonable decision by extracting a single factor. A scatter plot of the respective factor loadings (Figure 2.7), with the first two principal factors on the axis, shows that some items load very highly on the extracted locus of control factor (factor 1), while some other items have a loading close to zero (Q1, Q5, Q8 and Q10). Furthermore, the items with a close to zero loading are items that capture an internal attitude, while the other items mostly capture the external dimension of locus of control. Consequently, we can draw two conclusions from this exploratory factor analysis. First, researchers who use an index, constructed for example as the standardized mean of the items, instead of a latent factor, force each of the measurement items to enter the index with an equal weight. Doing this yields a locus of control measure that is flawed by measurement error, and the

coefficients are likely to be biased downward due to attenuation bias. Second, in this chapter we mostly capture the external attitude dimension of locus of control. For ease of interpretation, in our empirical application we normalize the model such that lower scores of the latent factor are associated with an external locus of control, and higher scores with an internal locus of control. To ensure that our results are not distorted by the inclusion of those items that have a low loading on the locus of control factor, we have conducted robustness checks using only those items loading highly on the first factor. We find that the use of the externalizing items only does not have a major impact on the results.[12]

Table 2.2: Samples and included covariates for the measurement system, education, employment and wage equations

	Type	**Meas.**	**Educ.**	**Empl.**	**Wage**
Samples					
Youth sample		✓	✓	(✓)	(✓)
Adult sample		—	✓	✓	✓
Covariates					
Number of siblings	D	✓	—	—	—
% of time in broken family	C	✓	✓	—	—
Father dropout	B	✓	✓	✓	✓
Father grammar school	B	✓	✓	✓	✓
Mother dropout	B	✓	✓	—	—
Mother grammar school	B	✓	✓	—	—
Region: North	B	✓	✓	✓	✓
Region: South	B	✓	✓	✓	✓
Childhood in large city	B	✓	✓	✓	✓
Childhood in medium city	B	✓	✓	✓	✓
Childhood in small city	B	✓	✓	✓	✓
Track recommendation (highest)	B	✓	—	—	—
Track recommendation (lowest)	B	✓	—	—	—
Local unemployment rate	C	—	—	✓	✓
Local unemployment rate (edu)	C	—	✓	—	—
Age of individual	C	—	—	✓	✓
Cohort 26/30	B	—	✓	—	✓
Cohort 31/35	B	—	✓	—	✓
Married	B	—	—	✓	✓
Number of Children	C	—	—	✓	✓

Note: B = Binary, C = Continuous, D = Discrete. *Local unemployment rate (edu)* denotes the local unemployment rate when the education decision is made.

[12]Results of the robustness check using only the externalizing items can be obtained from the authors upon request.

Covariates

Table 2.2 summarizes the covariates used for our analysis, and also shows how the two samples are linked by the schooling equation. To account for family background, socioeconomic status and labor market conditions, we control for a large range of background variables, as well as for local unemployment rates at the time of education decisions and labor market outcomes, respectively. In addition, Germany has an education system where tracking already takes place after the fourth grade. Hence, to proxy cognitive skills, and to account for the fact that these cognitive skills might affect the items revealing premarket locus of control, we include the primary school teacher track recommendation as a control variable in the measurement system. Because locus of control is estimated from the residual variance net of covariates in the measurement system, covariates included in the measurement equation are a means to purge locus of control of their influence. However, the inclusion of track recommendation only proxies cognitive skills and the resulting track type. It cannot account for other conflicting effects such as school quality. Hence, locus of control, as identified in this chapter, only captures *premarket* locus of control, and not necessarily *pre-compulsory-school* locus of control. Thus we control for track recommendation, parental education and a large set of other background variables to capture school quality, home investment and cognitive ability. Summary statistics of control variables in the measurement and outcome equations can be found in Tables 2.7 and 2.8.[13]

2.5 Empirical Results

The results are presented and discussed in two stages. We first provide a description of the main findings in Section 2.5.1, with an emphasis on the statistical significance of the impact of locus of control on the different outcomes, and on the fit of our model. Then, we gain more insights in Section 2.5.2 by conducting some simulations that make it possible to better grasp the magnitude of the impact of locus of control.

[13]A detailed description of the coding of all variables can be obtained from the authors upon request.

2.5.1 MCMC results

Factor loadings. The factor loadings express how the different measurements and outcomes are affected by the latent factor. The larger the magnitude of the loadings, the higher the contribution of the corresponding measurements to the distribution of the latent factor. In the education, employment and wage equations, the loadings measure the impact of the factor on the respective outcomes. Cross-model comparisons should however be carefully done: the factor loadings of the different models cannot be directly compared, as their magnitude and their sign depend on the normalization retained to set the scale of the factor. We normalize the factor loading of the fourth indicator to -1 in all models, which is a way of anchoring the factor distribution in a real measurement (Cunha and Heckman, 2008).[14] However, contrary to Cunha and Heckman (2008), who anchor the factor in earnings, we cannot give an interpretable metric to the latent factor, because of the ordinal nature of the measurement. Moreover, the respective item of the questionnaire used for the normalization might be perceived differently by males and females, and gender comparisons are therefore not straightforward.

Table 2.3 summarizes the estimation results for the factor loadings of the different models. The results of the measurement system are in line with our expectations. Typical questions associated with an external locus of control such as 'Success is a matter of fate or luck' (Q3) or 'I have not achieved what I deserve' (Q2) have negative factor loadings, whereas statements reflecting an internal locus of control, such as 'My life's course depends on me' (Q1), have a positive factor loading. Also, the heterogeneity of these factor loadings is worth noting, as well as the fact that some of them are not significantly different from zero.

In the outcome system of equations, the factor loading of the education equation is always significant and positive, indicating an actual impact of locus of control. When we do not control for education [columns (1) and (3)], wages appear to be affected by locus of control, whereas this impact vanishes when education is controlled for [columns (2) and (4)]. Hence, with respect to the theoretical framework laid out in Section 6.3.1, we can conclude that the impact of premarket locus of control on w_t^0 and w_t^1, denoted by α_0 and α_1 respectively, is zero. However, we find that locus of control does have an impact

[14] The fourth indicator is a typical externalizers' statement, hence the normalization to a negative integer.

Table 2.3: Factor loadings of the model estimated by conditioning labor market outcomes on education [(2) and (4)] and without conditioning on education [(1) and (3)]

	Males				**Females**			
	(1)		(2)		(3)		(4)	
Measurement system: Locus of control items								
Q1	0.354***	(0.087)	0.364***	(0.086)	0.423***	(0.095)	0.440***	(0.101)
Q2	-0.735***	(0.119)	-0.729***	(0.116)	-0.895***	(0.132)	-0.938***	(0.143)
Q3	-0.741***	(0.118)	-0.743***	(0.116)	-0.619***	(0.107)	-0.650***	(0.113)
Q4	-1.000	—	-1.000	—	-1.000	—	-1.000	—
Q5	0.013	(0.074)	0.024	(0.075)	0.026	(0.085)	0.025	(0.089)
Q6	-0.640***	(0.108)	-0.605***	(0.102)	-0.890***	(0.134)	-0.916***	(0.139)
Q7	-0.559***	(0.099)	-0.565***	(0.099)	-0.581***	(0.105)	-0.617***	(0.112)
Q8	-0.195***	(0.072)	-0.197***	(0.072)	-0.107*	(0.078)	-0.112*	(0.082)
Q9	-1.045***	(0.175)	-1.035***	(0.175)	-1.781***	(0.309)	-1.858***	(0.332)
Q10	-0.122**	(0.067)	-0.140**	(0.068)	0.143**	(0.078)	0.146**	(0.080)
Education choice								
S	0.634***	(0.134)	0.404***	(0.118)	0.444***	(0.123)	0.364***	(0.127)
Labor market participation								
E	0.055	(0.136)			-0.021	(0.131)		
E_0			0.757***	(0.287)			0.357**	(0.222)
E_1			-0.126	(0.331)			-0.268	(0.286)
log Wages								
Y	0.181***	(0.041)			0.121***	(0.048)		
Y_0			0.007	(0.060)			0.058	(0.064)
Y_1			-0.072	(0.086)			0.020	(0.087)
Variance of the latent factor								
σ^2_θ	0.635***	(0.138)	0.622***	(0.135)	0.446***	(0.092)	0.411***	(0.088)

Notes: Factor loading of item 4 (statement reflecting an external locus of control) fixed to -1 to set the scale of the latent factor. Standard errors in brackets. Significance check: */**/*** if zero lies outside the 90%/95%/99% confidence interval of the posterior distribution of the corresponding parameter.

on education decisions ($P(S = 1)$), and thus on wages in the end. Hence, reverting to Equation (6.3), we can conclude that locus of control does not affect education decisions via higher expected wages (α_0,α_1), but instead through its impact on the cost of education α_C.

So far, no firm conclusions have been made as to the magnitude of the impact of locus of control on education decisions and overall wages. In the following Section 2.5.2, the simulations we conduct make it possible to unravel and quantify the actual impact of locus of control on the different outcomes of interest.

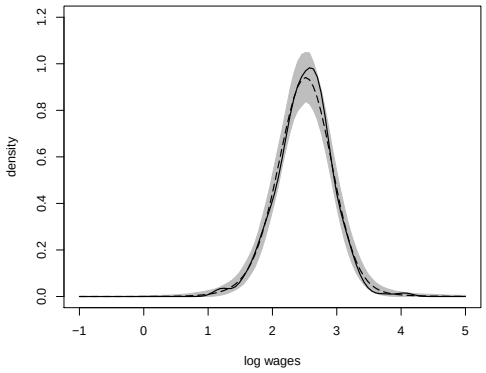
Model fit to actual data. Our model provides a good fit to the data, and especially to the distribution of wages. Figure 2.1 displays the observed distribution of wages, along with their posterior predictive distribution for the different specifications. The actual distribution is quite well approximated by the posterior predictive distribution, particularly in the case where the two schooling groups are pooled for the estimation of the wage equation (panels 2.1a and 2.1b). When the wage equation is estimated by level of schooling (panels 2.1c, 2.1d , 2.1e and 2.1f), the fit is somewhat less good. Nevertheless, the Kolmogorov-Smirnov tests we conduct to compare the actual distribution and the posterior predictive distribution never reject the null hypothesis of equal distribution. This result is in great part due to the use of normal mixtures for the error term, allowing for a flexible approximation of the true distribution.

To assess the goodness of fit to the education decision, Table 2.22 shows the proportion of correct predictions of education achievement for each decile of the latent factor distribution. The fit appears good overall, especially for the lower deciles of the distribution.

2.5.2 Simulation of the model

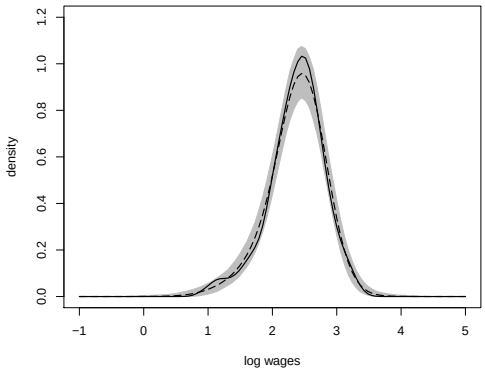
To shed more light on the implications of our model, we need to go beyond the mere interpretation of the factor loadings. Their statistical significance reveals an impact of locus of control on the outcomes, but is quite uninformative regarding the magnitude of this impact (McCloskey and Ziliak, 1996; Ziliak and McCloskey, 2004). Since the effects of premarket locus of control are intertwined and potentially operate through different channels on wages, the best way to understand our model is to simulate it.

Figure 2.1: Goodness-of-fit check for wages: posterior predictive (dashed) vs. actual distribution (solid) and Kolmogorov-Smirnov test for equal distributions.



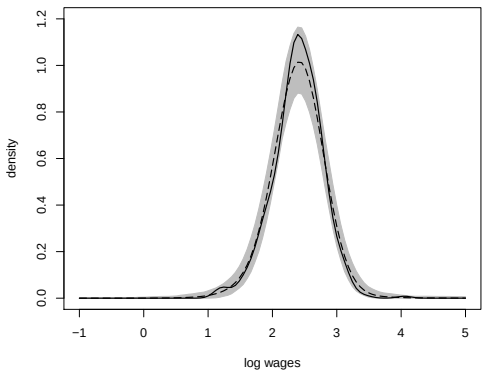
KS-test: 0.026 (0.745)

(a) Males, all education levels



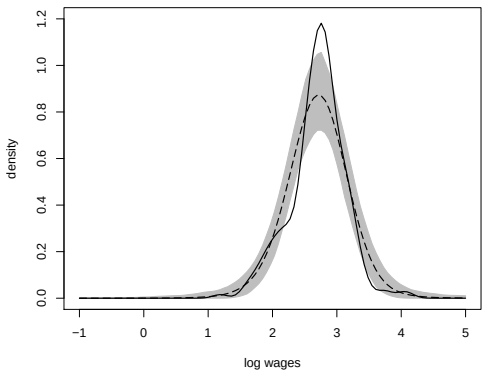
KS-test: 0.027 (0.799)

(b) Females, all education levels



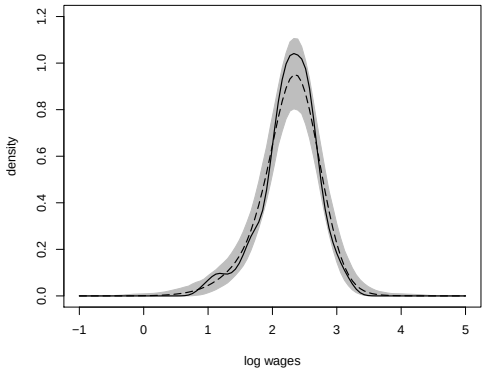
KS-test: 0.042 (0.418)

(c) Males, no higher education



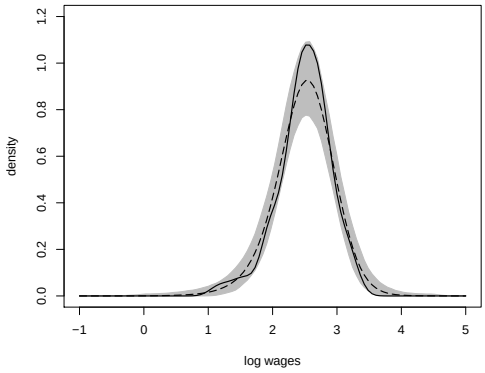
KS-test: 0.085 (0.072)

(d) Males, with higher education



KS-test: 0.047 (0.479)

(e) Females, no higher education



KS-test: 0.055 (0.398)

(f) Females, with higher education

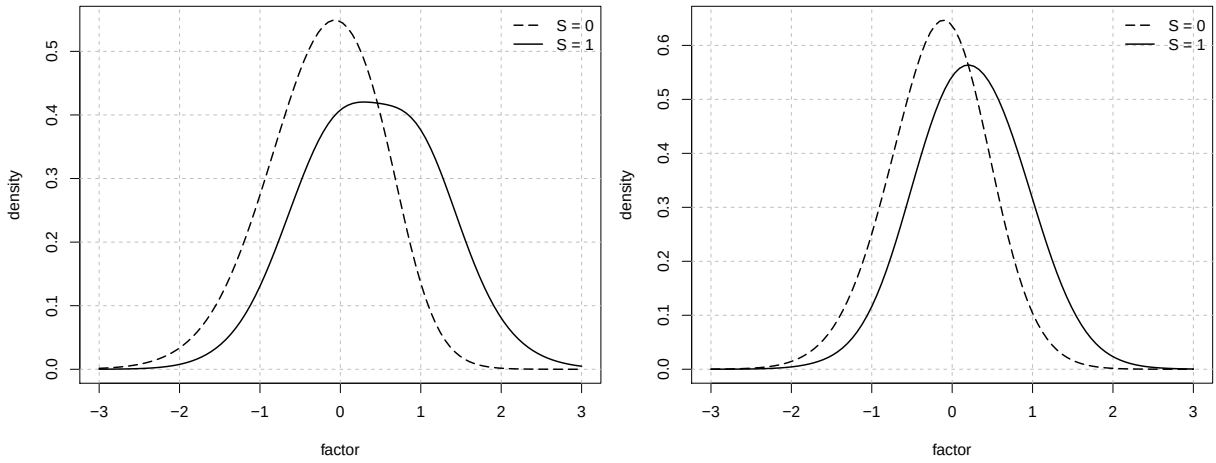
Notes: Model estimated by conditioning labor market outcomes on education (panels 2.1c to 2.1f) and without conditioning on education (panels 2.1a and 2.1b). Kernel density estimation implemented using a Gaussian kernel with bandwidth selected using Silverman's rule of thumb (Silverman, 1986) with the variation proposed by Scott (1992). Wages predicted from their posterior distribution using 1,000 replications of the sample. Shaded area represents 95% confidence interval of posterior predictive distribution. Kolmogorov-Smirnov test: Two-sample KS-test with null hypothesis that the actual sample and the posterior predictive sample have the same distribution. p-values in brackets. Exact p-values could not be computed due to ties in the distribution of actual wages.

Table 2.4: Goodness-of-fit check: proportion of correct predictions of education achievement for each decile of the latent factor distribution

		Deciles of latent factor distribution									
		10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
Males	(1)	0.827	0.816	0.804	0.785	0.764	0.737	0.708	0.673	0.645	0.699
		(0.033)	(0.031)	(0.032)	(0.033)	(0.035)	(0.035)	(0.035)	(0.038)	(0.039)	(0.044)
	(2)	0.788	0.785	0.777	0.762	0.745	0.728	0.711	0.691	0.671	0.669
		(0.033)	(0.033)	(0.034)	(0.034)	(0.035)	(0.036)	(0.036)	(0.039)	(0.040)	(0.040)
Females	(3)	0.771	0.746	0.728	0.714	0.702	0.689	0.677	0.667	0.667	0.692
		(0.033)	(0.035)	(0.036)	(0.037)	(0.037)	(0.036)	(0.037)	(0.038)	(0.041)	(0.040)
	(4)	0.756	0.736	0.722	0.709	0.698	0.689	0.678	0.666	0.659	0.675
		(0.034)	(0.037)	(0.037)	(0.037)	(0.037)	(0.039)	(0.039)	(0.039)	(0.039)	(0.039)

Notes: Model estimated by conditioning labor market outcomes on education [(2) and (4)] and without conditioning [(1) and (3)]. Proportions of correct predictions computed for each MCMC replication, corresponding means and standard errors (in brackets) are reported.

Figure 2.2: Latent factor distribution by levels of education: people with higher education ($S = 1$) and without higher education ($S = 0$).



KS-test: 0.298 (0.000) KS-test: 0.242 (0.000)

(a) Males **(b)** Females

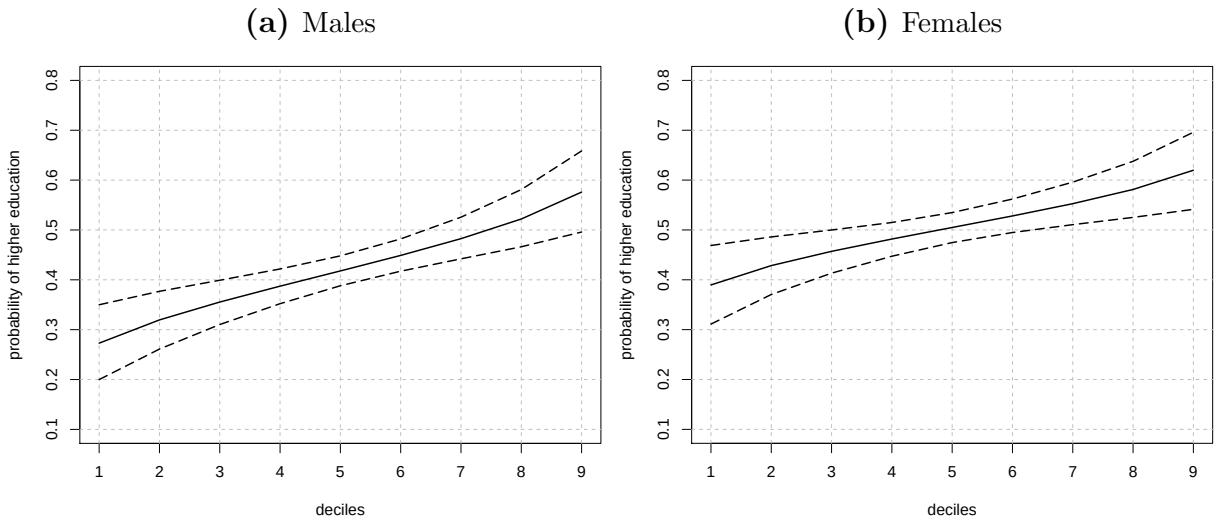
Notes: Simulation from the estimates of the model using 1,000 replications of the posterior sample. Model estimated without conditioning labor market outcomes on education. Predicted levels of education used ($\Pr(S = 1) > .5$). Kernel density estimation implemented using a Gaussian kernel with bandwidth selected using Silverman's rule of thumb (Silverman, 1986) with the variation proposed by Scott (1992). Kolmogorov-Smirnov test: Two-sample KS-test with null hypothesis that the two distributions are the same. p-values in brackets. Exact p-values could not be computed due to ties in the distribution of the latent factor.

Figure 2.2 plots the estimated posterior distribution of the latent factor by levels of education, and shows that people who achieve higher education have a more internal locus of control. For males, the gap between the two schooling groups is even wider, revealing some gender differences in the way locus of control influences education decisions. The Kolmogorov-Smirnov test confirms that the discrepancy between the two distributions is statistically significant for both genders.

To get more insight on the impact of premarket locus of control on later outcomes, we can investigate how the wage of a given individual would be affected if she were exogenously moved along the distribution of the latent factor, for a given set of observed characteristics X_Y (Heckman et al., 2006). For this purpose, we compute the expected wage for different quantiles of the distribution of the factor, conditional on a given set of covariates X_Y. The Gibbs algorithm we implement to estimate our model generates a sample of the model parameters from their conditional distribution that can be used as follows to approximate the expected wage for each quantile q_θ of the factor distribution:

$$\frac{1}{M}\sum_{m=1}^{M}\left(X_Y\beta_Y^{(m)} + q_\theta^{(m)}\alpha_Y^{(m)}\right),$$

Figure 2.3: Higher education probability for each decile of the factor distribution



Notes: Simulation from the estimates of the model using 10,000 replications of the posterior sample. Model estimated conditioning labor market outcomes on education. 95% confidence band between dashed lines.

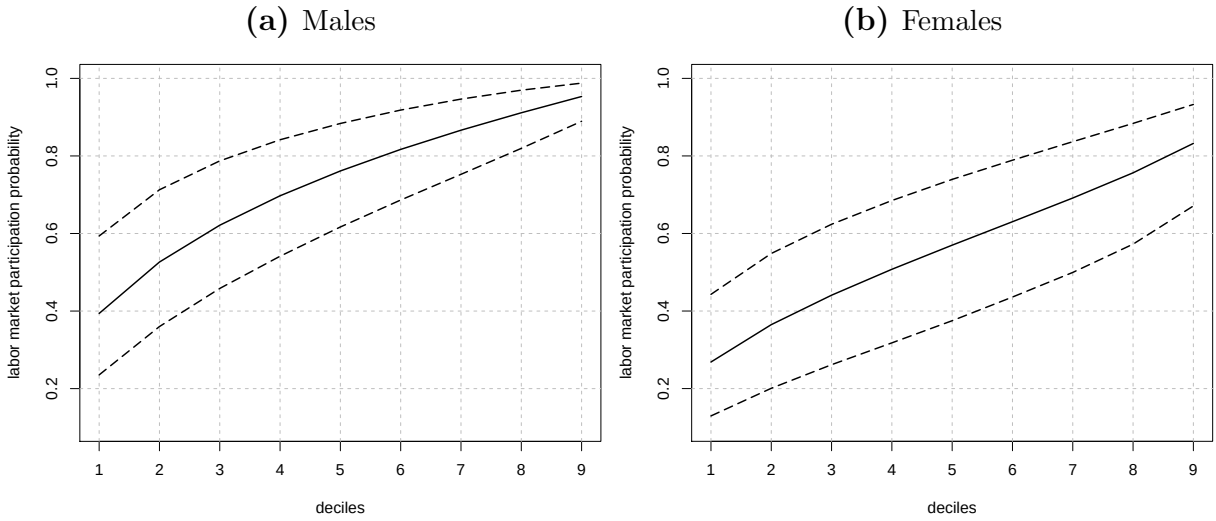
for a set of M simulated parameters $(\beta_Y^{(1)}, \alpha_Y^{(1)}), \ldots, (\beta_Y^{(M)}, \alpha_Y^{(M)})$. The quantile of the latent factor $q_\theta^{(m)}$ also has a superscript (m), since it depends on the variance of the factor $\sigma_\theta^{2(m)}$, and therefore varies during the MCMC sampling. Similarly, the schooling and labor market participation probabilities in the q^{th} quantile of the latent factor distribution can be approximated by:

$$\frac{1}{M}\sum_{m=1}^{M}\Phi\Big(X_S\beta_S^{(m)} + q_\theta^{(m)}\alpha_S^{(m)}\Big), \qquad \frac{1}{M}\sum_{m=1}^{M}\Phi\Big(X_E\beta_E^{(m)} + q_\theta^{(m)}\alpha_E^{(m)}\Big),$$

respectively, where $\Phi(\cdot)$ denotes the cdf of the standard normal distribution. More specifically, the simulations we present rely on the deciles of the distribution. In the following, our simulations are performed for the mean individual of the corresponding sample.

From Figure 2.3, locus of control appears to have a large impact on the schooling decision, since moving the mean individual from the first to the last decile of the distribution results in a 0.30 point increase in the probability of achieving higher education for males, and a 0.23 point increase for females. Similarly, Figure 2.4 shows that in the group of people who did not achieve higher education, locus of control has a huge impact on labor market participation. This effect is more or less linear for females, whereas for males the concavity of the curve indicates that people in the low deciles are more affected than people in the

Figure 2.4: Labor market participation for each decile of the factor distribution



Notes: Simulation from the estimates of the model using 10,000 replications of the posterior sample. Model estimated conditioning labor market outcomes on education. 95% confidence band between dashed lines.

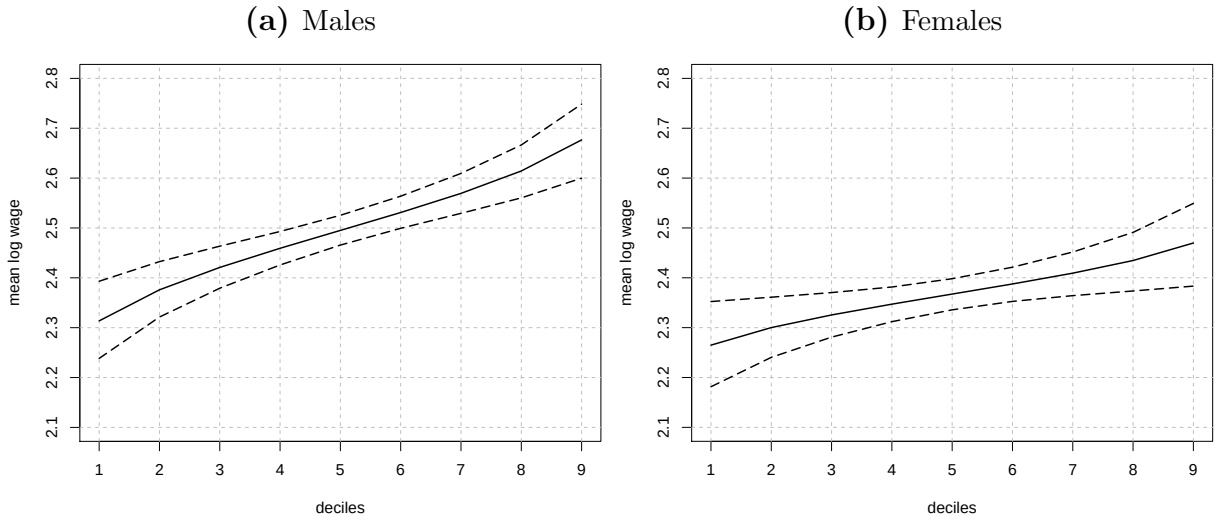
higher deciles of the distribution. Concerning wages, Figure 2.5 shows that if the mean individual could be moved exogenously from the first to the 9^{th} decile of the locus of control distribution, this would corresponds to an increase in hourly wages of roughly 4.40 Euros for the mean male individual, and of roughly 2.20 Euros for the mean female individual.

At first sight, the effect of locus of control on education choice and labor market outcomes seems large. For instance, the mean male individual would earn 36% more in the last decile than in the first one. However, it is unrealistic to see an individual move all the way across the distribution. People are more likely to make small moves from one decile to the adjacent ones, and Figures 2.3 to 2.5 show that in the middle of the distribution, the locus of control effect is much smaller.

2.5.3 Some remarks on the results

In summary, we find an effect of locus of control on schooling probabilities, where males are more affected than females. Moving the mean individual in the distribution of the latent factor substantially changes her/his wage. However, this overall effect only operates through the channel of schooling. This finding that *premarket* locus of control influences schooling is in line with Coleman and DeLeire (2003), although in their paper

Figure 2.5: Mean log wage for each decile of the factor distribution



Notes: Simulation from the estimates of the model using 10,000 replications of the posterior sample. Model estimated without conditioning labor market outcomes on education. 95% confidence band between dashed lines.

the mechanism through which locus of control affects schooling is different, as it only works through wage expectations.

Our results seem somewhat contrary to the more direct link between locus of control and wages that been found in some of the literature (Heckman et al., 2006; Heineck and Anger, 2010). Three different answers can be put forward to address this apparent contradiction. First, the term 'noncognitive skills' is very often used as a generic expression encompassing a lot of different personal abilities and traits, sometimes leading to confusion. A comparison of results is possible only if the same concept is used. For instance, Heckman et al. (2006) find a significant effect of noncognitive skills on wages. However, they use a single underlying factor for noncognitive skills constructed from two psychometric tests, namely the Rosenberg self-esteem scale and the Rotter scale. This composite factor thus captures a different dimension than our factor, especially since it loads more on the self-esteem scale than on the locus of control scale in their empirical study. Second, and more importantly, we focus on *premarket* locus of control as a measure of locus of control that is independent of labor market experience. As a consequence, our findings differ from the results presented by Heineck and Anger (2010) who find a strong and significant impact of locus of control on wages, even after controlling for education. One reason could be that the authors do not estimate separate models by education level. More likely, however, the difference in results arises because of the use of contemporaneous measurements in their study, while

we focus on the impact of *premarket* locus of control. Third, we only look at a sample of young labor market entrants. At this stage, wage setting is likely to be merely a function of formal qualifications. Hence, only after individuals have entered the labor market, a complex dynamic interaction process begins. While working on-the-job, individuals learn about their abilities, while at the same time employers adapt their knowledge about an individual's locus of control. As a result, a positive interdependence between locus of control and wages may arise (such as the one found by Heineck and Anger, 2010). Additional analyses not displayed in this chapter show that the correlation between locus of control and wages does indeed increase with age and experience of the agents. Whether this is the result of reverse causality or learning of employers is an interesting topic left for future research. One explanation may be that although early locus of control does not influence wages directly, it may influence late locus of control which in turn is directly rewarded on the labor market. We leave it for future research to find out whether there exists a constant and invariable component to personality traits in general, and to locus of control in particular. Such a component may be extracted using dynamic factor models, and would require repeated measurements of locus of control over large parts of the life-cycle.

2.6 Conclusion

In this chapter, we use Bayesian factor structure models to investigate how locus of control influences education decisions and wages. Using advanced econometric methods, we show that such recent methods can serve as a solution measurement error and endogeneity problems, especially if researchers are confronted with truncated life cycle data, as is very often the case for research at the intersection of psychology and economics.

We establish that an individual's *premarket* locus of control substantially raises the probability of choosing higher education. We also show that locus of control influences wages through schooling, but that there is no direct impact on wages once schooling is controlled for. Thus, in a framework where schooling decisions depend on relative lifetime earnings returns for each schooling level, net of the costs of obtaining either level of education, we can infer from our results that *premarket* locus of control, as measured at the age of 17, is not directly rewarded as a skill on the labor market. Instead, it is a personality trait that influences the non-pecuniary costs of education. Our work conveys

important policy implications. If some personality traits, such as locus of control, influence the cost of education but not outcomes directly, these individual characteristics may keep individuals from studying who, once they reach the labor market, are no less successful than other individuals. If these individuals are at high risk of dropping out of school, early personality tests and targeted mentoring of students with an external locus of control are a means to countervail skill shortages in society.

2.A Data Addendum

Our data come from the German Socioeconomic Panel (GSOEP), a representative longitudinal micro-dataset that contains a wide range of socio-economic information on individuals in Germany, comprising follow-ups for the years 1984-2008. Information was first collected from about 12,200 randomly selected adult respondents in West Germany in 1984. After German reunification in 1990, the GSOEP was extended to around 4,500 persons from East Germany, and subsequently supplemented and expanded by additional samples. The data are well-suited for our analysis in that they allow us to exploit information on a wide range of background variables, locus of control and wages, for a representative panel of individuals. Furthermore, the inclusion of a special youth survey, comprising information on 17-year-olds, allows us to obtain background variables and locus of control measures for individuals who have not yet entered the labor market.

2.A.1 Combining samples

Our focus is to analyze the impact of locus of control and to purge our estimates of measurement error and endogeneity problems. Hence, to investigate how locus of control affects schooling decisions and wages, respectively, we would ideally need a sample of individuals for whom locus of control measures are collected at several points in time: first, at the time when individuals make education decisions, and second, at a time just before they start the respective job for which labor market returns can be observed. In this way, we would obtain locus of control measures that are truly exogenous, and not influenced by previous on-the-job labor market experience. However, we only have access to one measure of what we term 'premarket' locus of control. This measure is taken when individuals are 17 years of age, just after compulsory schooling, but before they enter the labor market.[15] We then combine the sample of youth for which we have 'premarket' locus of control measures with a sample of young adults for whom we observe labor market outcomes. We draw our samples on the basis of selection criteria that are explained in the following.

[15]Locus of control measures have also been collected for a cross section of young adults in 2007, but we disregard this information, as we suspect it to be flawed by previous labor market experience.

Youth sample

Our youth sample is composed of 1,534 individuals born between 1984 and 1991, all of which are children of GSOEP panel members. A comprehensive set of background variables, schooling choices, as well as locus of control measures of these individuals, have been collected in the years 2001-2008, when the subjects were 17 years of age. After the first interview at age 17, all subjects are subsequently interviewed on a yearly basis until early adulthood. For example, in 2008, the oldest youth are 24 years of age. An exception to the age rule was made for the 2001 wave, such that some subjects were already 18 or 19 years of age when first completing the questionnaire. We exclude these individuals from our sample. Besides, to ensure that our results are not flawed by post 1991 schooling and labor market adjustments, all individuals who went to school in East Germany (the former German Democratic Republic) have been excluded. Last, we exclude all individuals with missing locus of control measures, missing schooling information, or missing information among the covariates.

Adult sample

The adult sample used for our analysis comprises information on 1,192 individuals, aged 26-35, who are drawn from all West German representative subsamples We construct a cross-section of individuals based on the most recent information available from the waves 2004-2008. Hence, most of our information on the adult sample stems from the 2008 wave. However, if some important pieces of information on certain individuals in that wave are missing, they are filled up with information from 2007. If the information in the 2007 wave is also missing, information from 2006 is used, and so on.

We want to ensure that labor market outcomes and cognitive measures are not related to language problems, post 1991 adjustments, or discrimination. Hence, we exclude non-German citizens, individuals who did not live in West Germany at the time of reunification, as well as individuals whose parents do not speak German as a mother tongue. We also exclude handicapped individuals and individuals in vocational training. Furthermore, we exclude individuals with missing schooling information, because the schooling equation is crucial as it links our two samples and ensures identification. Also, individuals with missings among the control variables are dropped from the sample.

2.A.2 'Premarket' locus of control

In the GSOEP, locus of control is measured by a 10-item questionnaire. However, the number of possible answers differs between the years 2001-2005, where a 7-point scale was used, and the years 2006-2008, where a 4-point item scale was used. To make the questionnaire comparable across samples, we transform the 7-point scale into a 4-point scale by assigning the middle category (4) either to category 2 or 3 of the 4-item scale, depending on the most probable answer. For example, if in the 2005 sample most youth answered "completely agree," people who answered "indifferent" in the 2006 sample are assumed to tend toward the "slightly agree" answer. After transforming answers to have the same scale, each question is answered on a Likert scale ranging from 1 ("completely disagree ") to 4 ("completely agree").

2.A.3 School choice

We group schooling into two broad categories: higher education and lower education. Individuals are classified as being highly educated whenever they have some kind of academic qualification. That is, to qualify as highly educated, individuals need to have passed at least those exams that mark the completion of secondary schooling, and which are obtained in tracks with an academic orientation (German high school diploma (Abitur) obtained either at Gymnasium or Gesamtschule). To identify the level of schooling obtained, we use the international Comparative Analysis of Social Mobility in Industrial Nations (CASMIN) Classification, which is a generated variable available in the GSOEP. We define individuals as being highly educated when their attained education level corresponds to CASMIN categories (2c, 3a, 3b). Similarly, individuals are low-educated if their education status is classified according to CASMIN classification categories (1b, 1c, 2a, 2b). Furthermore, for a subsample of youth who have not completed their education at the time of the last interview, we replace their final education status with their aspired (planned) level of education.

2.A.4 Wage construction and labor market participation

Wages are constructed by using most recent wage information available from the GSOEP. Whenever occurring, missing wage information was substituted by wage information

obtained in one of the earlier years. Wages have been inflation adjusted to match 2008 wage levels (inflation rates obtained from Eurostat). Wages are assigned a missing whenever the respective individual is indicating not to have a regular (full time or part time) job. We exclude other types of employment such as marginal employment, to ensure that we are not including typical student jobs.

Hourly wages have been constructed by dividing gross monthly wages by the actual number of hours worked in the last month before the interview. Log hourly wages are then obtained by taking the natural logarithm of the hourly wage variable. To account for outliers, we trim hourly wages below the first and above the ninety ninth percentiles. All individuals who indicate a positive wage and are full- or part-time employed are classified as labor market participants.

2.A.5 Covariates

In our measurements system, schooling equation and outcome equations, we control for a large set of background variables. The locus of control factor distribution is identified from the covariance structure of the unobservables of the model. Hence, any controls in the measurement system purge our measures of locus of control of any effects which are captured by the covariates. Thus, the covariates in place should be uncorrelated with the latent trait we want to capture, since in our model the latent factor has to be uncorrelated with these covariates by construction. In the following, a brief description of the different categories of covariates is provided.

Parental education and investment

Parental education variables have been constructed in the form of dummy variables for higher secondary degree (German Gymnasium), lower secondary degree (German Hauptschule or Realschule), dropout and other degree. This information was collected using the Biography Questionnaire, which every person answers when she is first interviewed in the GSOEP.

Apart from parental education, *Parental investment* is proxied by two variables: broken home and number of siblings. Our broken home variable reflects the percentage of childhood time spent in a broken home until the age of 15. This information was also obtained from the Biography Questionnaire. Last, the number of siblings is obtained for

the youth by counting the number of siblings living in the household. If an individual has many brothers and sisters, this may indicate that parental time is spread among more individuals, and that overall parental investment is lower.

Region dummies and city size

Because school quality and availability, culture and incomes may vary between large and small municipalities, we control for the size of the city where agents spent most of their childhood. Hence, we specify dummy variables for large city, medium city, small city and countryside. Furthermore, we specify four region variables to represent the current region of residence. Hereby, the German Länder are classified as follows:

- North: Berlin, Bremen, Hamburg, Lower Saxony, Schleswig-Holstein,
- South: Bavaria, Baden-Württemberg,
- West: Hessen, North Rhine-Westphalia, Rhineland-Palatinate, Saarland,
- East: Brandenburg, Mecklenburg Western Pomerania, Saxony, Saxony-Anhalt, Thuringia.

Unemployment rates

We construct unemployment rates at two different points in time. First, we use overall German unemployment at the time when individuals are 17, to have a rough measure of the business cycle when schooling decisions are made. Second, we use region (Länder) specific unemployment rates at the time when labor market outcomes are observed. The latter are important to explain the participation decision, as well as local wage rates. All local unemployment rates are obtained from the Federal Employment Office (Bundesagentur für Arbeit), and overall unemployment from the German Federal Statistical Agency (Bundesamt für Statistik).

Marital status and number of children

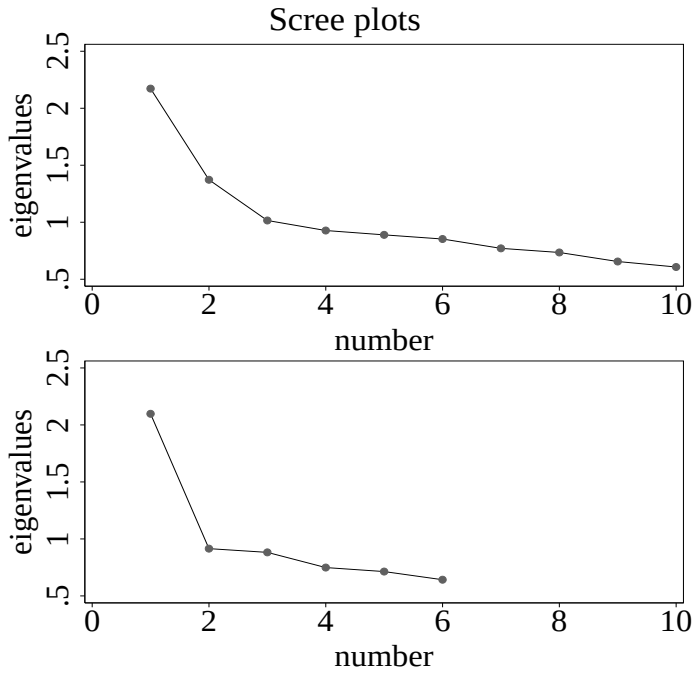
We construct a dummy variable for whether someone is married by looking at her current marital status. Furthermore, we identify the number of dependent children by counting all children for which child benefit payments (Kindergeld) are received by the household.

These variables are important, because previous studies show that being married and the number of dependent children have a positive impact on labor market participation and wages for males, and a negative one for females (see, e.g., Hill, 1979, among others).

Track recommendation after elementary school

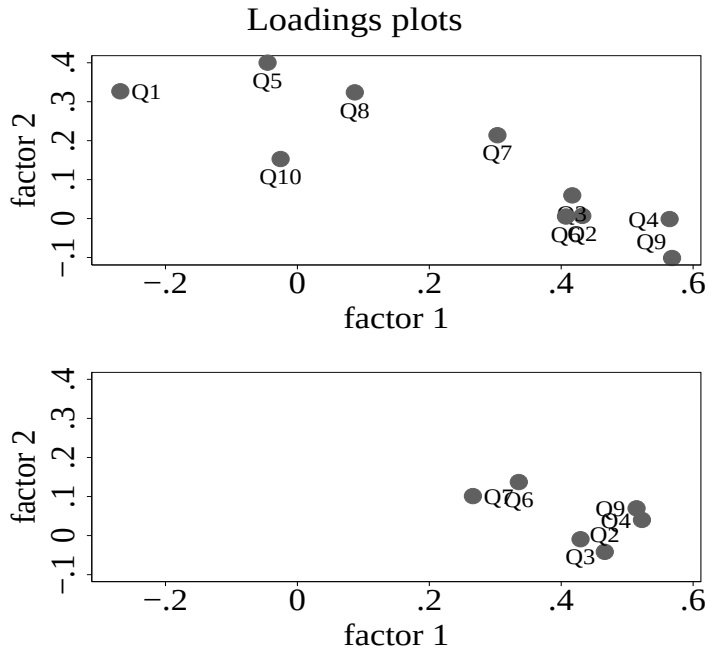
We acknowledge that both schooling decisions and locus of control measures may be correlated with cognitive skills. Hence, in order to proxy cognitive skills, and to account for the fact that schooling decisions may depend on prior track attendance, we include an individual's track recommendations after elementary school. In Germany, track recommendations are given to every student during 4^{th} grade by their elementary school teachers. In some of the German Länder, track recommendations are non-mandatory (but generally adhered to). In some other Länder, track recommendations are compulsory.

Figure 2.6: Scree plot: all measurements versus 6 'external' items only



GSOEP, own calculations.

Figure 2.7: Scatterplot of loadings: all measurements versus 6 'external' items only



GSOEP, own calculations.

Table 2.5: Proportion of people with higher education (all samples)

Variables	Mean	SD	N
Females (youth sample)	0.518	0.500	774
Males (youth sample)	0.459	0.499	760
Females (adult sample)	0.461	0.499	592
Males (adult sample)	0.368	0.483	600

2.B Descriptive Statistics

Table 2.6: Descriptive statistics: labor market outcomes by schooling

	High education			Low education			
Variables	**Mean**	**SD**	**N**	**Mean**	**SD**	**N**	***p*-value**
Labor market participation (males)	0.49	0.50	472	0.71	0.45	617	0.00
Hourly wage (males)	16.03	7.16	228	11.58	4.67	435	0.00
Labor market participation (females)	0.49	0.50	553	0.58	0.49	558	0.00
Hourly wage (females)	12.89	4.86	269	10.35	4.00	316	0.00

Source: GSOEP, cross section using most recent information from the waves 2004-2008. Own calculations. **Notes:** p-values of a two-sided t-test for differences in means are reported.

Table 2.7: Descriptive statistics: covariates in the measurement system

	Males		Females	
Variables	**Mean**	**SD**	**Mean**	**SD**
Childhood in large city	0.20	0.40	0.22	0.42
Childhood in medium city	0.19	0.40	0.19	0.40
Childhood in small city	0.29	0.45	0.25	0.44
North	0.26	0.44	0.24	0.43
South	0.31	0.46	0.34	0.47
Recommendation: grammar school	0.39	0.49	0.45	0.50
Recommendation: general secondary school	0.17	0.38	0.13	0.34
Number of siblings	0.98	1.27	1.01	1.22
Broken home	0.24	0.43	0.24	0.43
Father grammar school	0.29	0.45	0.33	0.47
Father dropout	0.03	0.16	0.04	0.19
Mother grammar school	0.23	0.42	0.25	0.44
Mothers dropout	0.01	0.10	0.03	0.16
# Observations	760		774	

Source: GSOEP, cross section using most recent information from the waves 2004-2008. Own calculations. **Notes:** p-values of a two-sided t-test for differences in means are reported.

Table 2.8: Descriptive statistics: covariates in the outcome equations (by schooling)

	Males					**Females**				
	High education		Low education			High education		Low eduction		
Variables	**Mean**	**SD**	**Mean**	**SD**	***p*-val**	**Mean**	**SD**	**Mean**	**SD**	***p*-val**
Age	24.96	5.88	26.52	5.53	0.00	25.31	5.86	25.82	5.50	0.14
Broken home	0.14	0.35	0.18	0.39	0.07	0.17	0.37	0.22	0.41	0.03
Father grammar school	0.44	0.50	0.07	0.26	0.00	0.41	0.49	0.08	0.28	0.00
Father dropout	0.01	0.09	0.03	0.16	0.03	0.01	0.07	0.04	0.20	0.00
Mother grammar school	0.33	0.47	0.08	0.27	0.00	0.30	0.46	0.08	0.27	0.00
Mother dropout	0.00	0.07	0.02	0.14	0.03	0.01	0.11	0.03	0.17	0.06
Childhood in large city	0.24	0.43	0.18	0.38	0.01	0.21	0.40	0.20	0.40	0.71
Childhood in medium city	0.21	0.41	0.20	0.40	0.58	0.23	0.42	0.18	0.38	0.03
Childhood in small city	0.26	0.44	0.24	0.42	0.29	0.25	0.44	0.25	0.43	0.77
North	0.24	0.43	0.21	0.41	0.25	0.23	0.42	0.20	0.40	0.27
South	0.31	0.46	0.33	0.47	0.58	0.28	0.45	0.34	0.47	0.06
Unemployment at schooling decision	9.01	1.30	8.93	1.37	0.35	9.03	1.39	9.03	1.34	0.97
Unemployment	7.47	2.90	7.72	3.25	0.18	7.70	3.06	7.52	3.12	0.33
Married	0.16	0.37	0.23	0.42	0.01	0.18	0.38	0.27	0.45	0.00
Number of children	1.03	1.12	0.79	1.01	0.00	0.96	1.17	0.92	1.12	0.55
# Observations	472		617			553		558		

Source: GSOEP, cross section using most recent information from the waves 2004-2008. Own calculations. **Notes:** p-values of a two-sided t-test for differences in means are reported.

2.C Identification of the Model

The overall model consists of $K + 5$ submodels, with the latent factor θ as only source of unobserved correlation between them. Let ψ be the set of all parameters. Stated as such, our model is not identified. The lack of identification has different sources. Some are typical in latent variable models and can be fixed by imposing appropriate restrictions (section 2.C). Others are due to the structure of our data set, and more specifically to the fact that the overall model is not identified if we use the two samples separately. Identification only arises from the combination of the two samples (Section 2.4.1).

Identifying restrictions

Independence assumption of the factor The latent factor is assumed to be independent of the covariates and of the error terms:

$$\theta \quad \perp\!\!\!\perp \quad (X, \varepsilon), \tag{2.8}$$

where $X = (X_S, X_Y, X_M)$ and $\varepsilon = (\varepsilon_S, \{\varepsilon_{Y,s}\}, \{\varepsilon_{M,k}\})$. This assumption is standard in factor analysis and is required for identification.

Ordinal models The first problem with these models is the invariance of the measurements to location transformations of the latent M^*. Adding any constant term to equation (2.6) and shifting the thresholds γ by the same quantity will not alter the observed measurements. This problem is typically solved by setting the first finite cut-point γ_1 to zero. In case no intercept term is included in the equation, this restriction is redundant. The second problem is the invariance to scale transformations. Multiplying the latent measurements by a constant will modify the covariance matrix, but will leave the likelihood unchanged if the cut-points are multiplied by the same constant. To better understand this point, consider the $(K \times 1)$-vector of latent measurements $M^* = (M_1^*, ..., M_K^*)'$ and stack the corresponding equations to obtain the following system:

$$\begin{aligned} M^* &= (I_K \otimes X_M)\beta_M + \vartheta, \\ \vartheta &= \alpha_M \theta + \varepsilon_M, \end{aligned}$$

where $\beta_M = (\beta'_{M,1}, ..., \beta'_{M,K})'$, $\alpha_M = (\alpha_{M,1}, ..., \alpha_{M,K})'$ and $\varepsilon_M = \text{diag}_{k=1}^{K} \{\varepsilon_k\}$. Because of the independence assumption (2.8), the covariance matrix of the error term ϑ is

$$\mathrm{V}[\vartheta] = \begin{pmatrix} \alpha_{M,1}^2\sigma_\theta^2 + \sigma_{\varepsilon_{M,1}}^2 & & & & \\ \alpha_{M,1}\alpha_{M,2}\sigma_\theta^2 & \alpha_{M,2}^2\sigma_\theta^2 + \sigma_{\varepsilon_{M,2}}^2 & & & \\ \alpha_{M,1}\alpha_{M,3}\sigma_\theta^2 & \alpha_{M,2}\alpha_{M,3}\sigma_\theta^2 & \alpha_{M,3}^2\sigma_\theta^2 + \sigma_{\varepsilon_{M,3}}^2 & & \\ \vdots & \vdots & & \ddots & \\ \alpha_{M,1}\alpha_{M,K}\sigma_\theta^2 & \alpha_{M,2}\alpha_{M,K}\sigma_\theta^2 & \dots & & \alpha_{M,K}^2\sigma_\theta^2 + \sigma_{\varepsilon_{M,K}}^2 \end{pmatrix} \tag{2.9}$$

The invariance to scale transformations problem is usually solved by fixing the diagonal elements of the covariance matrix to one. This restriction sets the scale of the latent measurements, and thus enables parameter identification from the polychoric correlation matrix. In our case, the structure of the covariance matrix is well-defined. Since the measurements are independent conditional on the controls X_M and on the factor θ, it is sufficient to restrict the variances of the error terms to one to achieve identification ($\sigma_{\varepsilon_{M,1}}^2 = ... = \sigma_{\varepsilon_{M,K}}^2 = 1$). With this restriction, there exists no scale transformation of the measurements yielding the same likelihood other than the identity transformation.
The last identification issue concerns the factor loadings. Looking at the covariance matrix (2.9), it becomes clear that only the ratios of the factor loadings can be identified from the covariances, for example:

$$\frac{\text{Cov}(M_1^*; M_2^*|X_M)}{\text{Cov}(M_1^*; M_3^*|X_M)} = \frac{\alpha_{M,2}}{\alpha_{M,3}}. \tag{2.10}$$

The factor loadings are thus identified up to a multiplicative constant. This is a typical problem in factor analysis which can be easily solved by fixing one of the loadings to one, so as to set the scale of the latent factor θ. In the ratio (2.10), fix for instance $\alpha_{M,3}$ to one to identify $\alpha_{M,2}$. With $\alpha_{M,2}$ in hand, $\alpha_{M,1}$ can then be retrieved from the following relation:

$$\frac{\text{Cov}(M_1^*; M_4^*|X_M)}{\text{Cov}(M_2^*; M_4^*|X_M)} = \frac{\alpha_{M,1}}{\alpha_{M,2}},$$

and all the other loadings can be identified by domino effect in the same way. Once the factor loadings have been identified, the variance of the latent factor σ_θ^2 can be recovered from any element of the covariance matrix (2.9).

Dichotomous model If the latent factor θ is treated as given, the schooling choice in equation (2.2) can be regarded as a simple probit model. The variance of the error term ε_S is fixed to one to ensure the identifiability of the parameters. For the factor loading α_S to be identified, the same argument as before can be applied using the covariance between the latent utility of schooling S^* and any latent measurement M_k^*.

Specification of latent factor distribution and error terms distributions

In their seminal papers, Heckman and co-authors (see e.g. Hansen et al. (2004), among others) achieve nonparametric identification by means of some independence and support assumptions. When the measurement system consists continuous outcomes, they first nonparametrically identify the joint distribution of the observed and latent measurements before turning to the identification of the latent factors and error terms using a theorem proposed by Kotlarski (1967). In our case, this identification strategy cannot be applied, insofar as we only dispose of discrete measurements. Nonparametric identification of the latent factor distribution as well as of the error term distributions would only be possible if we first managed to nonparametrically identify the joint distribution of the latent measurements. This preliminary stage appears to be difficult when dealing with discrete variables, if not impossible. In our case, there is actually no chance to nonparametrically identify this joint distribution, because the covariates used are sparse and common across measurement equations. The lack of variability for each measurement prevents any nonparametric identification. For this reason, we will use a fully parametric approach in our empirical application.

Following Heckman et al. (2006), and to remain as flexible as possible, a mixture of normals is specified for the distribution of the factor. Mixtures of gaussian distributions are widely used in applied work to approximate unknown densities (Escobar and West, 1995). This popularity is rooted in the seminal work of Ferguson (1983b) who showed that normal mixtures with a large number of components can approximate virtually any distribution. In most applications, only a small number of components can be estimated to keep the likelihood tractable. The use of mixtures instead of the usual gaussian distribution

for the factor makes it possible to capture some features of the latent factor distribution which would otherwise be neglected, thus reducing the bias in the estimation of the factor loadings.

For the sake of simplicity, standard normal distributions are assumed for the error terms of the schooling and measurement equations. For the wage equation, a mixture of normals is used to introduce more flexibility. The use of a normal mixture allows more flexibility in accounting for wage heterogeneity, and proves to be very helpful. Another alternative would be to introduce a second latent factor in the wage equation to capture this unobserved heterogeneity. In our case, we only have two equations in the adult sample and the identification of a second latent factor would be problematic.

2.D Estimation Results

Table 2.9: Shooling choice – males, common wage equation

	Schooling
cst	0.1912
	(0.1309)
nb siblings	-0.1250
	(0.0389)
broken family	-0.2835
	(0.1162)
father secondary school	-0.1750
	(0.1064)
father grammar school	0.9739
	(0.1351)
mother secondary school	-0.3475
	(0.1047)
mother grammar school	0.6228
	(0.1514)
north	-0.0013
	(0.1169)
south	-0.2238
	(0.1076)
large city	0.2390
	(0.1343)
medium city	0.1034
	(0.1318)
small city	0.1301
	(0.1172)
cohort 26/30	-0.3253
	(0.1403)
cohort 31/35	-0.2576
	(0.1366)
factor	0.3062
	(0.0890)
precision	1.0000
	(0.0000)

Table 2.10: Common wage equation – males

	Wage
cst	2.4501
	(0.1591)
unemp rate	-0.0040
	(0.0168)
large city	0.0854
	(0.0696)
medium city	-0.0651
	(0.0629)
small city	0.0738
	(0.0655)
north	-0.0857
	(0.0631)
south	0.0274
	(0.0836)
cohort 31/35	0.1661
	(0.0482)
factor	0.1738
	(0.0672)
precision	4.8042
	(0.9731)

Note: Estimates obtained with 100,000 draws from the posterior distribution. Standard deviations in brackets.

Table 2.11: Locus of control – males, common wage equation

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
cst	2.5567	0.8360	0.9142	0.8442	2.5676	0.7578	1.8601	2.2369	0.4580	1.1089
	(0.1921)	(0.1455)	(0.1424)	(0.1600)	(0.2088)	(0.1374)	(0.1530)	(0.1644)	(0.1574)	(0.1265)
nb siblings	-0.0208	0.0259	0.0845	0.0281	0.0538	0.0747	-0.0478	0.0010	0.0292	0.0178
	(0.0377)	(0.0386)	(0.0377)	(0.0423)	(0.0369)	(0.0367)	(0.0364)	(0.0344)	(0.0428)	(0.0331)
broken family	-0.0781	0.1385	0.3530	0.0857	-0.0245	0.2291	0.1382	0.0265	0.1455	-0.1582
	(0.1104)	(0.1140)	(0.1123)	(0.1261)	(0.1057)	(0.1080)	(0.1062)	(0.0990)	(0.1273)	(0.0961)
father secondary school	-0.0921	-0.2635	-0.1874	0.0717	0.0318	-0.2411	-0.1719	-0.0605	-0.0448	-0.0045
	(0.1164)	(0.1207)	(0.1165)	(0.1312)	(0.1115)	(0.1131)	(0.1113)	(0.1045)	(0.1316)	(0.1005)
father grammar school	-0.0244	-0.2271	-0.3145	-0.0518	-0.2537	-0.2956	-0.2554	-0.1753	-0.3212	0.2400
	(0.1253)	(0.1282)	(0.1249)	(0.1395)	(0.1169)	(0.1212)	(0.1191)	(0.1118)	(0.1430)	(0.1080)
mother secondary school	0.0099	0.0951	0.1166	0.1205	-0.0640	0.1063	-0.0344	0.0505	0.2847	0.0137
	(0.1138)	(0.1181)	(0.1144)	(0.1292)	(0.1088)	(0.1110)	(0.1092)	(0.1022)	(0.1304)	(0.0986)
mother grammar school	0.0849	-0.0251	-0.1819	-0.0422	-0.1652	-0.0729	0.0926	-0.0847	0.0179	0.1064
	(0.1303)	(0.1331)	(0.1293)	(0.1454)	(0.1195)	(0.1261)	(0.1232)	(0.1149)	(0.1480)	(0.1112)
north	-0.1650	-0.1611	0.0810	0.0108	0.0354	0.1147	0.2275	0.1954	-0.1060	-0.0685
	(0.1202)	(0.1235)	(0.1189)	(0.1349)	(0.1138)	(0.1161)	(0.1147)	(0.1066)	(0.1365)	(0.1031)
south	-0.2205	-0.0008	-0.0152	0.0635	-0.0260	0.0940	-0.0387	-0.0294	0.2986	-0.2494
	(0.1108)	(0.1131)	(0.1110)	(0.1245)	(0.1053)	(0.1078)	(0.1067)	(0.0999)	(0.1261)	(0.0966)
large city	0.1009	-0.1067	0.0797	-0.0591	0.2468	0.1348	-0.1236	-0.0248	-0.0504	0.0674
	(0.1397)	(0.1417)	(0.1383)	(0.1549)	(0.1329)	(0.1343)	(0.1317)	(0.1240)	(0.1575)	(0.1200)
medium city	-0.0273	-0.0281	0.1266	0.0995	-0.0408	0.0333	0.0346	-0.1421	0.3371	-0.1043
	(0.1393)	(0.1426)	(0.1397)	(0.1569)	(0.1328)	(0.1359)	(0.1339)	(0.1256)	(0.1584)	(0.1214)
small city	-0.0116	-0.1573	0.0617	-0.0164	0.0356	0.0030	-0.0241	-0.1230	0.0966	0.1465
	(0.1210)	(0.1251)	(0.1212)	(0.1365)	(0.1155)	(0.1181)	(0.1163)	(0.1081)	(0.1372)	(0.1054)
factor	0.3543	-0.7317	-0.6884	-1.0000	0.1090	-0.5929	-0.5114	-0.0994	-0.9788	-0.1047
	(0.0825)	(0.1192)	(0.1078)	—	(0.0665)	(0.0986)	(0.0866)	(0.0601)	(0.1677)	(0.0598)
precision	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	—	—	—	—	—	—	—	—	—	—
cutoff 1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	—	—	—	—	—	—	—	—	—	—
cutoff 2	0.7607	1.4047	1.4500	1.6300	0.9021	1.4602	1.3773	1.2832	1.8405	1.1350
	(0.1250)	(0.0727)	(0.0731)	(0.0906)	(0.1574)	(0.0696)	(0.0883)	(0.1123)	(0.1016)	(0.0602)
cutoff 3	2.1283	2.5735	2.7990	2.9618	2.4819	2.7185	3.0661	2.8821	3.2209	2.3959
	(0.1416)	(0.1150)	(0.1169)	(0.1461)	(0.1690)	(0.1113)	(0.1147)	(0.1223)	(0.1795)	(0.0821)

Note: Estimates obtained with 100,000 draws from the posterior distribution. Standard deviations in brackets.

Table 2.12: Schooling choice – males, wage equation by schooling

	Schooling
cst	0.1919
	(0.1292)
nb siblings	-0.1272
	(0.0384)
broken family	-0.2798
	(0.1137)
father secondary school	-0.1736
	(0.1046)
father grammar school	0.9717
	(0.1325)
mother secondary school	-0.3386
	(0.1026)
mother grammar school	0.6115
	(0.1479)
north	-0.0059
	(0.1149)
south	-0.2176
	(0.1065)
large city	0.2332
	(0.1311)
medium city	0.1015
	(0.1294)
small city	0.1237
	(0.1150)
cohort 26/30	-0.3175
	(0.1378)
cohort 31/35	-0.2523
	(0.1331)
factor	0.2093
	(0.0812)
precision	1.0000
	(0.0000)

Table 2.13: Wage equation by schooling – males

	Wage 0	Wage 1
cst	2.3595	2.6903
	(0.1892)	(0.2784)
unemp rate	-0.0032	-0.0082
	(0.0208)	(0.0278)
large city	-0.0045	0.0460
	(0.0895)	(0.1097)
medium city	-0.0835	-0.0644
	(0.0721)	(0.1076)
small city	0.1048	-0.0698
	(0.0763)	(0.1108)
north	-0.0491	-0.1546
	(0.0749)	(0.1084)
south	0.0417	0.0178
	(0.0952)	(0.1452)
cohort 31/35	0.1191	0.2138
	(0.0558)	(0.0863)
factor	-0.0002	-0.0139
	(0.0888)	(0.0901)
precision	4.9933	3.2207
	(1.1684)	(0.9281)

Note: Estimates obtained with 100,000 draws from the posterior distribution. Standard deviations in brackets.

Table 2.14: Locus of control – males, wage equation by schooling

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
cst	2.5581	0.8319	0.9097	0.8378	2.5692	0.7540	1.8649	2.2374	0.4535	1.1087
	(0.1927)	(0.1447)	(0.1412)	(0.1600)	(0.2096)	(0.1374)	(0.1524)	(0.1644)	(0.1577)	(0.1262)
nb siblings	-0.0206	0.0255	0.0842	0.0276	0.0537	0.0743	-0.0482	0.0011	0.0287	0.0176
	(0.0377)	(0.0385)	(0.0378)	(0.0424)	(0.0367)	(0.0367)	(0.0365)	(0.0344)	(0.0427)	(0.0330)
broken family	-0.0794	0.1412	0.3557	0.0895	-0.0238	0.2312	0.1404	0.0271	0.1501	-0.1582
	(0.1103)	(0.1136)	(0.1119)	(0.1257)	(0.1058)	(0.1075)	(0.1064)	(0.0994)	(0.1275)	(0.0962)
father secondary school	-0.0917	-0.2626	-0.1862	0.0711	0.0317	-0.2400	-0.1718	-0.0599	-0.0448	-0.0042
	(0.1158)	(0.1197)	(0.1160)	(0.1307)	(0.1116)	(0.1134)	(0.1115)	(0.1037)	(0.1320)	(0.1004)
father grammar school	-0.0241	-0.2264	-0.3145	-0.0530	-0.2543	-0.2960	-0.2562	-0.1751	-0.3235	0.2394
	(0.1248)	(0.1280)	(0.1248)	(0.1385)	(0.1169)	(0.1212)	(0.1194)	(0.1113)	(0.1433)	(0.1079)
mother secondary school	0.0080	0.0986	0.1197	0.1254	-0.0648	0.1091	-0.0313	0.0514	0.2905	0.0149
	(0.1141)	(0.1170)	(0.1143)	(0.1289)	(0.1088)	(0.1113)	(0.1097)	(0.1023)	(0.1307)	(0.0988)
mother grammar school	0.0852	-0.0258	-0.1817	-0.0421	-0.1664	-0.0726	0.0926	-0.0850	0.0182	0.1064
	(0.1305)	(0.1326)	(0.1294)	(0.1446)	(0.1205)	(0.1258)	(0.1226)	(0.1147)	(0.1475)	(0.1108)
north	-0.1650	-0.1621	0.0787	0.0085	0.0348	0.1134	0.2271	0.1952	-0.1077	-0.0682
	(0.1200)	(0.1229)	(0.1181)	(0.1336)	(0.1142)	(0.1155)	(0.1148)	(0.1073)	(0.1368)	(0.1027)
south	-0.2197	-0.0021	-0.0175	0.0613	-0.0252	0.0930	-0.0407	-0.0297	0.2978	-0.2500
	(0.1109)	(0.1132)	(0.1104)	(0.1239)	(0.1055)	(0.1080)	(0.1068)	(0.0996)	(0.1265)	(0.0968)
large city	0.1007	-0.1048	0.0819	-0.0560	0.2460	0.1367	-0.1213	-0.0237	-0.0480	0.0676
	(0.1392)	(0.1409)	(0.1379)	(0.1543)	(0.1331)	(0.1338)	(0.1326)	(0.1239)	(0.1570)	(0.1194)
medium city	-0.0271	-0.0273	0.1271	0.0997	-0.0417	0.0343	0.0353	-0.1418	0.3382	-0.1045
	(0.1396)	(0.1424)	(0.1392)	(0.1558)	(0.1327)	(0.1357)	(0.1338)	(0.1261)	(0.1585)	(0.1217)
small city	-0.0124	-0.1541	0.0646	-0.0124	0.0345	0.0058	-0.0205	-0.1227	0.1014	0.1479
	(0.1205)	(0.1248)	(0.1208)	(0.1362)	(0.1155)	(0.1181)	(0.1164)	(0.1086)	(0.1377)	(0.1054)
factor	0.3540	-0.7344	-0.6876	-1.0000	0.1140	-0.5906	-0.5196	-0.0963	-0.9970	-0.1103
	(0.0836)	(0.1214)	(0.1093)	—	(0.0668)	(0.0987)	(0.0882)	(0.0599)	(0.1753)	(0.0599)
precision	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	—	—	—	—	—	—	—	—	—	—
cutoff 1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	—	—	—	—	—	—	—	—	—	—
cutoff 2	0.7625	1.4036	1.4481	1.6261	0.9012	1.4584	1.3846	1.2852	1.8460	1.1352
	(0.1257)	(0.0726)	(0.0729)	(0.0910)	(0.1568)	(0.0694)	(0.0885)	(0.1126)	(0.1033)	(0.0608)
cutoff 3	2.1292	2.5707	2.7950	2.9537	2.4827	2.7124	3.0756	2.8834	3.2321	2.3970
	(0.1418)	(0.1144)	(0.1166)	(0.1463)	(0.1687)	(0.1102)	(0.1150)	(0.1222)	(0.1832)	(0.0822)

Note: Estimates obtained with 100,000 draws from the posterior distribution. Standard deviations in brackets.

Table 2.15: Shooling choice – females, common wage equation

	Schooling
cst	0.6001
	(0.1398)
nb siblings	-0.1560
	(0.0383)
broken family	-0.4022
	(0.1136)
father secondary school	-0.0588
	(0.1126)
father grammar school	1.0399
	(0.1388)
mother secondary school	-0.4182
	(0.1088)
mother grammar school	0.5096
	(0.1558)
north	-0.3204
	(0.1191)
south	-0.2768
	(0.1100)
large city	-0.0379
	(0.1318)
medium city	0.0772
	(0.1361)
small city	0.0168
	(0.1237)
cohort 26/30	-0.3119
	(0.1389)
cohort 31/35	-0.0365
	(0.1410)
factor	0.4840
	(0.1093)
precision	1.0000
	(0.0000)

Table 2.16: Common wage equation – females

	Wage
cst	2.3756
	(0.1743)
unemp rate	0.0009
	(0.0195)
large city	-0.0412
	(0.0767)
medium city	0.0654
	(0.0715)
small city	0.0361
	(0.0660)
north	-0.0928
	(0.0715)
south	0.0676
	(0.0918)
cohort 31/35	0.0614
	(0.0517)
factor	0.2125
	(0.0812)
precision	5.2990
	(0.9568)

Note: Estimates obtained with 100,000 draws from the posterior distribution. Standard deviations in brackets.

Table 2.17: Locus of control – females, common wage equation

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
cst	3.1121	0.3970	1.4508	0.6336	2.9032	0.9958	1.6040	2.5158	0.6524	1.4407
	(0.2980)	(0.1451)	(0.1494)	(0.1533)	(0.2441)	(0.1462)	(0.1454)	(0.1944)	(0.1857)	(0.1343)
nb siblings	-0.0381	0.0385	-0.0295	0.0341	-0.0259	-0.0113	-0.0040	-0.0309	0.0987	0.0390
	(0.0372)	(0.0390)	(0.0374)	(0.0406)	(0.0360)	(0.0379)	(0.0349)	(0.0340)	(0.0494)	(0.0335)
broken family	0.0985	0.3475	0.0601	0.2547	0.0776	-0.0798	0.0638	0.0513	0.0958	-0.1858
	(0.1077)	(0.1108)	(0.1060)	(0.1147)	(0.1039)	(0.1076)	(0.1000)	(0.0970)	(0.1391)	(0.0948)
father secondary school	-0.0787	-0.0219	0.0533	0.0388	-0.1657	0.0886	0.0510	-0.0392	-0.2486	0.0145
	(0.1179)	(0.1214)	(0.1159)	(0.1259)	(0.1133)	(0.1182)	(0.1097)	(0.1064)	(0.1539)	(0.1034)
father grammar school	-0.1300	-0.0236	-0.1869	-0.0068	-0.1114	-0.2497	-0.0724	-0.0531	-0.2160	0.3564
	(0.1224)	(0.1273)	(0.1219)	(0.1319)	(0.1177)	(0.1240)	(0.1142)	(0.1103)	(0.1603)	(0.1087)
mother secondary school	0.0597	0.1957	-0.1433	0.1199	0.0725	0.2845	0.0236	0.1700	0.1431	-0.1785
	(0.1131)	(0.1169)	(0.1121)	(0.1216)	(0.1090)	(0.1144)	(0.1055)	(0.1020)	(0.1485)	(0.0994)
mother grammar school	-0.0366	-0.1710	-0.2000	0.0023	-0.1428	0.0842	0.1472	-0.0173	0.0002	-0.1842
	(0.1278)	(0.1332)	(0.1269)	(0.1370)	(0.1218)	(0.1287)	(0.1202)	(0.1152)	(0.1666)	(0.1130)
north	0.1873	-0.0427	-0.1056	-0.1202	-0.1214	0.0934	-0.0619	0.0569	0.0044	0.0199
	(0.1194)	(0.1214)	(0.1160)	(0.1267)	(0.1134)	(0.1185)	(0.1095)	(0.1057)	(0.1537)	(0.1032)
south	-0.1126	-0.0354	-0.0116	0.0464	-0.2097	0.2756	-0.0311	-0.1395	0.2759	-0.0234
	(0.1073)	(0.1122)	(0.1075)	(0.1165)	(0.1046)	(0.1093)	(0.1010)	(0.0979)	(0.1423)	(0.0963)
large city	0.0538	-0.1025	-0.0946	0.0169	0.0798	0.1227	0.0246	-0.0118	0.0365	-0.2481
	(0.1282)	(0.1325)	(0.1276)	(0.1380)	(0.1230)	(0.1287)	(0.1191)	(0.1153)	(0.1666)	(0.1140)
medium city	0.0236	0.2011	0.0839	0.2739	0.0150	0.2475	0.3267	0.0433	0.2911	-0.1166
	(0.1338)	(0.1381)	(0.1324)	(0.1438)	(0.1284)	(0.1349)	(0.1253)	(0.1214)	(0.1744)	(0.1176)
small city	-0.1187	0.0895	-0.1318	0.2642	0.0821	-0.0070	0.1493	-0.0973	-0.0864	0.0008
	(0.1240)	(0.1296)	(0.1233)	(0.1348)	(0.1203)	(0.1258)	(0.1168)	(0.1131)	(0.1644)	(0.1103)
factor	0.3889	-0.8246	-0.7491	-1.0000	-0.0129	-0.8355	-0.4580	-0.1120	-1.5261	0.1469
	(0.0900)	(0.1234)	(0.1161)	—	(0.0764)	(0.1223)	(0.0921)	(0.0731)	(0.2415)	(0.0714)
precision	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	—	—	—	—	—	—	—	—	—	—
cutoff 1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	—	—	—	—	—	—	—	—	—	—
cutoff 2	1.2507	1.5501	1.6154	1.6060	0.9175	1.3964	1.3408	1.6079	2.0780	1.3541
	(0.2630)	(0.0762)	(0.0780)	(0.0807)	(0.1971)	(0.0741)	(0.0831)	(0.1548)	(0.1394)	(0.0690)
cutoff 3	2.8774	2.6001	2.9842	2.8058	2.6444	2.7268	3.0106	3.2096	3.5597	2.8312
	(0.2706)	(0.1188)	(0.1158)	(0.1206)	(0.2081)	(0.1078)	(0.1063)	(0.1610)	(0.2395)	(0.0938)

Note: Estimates obtained with 100,000 draws from the posterior distribution. Standard deviations in brackets.

Table 2.18: Shooling choice – females, wage equation by schooling

	Schooling
cst	0.6050
	(0.1377)
nb siblings	-0.1578
	(0.0378)
broken family	-0.4041
	(0.1120)
father secondary school	-0.0687
	(0.1108)
father grammar school	1.0359
	(0.1370)
mother secondary school	-0.4173
	(0.1071)
mother grammar school	0.5052
	(0.1543)
north	-0.3181
	(0.1180)
south	-0.2755
	(0.1083)
large city	-0.0399
	(0.1298)
medium city	0.0747
	(0.1339)
small city	0.0136
	(0.1222)
cohort 26/30	-0.3082
	(0.1367)
cohort 31/35	-0.0357
	(0.1383)
factor	0.4178
	(0.1077)
precision	1.0000
	(0.0000)

Table 2.19: Wage equation by schooling – females

	Wage 0	Wage 1
cst	2.4813	2.4444
	(0.2609)	(0.2444)
unemp rate	-0.0186	0.0018
	(0.0296)	(0.0272)
large city	-0.0739	0.0233
	(0.1099)	(0.1068)
medium city	0.0528	0.0821
	(0.1025)	(0.0983)
small city	-0.0191	0.0710
	(0.0965)	(0.0900)
north	-0.0450	-0.0752
	(0.0978)	(0.1068)
south	-0.0741	0.1291
	(0.1304)	(0.1267)
cohort 31/35	-0.0554	0.1223
	(0.0730)	(0.0719)
factor	-0.0617	-0.0430
	(0.1394)	(0.1102)
precision	4.2845	5.0350
	(1.0479)	(1.1527)

Note: Estimates obtained with 100,000 draws from the posterior distribution. Standard deviations in brackets.

Table 2.20: Locus of control – females, wage equation by schooling

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
cst	3.1195	0.3963	1.4487	0.6324	2.9062	0.9968	1.6051	2.5133	0.6526	1.4405
	(0.3016)	(0.1454)	(0.1491)	(0.1531)	(0.2443)	(0.1466)	(0.1457)	(0.1931)	(0.1873)	(0.1336)
nb siblings	-0.0382	0.0383	-0.0297	0.0338	-0.0259	-0.0118	-0.0043	-0.0306	0.0986	0.0390
	(0.0371)	(0.0390)	(0.0373)	(0.0405)	(0.0360)	(0.0379)	(0.0350)	(0.0341)	(0.0497)	(0.0335)
broken family	0.0974	0.3490	0.0619	0.2566	0.0782	-0.0785	0.0634	0.0515	0.0988	-0.1866
	(0.1082)	(0.1107)	(0.1058)	(0.1145)	(0.1035)	(0.1077)	(0.1002)	(0.0968)	(0.1401)	(0.0946)
father secondary school	-0.0788	-0.0215	0.0540	0.0402	-0.1663	0.0888	0.0513	-0.0387	-0.2495	0.0144
	(0.1181)	(0.1212)	(0.1160)	(0.1259)	(0.1140)	(0.1179)	(0.1094)	(0.1059)	(0.1548)	(0.1036)
father grammar school	-0.1293	-0.0254	-0.1881	-0.0082	-0.1111	-0.2509	-0.0722	-0.0533	-0.2200	0.3568
	(0.1223)	(0.1273)	(0.1218)	(0.1318)	(0.1181)	(0.1236)	(0.1140)	(0.1105)	(0.1611)	(0.1087)
mother secondary school	0.0592	0.1988	-0.1401	0.1239	0.0728	0.2881	0.0255	0.1694	0.1510	-0.1789
	(0.1134)	(0.1171)	(0.1121)	(0.1215)	(0.1092)	(0.1143)	(0.1057)	(0.1019)	(0.1496)	(0.0993)
mother grammar school	-0.0352	-0.1711	-0.1994	0.0021	-0.1430	0.0831	0.1453	-0.0175	0.0013	-0.1843
	(0.1277)	(0.1330)	(0.1269)	(0.1375)	(0.1215)	(0.1287)	(0.1194)	(0.1152)	(0.1675)	(0.1135)
north	0.1858	-0.0409	-0.1032	-0.1177	-0.1214	0.0953	-0.0601	0.0572	0.0083	0.0193
	(0.1191)	(0.1216)	(0.1158)	(0.1264)	(0.1133)	(0.1182)	(0.1093)	(0.1054)	(0.1545)	(0.1037)
south	-0.1120	-0.0369	-0.0125	0.0444	-0.2098	0.2738	-0.0321	-0.1396	0.2749	-0.0233
	(0.1073)	(0.1120)	(0.1080)	(0.1158)	(0.1047)	(0.1094)	(0.1015)	(0.0982)	(0.1426)	(0.0964)
large city	0.0552	-0.1048	-0.0963	0.0132	0.0802	0.1196	0.0226	-0.0120	0.0321	-0.2465
	(0.1285)	(0.1323)	(0.1270)	(0.1373)	(0.1232)	(0.1289)	(0.1191)	(0.1154)	(0.1670)	(0.1144)
medium city	0.0245	0.1996	0.0828	0.2720	0.0149	0.2455	0.3258	0.0436	0.2897	-0.1160
	(0.1336)	(0.1379)	(0.1320)	(0.1436)	(0.1282)	(0.1347)	(0.1248)	(0.1213)	(0.1753)	(0.1181)
small city	-0.1178	0.0873	-0.1345	0.2606	0.0820	-0.0105	0.1475	-0.0975	-0.0918	0.0014
	(0.1243)	(0.1293)	(0.1238)	(0.1346)	(0.1197)	(0.1260)	(0.1170)	(0.1125)	(0.1651)	(0.1110)
factor	0.3906	-0.8265	-0.7477	-1.0000	-0.0116	-0.8401	-0.4597	-0.1091	-1.5517	0.1479
	(0.0912)	(0.1241)	(0.1163)	—	(0.0766)	(0.1237)	(0.0923)	(0.0733)	(0.2504)	(0.0718)
precision	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	—	—	—	—	—	—	—	—	—	—
cutoff 1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	—	—	—	—	—	—	—	—	—	—
cutoff 2	1.2590	1.5486	1.6140	1.6041	0.9213	1.3965	1.3411	1.6053	2.0872	1.3542
	(0.2678)	(0.0763)	(0.0778)	(0.0806)	(0.1978)	(0.0740)	(0.0834)	(0.1532)	(0.1426)	(0.0689)
cutoff 3	2.8848	2.5972	2.9801	2.8007	2.6478	2.7279	3.0112	3.2073	3.5734	2.8311
	(0.2750)	(0.1184)	(0.1150)	(0.1206)	(0.2092)	(0.1075)	(0.1059)	(0.1603)	(0.2455)	(0.0931)

Note: Estimates obtained with 100,000 draws from the posterior distribution. Standard deviations in brackets.

2.E Goodness-of-fit Tests

Table 2.21: Goodness-of-fit test for log wages (Kolmogorov-Smirnov test)

	Males		**Females**	
	(1)	(2)	(3)	(4)
overall	0.026		0.027	
	(0.745)		(0.799)	
$S = 0$		0.085		0.055
		(0.072)		(0.398)
$S = 1$		0.042		0.047
		(0.418)		(0.479)

Notes: Model estimated by conditioning labor market outcomes on education [(2) and (4)] and without conditioning on education [(1) and (3)]. Two-sample K-S test with null hypothesis that the actual sample and the posterior predictive sample have the same distribution. p-values in brackets. Exact p-values could not be computed due to ties in the distribution of actual wages.

Table 2.22: Goodness-of-fit check: proportion of correct predictions of education achievement for each decile of the latent factor distribution

		Deciles of latent factor distribution									
		10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
Males	(1)	0.827	0.816	0.804	0.785	0.764	0.737	0.708	0.673	0.645	0.699
		(0.033)	(0.031)	(0.032)	(0.033)	(0.035)	(0.035)	(0.035)	(0.038)	(0.039)	(0.044)
	(2)	0.788	0.785	0.777	0.762	0.745	0.728	0.711	0.691	0.671	0.669
		(0.033)	(0.033)	(0.034)	(0.034)	(0.035)	(0.036)	(0.036)	(0.039)	(0.040)	(0.040)
Females	(3)	0.771	0.746	0.728	0.714	0.702	0.689	0.677	0.667	0.667	0.692
		(0.033)	(0.035)	(0.036)	(0.037)	(0.037)	(0.036)	(0.037)	(0.038)	(0.041)	(0.040)
	(4)	0.756	0.736	0.722	0.709	0.698	0.689	0.678	0.666	0.659	0.675
		(0.034)	(0.037)	(0.037)	(0.037)	(0.037)	(0.039)	(0.039)	(0.039)	(0.039)	(0.039)

Notes: Model estimated by conditioning labor market outcomes on education [(2) and (4)] and without conditioning [(1) and (3)]. Proportions of correct predictions computed for each MCMC replication, corresponding means and standard errors (in brackets) are reported.

Chapter 3

Maternal Endowments, Investments, and Birth Outcomes

3.1 Introduction

Maternal cognitive, social and health endowments are important determinants of prenatal behaviors and a ring in the chain of the intergenerational transmission of inequality. Using UK data from the National Child Development Study (NCDS), the top panel of Figure 3.1 shows that mothers at the lowest quartile of the distribution of cognitive skills are almost 10 percentage points more likely to give birth to Small for Gestational Age (SGA)[2] babies than mothers in the upper quartile of the distribution.[3] Similarly, mothers with low social skills and mothers with less healthy physical constitution are more likely to give birth to small for gestational age babies than their counterparts in the upper tail of the distribution. These maternal endowments can affect newborn health either directly, or through prenatal choices, such as education and smoking.[4] The bottom panel of Figure 3.1 shows that both choices are highly correlated with the probability of giving birth to an SGA baby. These figures only show correlations. In practice, we need to understand causal mechanisms to develop effective policies. This is our aim in this chapter.

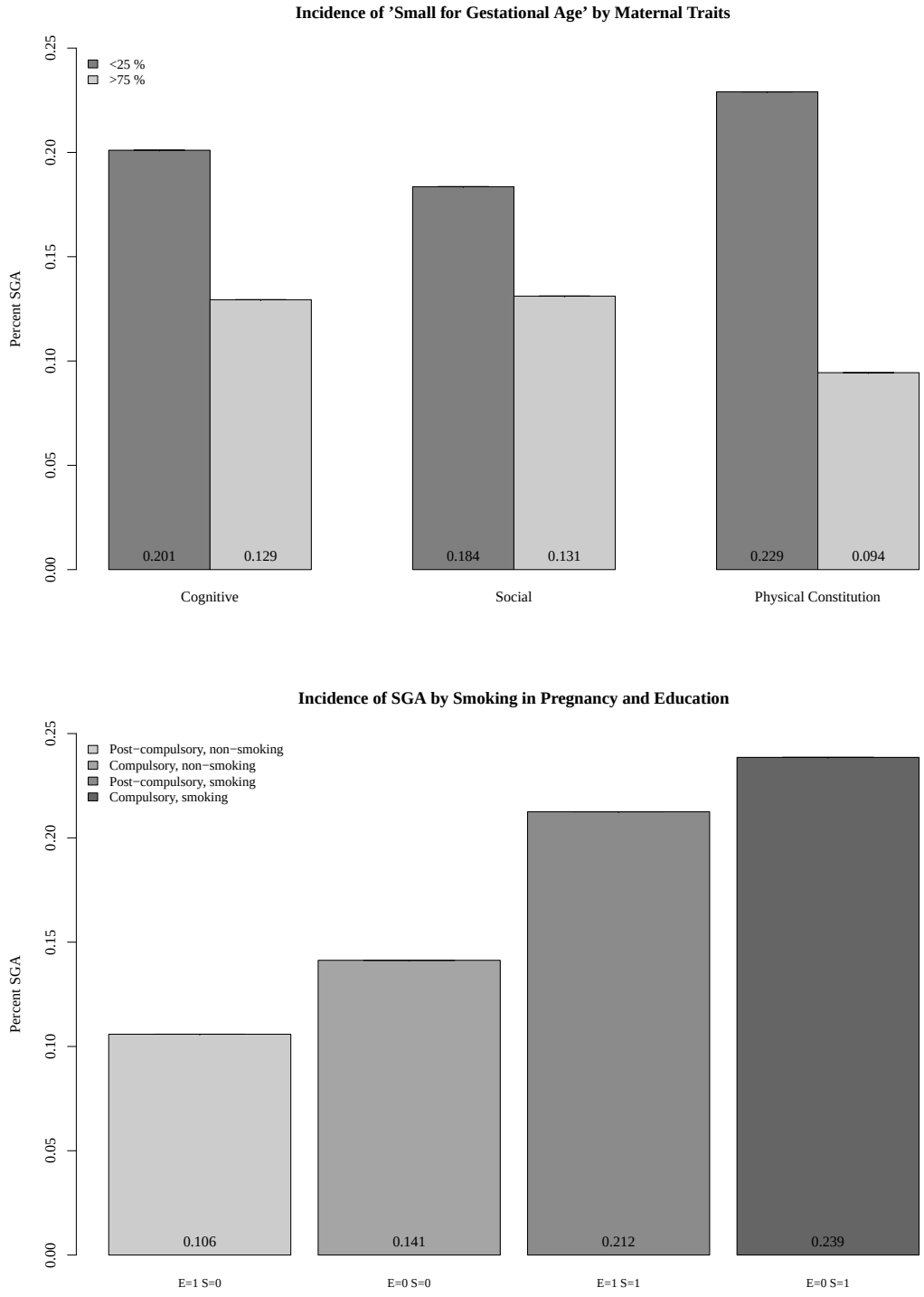
[1]This chapter is joint work with Gabriella Conti, James J. Heckman and Arianna Zanolini.

[2]SGA is defined in this chapter as being below the 10^{th} percentile of the distribution of birthweight by gestational age.

[3]Endowments here are latent factor scores predicted from measurements of cognitive skill, social skill and physical constitution, respectively. All details of our estimation are reported in Section 2.

[4]Unfortunately, we are unable to examine other maternal prenatal behaviors due to data limitations.

Figure 3.1: Probability of being born small for gestational age (SGA), by maternal traits and treatment status



Note: National Child Development Study (NCDS), Birth cohort 1958. The analytical sample on which these figures are based consists of all female cohort members that have no missings in any of the covariates.

First, examining the effects of maternal endowments on newborn health is important because it helps us understanding how intergenerational transmission of disadvantage arises. On the one hand, more disadvantaged mothers (Currie (2009), Finch (2003), Kramer et al. (2001)) tend to give birth to babies in poorer health. On the other hand, in a circle of intergenerational transmission of disadvantage, infants with worse health also have worse labor market outcomes, lower test scores, poorer health later in life and even a higher probability of themselves giving birth to babies with worse health. This relationship persists even after controlling for maternal characteristics, environments, and for genetic endowments (Behrman and Rosenzweig (2004), Royer (2009), Oreopoulos et al. (2008), Currie and Hyson (1999)). In this chapter, we study how disadvantage is transmitted from mother to child, and consider low maternal cognitive, social and physical endowments as one particular form of disadvantage.

Critically, we go beyond simply assessing the relationship between maternal endowments and newborn health, to examine the mechanisms behind it. We both estimate the total impact of maternal endowments on birth outcomes, and we also decompose it into a direct, "biological" effect, and into a "choice" effect, which is mediated by maternal behaviors. Additionally, we estimate the causal effects of maternal education and smoking in pregnancy, and we investigate whether women endowed with different traits have different returns.

The chapter is structured as follows. We discuss the relationship of our work with the previous literature in Section 3.2. In Section 3.3 we present our model and estimation strategy. We describe our data in Section 3.4, and discuss our results in Section 6.6. Section 6.7 concludes.

3.2 Previous Literature

[Literature]

Our work is related to several different literatures. The first strand of literature we refer to is the one studying the effects of early endowments. Maternal endowments have been found to be predictive of risky behaviors (Carneiro et al. (2007), Heckman et al. (2006), Conti et al. (2010)) and of child health (Rubalcava and Teruel (2004), Bhargava and Fox-Kean (2003), Currie and Moretti (2007)), both in the eco-

nomic and in the medical/epidemiologic literature. The latter has mostly focused on the link between maternal birth weight and height and offspring birth weight (Gluckman and Hanson (2004b), Subramanian et al. (2009), Kramer (1987)). However, none of these papers analyzes the mechanisms through which maternal endowments operate. Nonetheless, the importance of accounting for them has long been recognized. For example, Rosenzweig and Schultz (1983) report that not accounting for unobserved parental health endowments leads to a substantial under-estimation of the benefits of early prenatal care, while Fertig (2010) and Tominey (2007) find that selection explains between a third and a half of the association between prenatal smoking and birth weight.

The second strand of literature we contribute to is the one examining the effects of maternal prenatal behaviors on newborn health. Such literature reports estimates of varying magnitude of the negative effects on the baby, especially for what concerns prenatal smoking behavior. On the one hand, using panel data methods, prenatal smoking has been found to reduce birth weight on average by 150-250 grams (see e.g. Abrevaya (2006); Almond et al. (2005)). On the other hand, using instrumental variable (IV) techniques and maternal fixed effects, the magnitude of the effect is reduced to 50g (Lien and Evans (2005), Rosenzweig and Wolpin (1991), Walker et al. (2009), Abrevaya (2006)). Smoking also increases the probability of low birth weight by 2% to 7% (Abrevaya (2006), Almond et al. (2005), Evans and Ringel (1999), Tominey (2007), Walker et al. (2009)) and reduces the length of gestation by around 0.3-0.7 weeks (Abrevaya (2006), Rosenzweig and Wolpin (1991)).

A third strand of literature our work is related to is the one on the effects of education. Such literature has not reached a consensus yet on the effectiveness of educational policy as health policy. On the one hand, Currie and Moretti (2003) find that one extra year of maternal education, instrumented with college openings, reduces the probability of having a child with low birth weight by 1 percentage point; this is consistent with what found in Chou et al. (2010). On the other hand, Carneiro et al. (2007) and McCrary and Royer (2011) find no significant impact of education on birth weight. Also, Currie and Moretti (2003) find that one extra year of education decreases the probability of smoking in pregnancy by 6 percentage points (equivalent to a 30% reduction). An effect of similar

magnitude is found in Carneiro et al. (2007), while McCrary and Royer (2011), using age at school entry policies, find less evidence of an effect. It is well known, however, that instrumental variable estimators only identify a Local Average Treatment Effect (LATE), which measures the effect for those individuals induced into the treatment by the change in the instrument. If the effects of a treatment vary across individuals (given observed variables) and the agents act on the basis of their idiosyncratic returns, then marginal and average ex post returns will not be the same (see McCrary and Royer (2011) for a discussion of this in relation to maternal education and child health).

Methodologically, our approach is close to that in a series of papers by Li and Poirier (Li and Poirier, 2003b,a; Li et al., 2003), in which the authors estimate a structural equations input-output model using Bayesian methods. However, they neither explicitly model maternal endowments, nor the mechanisms through which they affect the health of the newborn.

Our chapter also differs from the majority of the literature on the causes and consequences of health at birth, because we use Small for Gestational Age (SGA), instead of low birth weight, as measure of newborn health.[5] The literature on the long term impact of SGA in the medical field is instead very long: SGA has been found to be associated with adult educational attainment, income and also with height (Strauss, 2000), cognitive outcomes in adolescence (Ido et al., 1995), height in adolescence (Frisancho et al., 1994), as well as hypertension at 50 years (Spence et al., 2012). SGA is arguably a better indicator of fetal growth than low birth weight, because it identifies babies who are small even after accounting for the time spent in utero.[6] Indeed, the term SGA is often used interchangeably with Intrauterine Growth Restriction (IUGR) (Meyer et al. (2009), Karlberg et al. (1995)); however, SGA does not necessarily imply IUGR, and vice versa

[5]Among the papers in economics which control for the time spent in the womb, see Behrman and Rosenzweig (2004) and Oreopoulos et al. (2008).

[6]Note that this is effectively one of the advantages of using twin-based designs: since twins have the same gestational age, any difference in birth weight between them is informative about differences in fetal growth.

(Lee et al. (2003)).[7]

In the following section we describe how we account for all these mechanisms into a unified modeling and estimation framework, which allows us to investigate the channels through which disadvantage is transmitted through generations.

3.3 Model

In this chapter we estimate a sequential selection model with a factor structure,[8] building upon, and expanding, the approach of Hansen et al. (2004), Heckman et al. (2006) and Conti et al. (2010). In our model, women first choose whether to continue education beyond the minimum compulsory level, and then, conditional on the educational choice, they choose their prenatal smoking behavior, which ultimately affects the probability of delivering a SGA baby.[9] We model both maternal choices and newborn outcomes as function of observable characteristics and latent traits. Maternal endowments $\Theta = (\theta_C, \theta_S, \theta_P)$ are composed of cognitive (C) and social skills (S)[10] of the mother, as well as her physical constitution (P).[11]

While the inclusion of cognitive and personality traits in economic models of behavior is well-established by now, the use of a maternal factor to capture the biological propensity to have a baby of a certain size (given the constraint imposed by the physical constitution of the mother) is new.[12] Additionally, it is based on a solid economic and medical

[7] We re-estimate all the models in this chapter using low birthweight instead of SGA, present the full results in Appendix C, and discuss them in the chapter when significant differences with the SGA results emerge.

[8] Factor models have become increasingly used in the economics literature (see e.g. Goldberger (1972); Joreskog and Goldberger (1975) and more recently Heckman et al. (2006); Conti et al. (2010); Piatek (2010); Heckman et al. (2011)).

[9] In a previous version of the chapter, we also modeled the pregnancy choice. However, its inclusion does not affect the results, so it has been omitted. This suggests that there are no other (observed or unobserved) determinants which affect the pregnancy choice, outside from those already included in the model.

[10] We use the term "social skills" for consistency with the previous literature in Economics and for simplicity, but the scale we use more specifically measures behavioral problems and social adjustment, as we will detail in Section 3.4. Social skills could alternatively be called socio-emotional endowments or social abilities.

[11] Note that the number of latent factors in our model is specified a priori; see Conti et al. (2012) for a model in which the number of latent of factors and the measurements they are proxied by is not pre-specified.

[12] See however Conti et al. (2010) for the use of a health factor.

literature, which establishes that maternal birth weight and height affect the health of the newborn (Lumey and Stein, 1997). For example, using the same data that we use, Hennessy and Alberman (1998) show that maternal physical constitution is a key determinant of SGA. It is by now well documented that conditions in utero are determined by factors going back two generations (Kuzawa and Quinn (2009)) and that maternal phenotype embodies her own cumulative environmental experiences which are in turn transmitted through biological vectors to the fetus, determining its health (Kuzawa, 2005). Consistent with this literature, we use maternal height in childhood and her birthweight, and grandmaternal height, as measurements for this factor, and we interpret it as a proxy for healthy physical constitution, arising from good fetal environments.[13] Importantly, controlling for maternal physical constitution also allows us to interpret the effects of maternal choices on the probability of delivering a small for gestational age baby as the result of environmentally-driven growth restrictions, rather than of a biological predisposition.

3.3.1 The measurement system

We assume that the observed measurements of maternal endowments are a function of observable characteristics X_C, X_S, X_P (proxies for early family environments), and of latent traits Θ=(θ_C, θ_S, θ_P), where θ_C indicate cognitive skills, θ_S social skills and θ_P physical constitution. The measurements are dedicated, but the factors are allowed to be correlated ($Cov(\theta_i, \theta_j) \neq 0$) $\forall\ i \neq j$. These traits are assumed to be fully determined at the time of the measurement (age eleven of the mother) and imperfectly observed by the researcher, but known to the mother.

[13] As described in the next section, we also condition on maternal weight, in order to avoid capturing the effect of a higher body mass, which is associated with bigger stature.

For individual i and measurement n, the cognitive skill and physical constitution measurement systems are:

$$M_{Cn,i} = X_{Cn,i}\beta_{Cn} + \lambda_{Cn}\theta_{C,i} + \epsilon_{Cn,i} \qquad \text{for } n = 1, ..., N_C$$

and

$$M_{Pn,i} = X_{Pn,i}\beta_{Pn} + \lambda_{Pn}\theta_{P,i} + \epsilon_{Pn,i} \qquad \text{for } n = 1, ..., N_P,$$

respectively.

Social skill measures are binary, so we write the measurement system in terms of a latent index structure:

$$M^*_{Sn,i} = X_{Sn,i}\beta_{Sn} + \lambda_{Sn}\theta_{S,i} + \epsilon_{Sn,i} \qquad \text{for } n = 1, ..., N_S,$$

$$M_{Sn,i} = I[M^*_{Sn,i} > 0],$$

where $X_{Cn,i}$, $X_{Sn,i}$, $X_{Pn,i}$, are vectors of covariates and λ_{Cn}, λ_{Sn} and λ_{Pn} are factor loadings, for the n^{th} measurement of the cognitive, social skill and physical constitution factors, respectively.[14] M_{Cn}, M_{Sn} and M_{Pn} are sets of measurements specific to each trait $T = \{C, S, P\}$. Continuous measurements are normalized to have mean zero and variance 1. Since the scale of each factor is arbitrary, we set the factor loading in the first measurement equation specific to each factor to unity to set the scale ($\lambda_{C1} = 1$, $\lambda_{S1} = 1$ and $\lambda_{P1} = 1$). Furthermore, we assume $E[\epsilon_{\theta_T n}] = 0 \ \forall \ \theta$ and $\epsilon_{\theta_T n} \perp\!\!\!\perp \epsilon_{\theta_{T'} m} \ \forall \ m, n$ and $\theta_T \neq \theta_{T'}$. Last, to ensure identification we require $N_T > 2$.[15]

3.3.2 The educational choice

We model the binary choice of obtaining compulsory ($E = 0$) versus post-compulsory education ($E = 1$) using a standard latent index model to characterize the decision rule:

[14] Factor loadings are allowed to differ across equations, so that measurements are given different weights.

[15] Identification of a model with three correlated factors is laid out in Appendix A of Conti et al. (2009).

$E_i = \mathbb{1}[E_i^* > 0]$, where E_i^* is the net utility derived from post-compulsory schooling, and E_i is the observed choice.[16] Latent utility from education is determined by observable covariates and by the latent factors. We model utility as a linear function of observable covariates and latent factors, with additive separability in the error term:

$$E_i^* = Z_{E,i}\beta_E + U_{E,i}$$

where $Z_{E,i}$ is a vector of observed characteristics and $U_{E,i}$ is a random variable that determine an individual's latent utility from education, and $Z_E \perp\!\!\!\perp U_E$. $U_{E,i}$ is itself a linear and additive function of the latent factors and of an idiosyncratic error term:

$$U_{E,i} = \lambda_{CE}\theta_{C,i} + \lambda_{SE}\theta_{S,i} + \lambda_{PE}\theta_{P,i} + \epsilon_{E,i}.$$

The idiosyncratic error terms are assumed to be independent of the observables and of the latent factors, i.e. $\epsilon_E \perp\!\!\!\perp (\theta_C, \theta_S, \theta_P, Z_E)$. Furthermore, they are independent of the idiosyncratic error terms of all the other equations (i.e. $\epsilon_E \perp\!\!\!\perp \epsilon_j \quad \forall \quad j \neq E$).

3.3.3 The smoking choice

Conditional on education, we then model the choice to smoke during pregnancy (S).[17] For mothers, this decision is the result of a utility maximization process, where both the health of the newborn and her own consumption of cigarettes are arguments of her utility function.[18] The net utility from smoking for a mother with education level $E \in \{0, 1\}$ is:

$$S_i^{*E} = Z_{S,i}^E \beta_S^E + U_{S,i}^E \qquad for\ E \in \{0, 1\},$$

where, as before, $Z_{S,i}^E$ is a vector of observed characteristics, and $U_{S,i}^E$ is a random variable which affects the mother's utility from smoking, with $Z_S^E \perp\!\!\!\perp U_S^E$. The latter is specified

[16] So agents select into the treatment if the net utility from doing so is positive.

[17] Note that the Surgeon General Report on Smoking and Health came out in January 1964, when our mothers were 5 years old. Hence, we assume that the adverse effects of smoking were known by the time they were pregnant with their first child.

[18] The model we have in mind is one like in Rosenzweig and Schultz (1983). Here, the child birth weight (H), in addition to consumption of non-health (X) and health goods (Y), enter the utility function, denoted by $U = U(X, Y, H)$. Newborn health is produced according to $H = \Gamma(Y, I, \mu)$, using health consumption goods (Y), health investment goods (I) and parental endowments μ as inputs. In their framework, maternal endowments are all captured by μ, and maternal smoking is one of the components of Y which enter both the health production function and the utility function.

as: $U^E_{S,i} = \lambda^E_{CS}\theta_{C,i} + \lambda^E_{SS}\theta_{S,i} + \lambda^E_{PS}\theta_{P,i} + \epsilon^E_{S,i}\ for\ E \in \{0,1\}$, where, as before, $\epsilon^E_S \perp\!\!\!\perp (\theta_C, \theta_S, \theta_P, Z^E_S)$.

3.3.4 Birth Outcomes

Finally, we model the probability of delivering a baby who is small for gestational age, fixing the education and smoking choice. We focus on outcomes of firstborn children only, since extending the analysis to higher order parities would require to model the fertility choice (since the decision to give birth to a second child might be a function of the health of the first child), which is beyond the scope of this chapter. Let $\mathrm{P}(SGA^{E,S}_i | X^{E,S}_{SGA,i}, \theta_C, \theta_S, \theta_P)$ denote the conditional probability of delivering a SGA baby for the two education and smoking statuses, respectively. Furthermore, assume $SGA^{E,S}_i = \mathbb{1}\left[SGA^{E,S*}_i > 0\right]$, where:

$$SGA^{E,S*}_i = X^{E,S}_{SGA,i}\beta^{E,S}_{SGA} + U^{E,S}_{SGA,i}$$

and

$$U^{E,S}_{SGA,i} = \lambda^{E,S}_{CSGA}\theta_{C,i} + \lambda^{E,S}_{SSGA}\theta_{S,i} + \lambda^{E,S}_{PSGA}\theta_{P,i} + \epsilon^{E,S}_{SGA,i} \quad .$$

for $E \in \{0,1\}\ and\ S \in \{0,1\}$, $X^{E,S}_{SGA} \perp\!\!\!\perp U^{E,S}_{SGA}$, $\epsilon^{E,S}_{SGA} \perp\!\!\!\perp (\theta_C, \theta_S, \theta_P, Z^{E,S}_{SGA})$, and $X^{E,S}_{SGA}$ is a vector of observables. The baby SGA outcome for individual i can thus be written in switching regression representation (Quandt, 1972) as follows: $SGA_i = E_i\left[S^1_i SGA^{1,1}_i + (1 - S^1_i)SGA^{1,0}_i\right] + (1 - E_i)\left[S^0_i SGA^{0,1}_i + (1 - S^0_i)SGA^{0,0}_i\right]$.

3.3.5 Identification and Estimation

Identification in our model is based on the following conditional independence assumption:

$$(SGA_0, SGA_1) \perp\!\!\!\perp ((E,S)|X, Z, \theta_C, \theta_S, \theta_P),$$

where (SGA_0, SGA_1) are potential outcomes for the untreated and treated state, respectively.[19] This is a standard matching assumption, with the difference that a subset of

[19] In practice, we have two sequential treatments and four potential outcomes, however here we refer only to two for simplicity.

the matching variables (Θ) is imperfectly observed, and proxied by several measurements with error.

We implement the conditional independence assumption estimating the sequential selection model with a factor structure, as described in the previous section, by means of Bayesian methods. However, we also provide evidence from a stepwise approach, where factor scores are estimated in a first step and then used as observed covariates in the choice and outcome models in a second one.[20] As mentioned, we use Markov Chain Monte Carlo (MCMC) methods to estimate the parameters of our sequential selection model.[21] In MCMC estimation, the latent factors are sampled from their joint posterior distribution in each iteration.[22] These draws are then treated as additional data in the estimation of the other model parameters. The estimation follows Carneiro et al. (2003); Hansen et al. (2004) and is described in those papers.[23]

Finally, we make the following assumptions on the latent factors and the error terms:

1. The distribution of $\Theta = (\theta_C, \theta_S, \theta_P)$ is flexibly approximated by a trivariate mixture with two components.[24] The probability density function is:

$$f_\Theta(\theta_C, \theta_S, \theta_P) \sim p_1 N(\mu_1, V_1) + p_2 N(\mu_2, V_2),$$

where μ_1 and μ_2 are vectors of dimensions (3×1), and V_1 and V_2 are matrices of dimensions (3×3). We do not restrict the variance-covariance matrices to be diagonal, so as mentioned we allow the underlying factors to be correlated.

[20] Here we apply the method proposed by Iwata (1992) to correct for attenuation bias, and we use bootstrapped standard errors. All the details of this alternative method are provided in Section D in the Appendix.

[21] We prefer Bayesian MCMC methods to classical Maximum Likelihood Estimation (MLE) for its computational convenience.

[22] We run 100,000 iterations in total, of which we discard the first 20,000 as burn-in period, and we then retain one out of 40 of the remaining 80,000.

[23] We use a Bayesian MCMC sampler and code written mostly by Karsten Hansen, Salvador Navarro and Sergio Urzua.

[24] It has been shown that mixtures of normals are able to closely approximate any smooth density (Ghosal and Van Der Vaart, 2001).

2. The idiosyncratic errors[25] associated with binary choice and outcome models are assumed to be distributed as follows:

$$\epsilon_{E,i}, \epsilon^{E}_{S,i}, \epsilon^{E,S}_{SGA,i}, \epsilon_{Sn,i} \sim N(0,1) \; for \; E \in \{0,1\} \; and \; S \in \{0,1\}.$$

The idiosyncratic errors of the continuous cognitive and physical constitution measurements equations are assumed to be distributed as follows:

$$\epsilon_{Cn,i}, \epsilon_{Pn,i} \sim N(0, \sigma^2) \qquad \text{for } n = 1, ..., N_C \text{ and } n = 1, ..., N_P.$$

3. Uninformative normal priors with mean and precision zero are used for all factor loadings.

We can then write the density of outcomes given observables as:

$$f(SGA, E, S, M_C, M_S, M_P | X, Z),$$

where $f(.)$ is the joint density of choices, outcomes and measurements. Written in terms of unobservables, the density is:

$$\int \int \int_{\theta_C, \theta_S, \theta_P} f(SGA, E, S, M_C, M_S, M_P | X, Z, t_C, t_S, t_P) dF_\theta(t_C, t_S, t_P),$$

where $F_\theta(.)$ denotes the joint cumulative distribution function associated with unobserved cognitive, social and physical endowments. Notice that, conditional on unobserved factors and observed characteristics, (E, S, M_C, M_S, M_P) are independent, and the sample likelihood simplifies accordingly.

3.4 Data

We use data from the British National Child Development Study (NCDS), which follows a cohort of individuals born in Great Britain during the week of March 3-9, 1958. The female members of this cohort are the mothers we study. The first wave of the NCDS,

[25] Defined as uniquenesses in relation to the measurement equations of factor models, see Aigner et al. (1984).

called the perinatal mortality survey, was administered at birth, and collected a rich array of information on birthweight, gestational age and other birth health conditions, as well as family background characteristics. Subsequent follow-ups were conducted in 1965, 1969, and 1974. Each administered a parental interview, a medical and a school questionnaire, as well as an achievement test battery including measures on cognitive and social skills at ages 7, 11, and 16. Additional follow-ups were conducted in 1981 (age 23), 1991 (age 33), and 2001 (age 43), and administered questions on fertility, partnerships, employment, and children outcomes.

By restricting our sample to all the female cohort members with nonmissing information on the covariates and at least three measurements for each factor, we are left with a sample of 3,217 observations.

3.4.1 Measurements

The measurements we use to proxy maternal endowments have been collected in the age 11 sweep. On of the advantages of the NCDS data is the availability of measures of cognition, personality and health long before the post-compulsory educational choice we model takes place, at a time when the schooling system in the U.K. is still homogeneous.[26]

The tests to measure cognition (which can be considered relatively stable at 11 years) were designed by the National Foundation for Educational Research in England and Wales (NFER). They include an IQ-type test with verbal and non-verbal subscales, a 35-items reading comprehension test, and a 40-items arithmetic/mathematics test.

As measures of social skills, we use items from the Bristol Social Adjustment Guide (BSAG).[27] This test evaluates the type and extent of behavioral disturbance in children as rated by the children's teachers, who were asked to indicate whether their pupils

[26]Then, at age 11 all pupils had to undertake an exam (now abolished almost everywhere) – the so-called "11-plus" – to be admitted to a selective school.

[27]We preferred the BSAG over the Rutter scale, which also measures behavioral adjustment problems and was administered at age 11, because this latter was administered to the mother, and Achenbach et al. (1987) have shown that teacher assessments are better predictors of childhood problems than parental ones. Achenbach et al. (1987) estimated correlations between child behavioral problems and assessments of teachers, parents, health visitors, and of the children themselves, and found them to be higher for teachers.

scored positively on a range of 146 items of social adjustment, behavior and attitudes. The behaviors were then recoded into 12 so-called 'syndromes': inconsequential behavior, nervous symptoms, anxiety towards adults, anxiety towards children, writing off adults, hostility towards adults, miscellaneous symptoms, restlessness, unforthcomingness,[28] depression and withdrawal.[29] All the items have been recoded, so that the variables we use take the value of 1 in case of absence of a particular syndrome.

Finally, to proxy for the maternal physical constitution factor, we use maternal height at age 11, maternal birth weight, and the adult height of the grandmother. Importantly, we control for maternal weight at age 11 in the measurement system (as described in the next section), to isolate the impact of a healthy physical constitution (reflecting a good early nutritional environment) from that of a bigger body mass which is associated with being taller. We provide evidence of the validity of this factor in capturing the effect of a healthy early nutritional environment when comparing its effects on the probability of delivering a SGA versus a low birth weight baby in Section 6.6.

Summary statistics for all the measurements are displayed in Table 3.1. They show, not surprisingly, that women who have chosen to continue education beyond the compulsory level and not to smoke in pregnancy are endowed with better traits under all the different dimensions.

[28] Unforthcomingness describes a collection of behaviors characterizing a child's unassertiveness, interpersonal and academic passivity, and avoidance of competition (McDermott and Watkins, 1981).

[29] Withdrawal describes behaviors reflecting general social detachment or induced isolationism (McDermott and Watkins, 1981).

Table 3.1: Summary Statistics, measurements, whole sample, Cohort 1958 [administered at age 11]

Variables	TOTAL		Education				Smoking			
	ALL		**Post-compulsory**		**Compulsory**		**Smoking in preg**		**Non-smoking in preg**	
Cognitive Skill Measures										
Verbal IQ	24.482	(8.67)	28.967	(6.99)	22.109	(8.53)	20.964	(8.66)	24.692	(8.45)
Non-verbal IQ	22.134	(7.22)	25.704	(6.27)	20.245	(6.97)	19.548	(6.88)	22.255	(7.15)
Reading comprehension	16.799	(5.76)	20.182	(5.15)	15.010	(5.23)	14.281	(5.29)	16.829	(5.52)
Mathematics	17.909	(9.89)	23.913	(8.93)	14.733	(8.84)	13.803	(8.58)	17.980	(9.75)
Copying designs test	8.394	(1.34)	8.704	(1.37)	8.229	(1.29)	8.118	(1.27)	8.407	(1.34)
Social Skill Measures										
Withdrawal	0.416	(0.49)	0.298	(0.46)	0.478	(0.50)	0.565	(0.50)	0.392	(0.49)
Nervous symptoms	0.067	(0.25)	0.040	(0.19)	0.081	(0.27)	0.099	(0.30)	0.062	(0.24)
Anxiety acceptance, adults	0.288	(0.45)	0.245	(0.43)	0.310	(0.46)	0.330	(0.47)	0.280	(0.45)
Anxiety acceptance, children	0.129	(0.33)	0.102	(0.30)	0.143	(0.35)	0.179	(0.38)	0.119	(0.32)
Hostility toward children	0.130	(0.34)	0.091	(0.29)	0.151	(0.36)	0.207	(0.41)	0.110	(0.31)
Writing off adults	0.324	(0.47)	0.233	(0.42)	0.372	(0.48)	0.471	(0.50)	0.293	(0.46)
Hostility towards adults	0.232	(0.42)	0.173	(0.38)	0.263	(0.44)	0.361	(0.48)	0.205	(0.40)
Miscellaneous symptoms	0.318	(0.47)	0.230	(0.42)	0.365	(0.48)	0.395	(0.49)	0.315	(0.46)
Restlessness	0.112	(0.32)	0.058	(0.23)	0.141	(0.35)	0.194	(0.40)	0.100	(0.30)
Unforthcomingness	0.495	(0.50)	0.430	(0.50)	0.529	(0.50)	0.515	(0.50)	0.487	(0.50)
Depression	0.364	(0.48)	0.243	(0.43)	0.429	(0.50)	0.498	(0.50)	0.336	(0.47)
Withdrawal	0.159	(0.37)	0.118	(0.32)	0.181	(0.38)	0.192	(0.39)	0.156	(0.36)
Measures of Body Size										
Maternal birthweight (grams)	3271.581	(501.80)	3318.091	(461.23)	3246.977	(520.42)	3264.149	(536.95)	3255.021	(490.09)
Grandmaternal height (inches)	63.484	(2.49)	63.614	(2.42)	63.415	(2.53)	63.366	(2.58)	63.506	(2.49)
Height, age 11 (cm)	57.056	(2.92)	57.497	(2.84)	56.823	(2.93)	56.580	(3.08)	57.065	(2.83)
N	3217		1113		2104		588		1797	

Source: National Child Development Study (NCDS), Birth cohort 1958.

Notes: The analytical sample on which these estimates are based consists of all female cohort members that have no missings in any of the covariates.

We measure cognitive and social skills at age 11 of the mothers. Standard errors in brackets. Own calculations.

3.4.2 Outcome variables and covariates

Our main outcome of interest is the probability of having, by the age of 33,[30] a firstborn child who is small for gestational age. As argued above, we believe SGA to be a more appropriate measure of newborn health than low birth weight, as reflecting in utero conditions: conditioning on gestational age allows to identify those children who are small not because born early, but because restricted in their growth.[31]

An obvious issue which arises when using SGA, however, is the presence of measurement error, since at the time the survey was carried out, the use of ultrasounds to check accuracy was still not diffused. In the NCDS data, gestational age is computed based on the mother's self-report of the date of her last menstrual period (LMP), which is then checked against general practitioner records (Jefferis et al., 2002). This recall problem has been pointed out both in the medical (Campbell et al. (1985); Chervenak et al. (1998); Kramer et al. (1988); Harville et al. (2010)) and in the economic literature (Royer, 2009). However, error in recall is mostly problematic for gestational periods longer than 42 weeks and for very early deliveries, while for deliveries within the 37-41 weeks range the accuracy has been shown to be high (Poulsen et al. (2011), Wingate et al. (2007), Mustafa and David (2001)). In our main specification we restrict our sample to deliveries occurred between 26 and 42 weeks of gestation (we limit this range further in Appendix C.3 of the paper); in doing so, we retain 97% of the women.[32] The distribution of gestational age is calculated using a growth chart published by Fenton (2003)[33] in BMC Pediatrics, which is based on a meta-analysis of published reference studies. We adopt the most common definition used in the literature and define a baby to be SGA if she lies below the 10^{th} percentile of the birthweight by gestational age distribution. Table 3.2 shows that 14% of the children in our sample are small for gestational age; this relatively high

[30] In our sample, 75% of the women have given birth at least once by this age, and, given the historical period of our sample, this reasonably approximates all women who will ever give birth.

[31] As already mentioned, we re-estimate all the models in this chapter using low birthweight instead of SGA. While all the results are presented in Appendix C, in the text we only refer to them when they significantly differ from those using SGA.

[32] Results for the complete sample are very similar and are available upon request.

[33] The chart is the most recent version of Babson and Benda (1976), which is the most common chart used in neonatology reference books. Different specifications using different charts or using the 5^{th} percentile to define SGA did not change the results in any significant way, as documented in Appendix C.3.

prevalence is caused in part by the fact that firstborn[34] children are, ceteris paribus, more likely to be SGA than later born children (see e.g. Shah (2010) for a meta study on the association between parity and pregnancy outcomes). We provide extensive robustness tests in Appendix C (by restricting the sample to full-term babies and adopting different definitions of SGA) which confirm our main results.

The educational decision (derived from the 1981 questionnaire) is defined as staying-on after the minimum compulsory school-leaving age. We see from Table 3.2 that in our sample about 35% of the women have continued schooling beyond age 16, and that this percentage is halved among the smokers. Finally, the decision to smoke during pregnancy after the first trimester[35] is also derived from the 1981 questionnaire, so it is asked retrospectively. While we are not worried about selective recall bias, still there is the possibility of women "lying" to the interviewer; however, if present, this should bias our results downwards.

The covariates we include in the measurement equations to control for family background characteristics are specified in Table 2.2.[36] We also include additional variables in the choice equations: regional smoking prevalence at the time the mother was pregnant and whether the grandmother was herself a smoker at the time she was pregnant for the smoking choice, and local labor market conditions (the change in the unemployment rate) in the region of residence at the time of the educational decision. Summary statistics for the covariates and outcomes of our model are reported in Table 3.2.

3.5 Empirical Results and Simulations

We now present our results. In this section we first provide the estimated posterior means of the coefficients and we describe our simulation algorithm (Section 3.5.1). We then present and comment the results of the simulation exercise in three parts. First, we present how mothers sort into education and smoking decisions on the basis of their traits, and also

[34] As mentioned above, we focus on firstborn children only.

[35] Women might not know they are pregnant in the first weeks of gestation, so smoking might not be a choice then.

[36] Note that some covariates are excluded from the choice and outcome equations due to insufficient variation to identify their effects, given the smaller sample sizes in the conditional models.

the effects of these traits on the choices (Section 3.5.2). Second, we show in Section 3.5.3 the results from our decomposition exercise of the overall effects of the maternal traits (displayed in the top panel of Figure 3.1) into the channels through which they operate. Third, we present the treatment effects of education and smoking (Section 3.5.4), we decompose the observed SGA differences by choice (displayed in the bottom panel of Figure 3.1) into selection and causal components, and we examine heterogeneity in the treatment effects.

3.5.1 MCMC results and Simulation Algorithm

The posterior means of the factor loadings in the measurement system, which reflect the correlation between the measurements and the latent factors, are reported in Table 3.3. We notice that the copying designs test contributes to the cognitive factor less than the other tests; that the social skills factor is mostly reflecting lack of hostile behavior rather than of anxious one;[37] and that the maternal physical constitution factor has the loading with the biggest magnitude on maternal height at age 11.

We then use the estimated distributions of parameters and the data[38] to simulate counterfactual choices and outcomes, in order to shed more light on the effects of maternal traits, the mechanisms through which they operate, and the treatment effects of smoking and education. More specifically, the Bayesian MCMC algorithm we use for estimation generates a sample of size K model parameters from their conditional posterior distributions that we can use to simulate (binary) outcomes for each individual according to the following expression:

$$\begin{aligned} \hat{y}_i &= \frac{1}{K}\sum_{k=1}^{K} Pr(U^*(k)_{y,i} \geq 0 | X_{y,i}, \beta_y(k), \lambda_y(k), \Theta_i(k)) \\ &= \frac{1}{K}\sum_{k=1}^{K} \mathbb{1}\left[(X_{y,i}\beta_y(k) + \lambda_y(k)\Theta_i(k) + \epsilon) > 0\right], \end{aligned} \tag{3.1}$$

[37] The items have been recoded, so that they take the value of 1 in case of the absence of a particular syndrome.

[38] We randomly draw with replacement individuals from our sample. This randomly drawn data has the same sample size as the actual data. Then, for each individual, we make $K = 10$ draws from the posterior distributions of the parameters, so to take estimation uncertainty into account.

where $\mathbb{1}[]$ denotes the indicator function and ϵ is drawn from a normal distribution. Furthermore, given independence between the factors and the covariates, we can simulate outcomes for each individual at each percentile of the factor distribution, according to the following:

$$\hat{y}_{i_{\theta_p},} = \frac{1}{K}\sum_{k=1}^{K} \mathbb{1}\left[(X_i\beta(k) + \lambda(k)\Theta_{p,i}(k) + \epsilon) > 0\right] \qquad \text{for } p = 1, ..., 99. \tag{3.2}$$

Furthermore, we can use the predicted probabilities from the choice equations to weight the predicted outcomes, so to generate expected outcomes for each individual, conditional or unconditional on choices.

The first thing we do using these simulated outcomes is to assess whether our model provides a good fit to the data. To do so, we compare the simulated means of the education and smoking decisions, and of child SGA, to the ones in the data. As shown in Table 3.4, the means are not statistically significantly different from each other; additionally, also the proportion of outcomes correctly predicted by the model is remarkably high, especially considering the small number of observations we have in the conditional SGA equations.

Table 3.4: Model fit

	Simulated	Data	Difference	P-val Chi2	% correcly predicted
Education	0.343	0.346	0.003	0.425	0.612
Smoking(E=1)	0.130	0.112	-0.018	0.969	0.732
Smoking(E=0)	0.294	0.304	0.011	0.839	0.631
SGA(E=1, S=1)	0.194	0.212	0.018	0.274	0.669
SGA(E=0, S=1)	0.227	0.239	0.012	0.916	0.646
SGA(E=1, S=0)	0.098	0.106	0.008	0.920	0.758
SGA(E=0, S=0)	0.143	0.141	-0.002	0.882	0.713

Note: National Child Development Study (NCDS), Birth cohort 1958. The analytical sample on which these estimates are based consists of all female cohort members that have no missings in any of the covariates. Table displays model fit.

3.5.2 The Intergenerational Effects of Maternal Traits

In order to examine the effects of maternal traits, we first compute and report the average marginal effects of a one standard deviation change in each of them on the outcomes. As shown in Table 3.5, both cognitive and social skills are significant determinants of the educational choice. However, social skills are the only traits which are significant determinants of the smoking choice. Lastly, neither cognitive nor social skills are significant

determinants of the probability of delivering a SGA baby, after conditioning on the education and smoking choices: the physical constitution of the mother, nonetheless, remains a strong determinant, in particular for women with a low level of education. Importantly, this is no longer the case when we use low birth weight rather than SGA as measure of newborn health: this suggests that our maternal physical constitution factor is indeed capturing the capacity of the mother to deliver a healthy baby, rather than the sheer size of her body (see Table C-1 in the Appendix for the LBW results).

Table 3.6: Correlation of maternal traits

	Cognitive	Social	Body Size
Cognitive	1.000	0.369	0.172
Social	0.369	1.000	-0.003
Body Size	0.172	-0.003	1.000

Note: National Child Development Study (NCDS), Birth cohort 1958. The analytical sample on which these estimates are based consists of all female cohort members that have no missings in any of the covariates.

In order to understand better the sorting decisions of the mothers, Figure 3.2 and Figure 3.3 display the marginal distributions of cognitive, social skills and physical constitution, conditional on the education and smoking choices.[39] We first notice that the spread of the distributions of the three endowments is different: the standard deviations amount to 0.84 for cognition, 1.62 for social skills, and to 0.35 for the maternal physical constitution factor; this has implications when considering the impacts of interventions aimed at changing the relative position of individuals along these distribution. Second, we see that mothers with higher cognition, better social skills and a healthier physical constitution sort into post-compulsory education, as confirmed by a Kolmogorov-Smirnov test. Additionally, women with a higher level of cognition and better social skills decide not to smoke after the third trimester of pregnancy; maternal body fitness, instead, seems to play no role. As mentioned in Section 3.3, we also estimate the model using a stepwise

[39]The unconditional joint distributions are presented in Figure 3.11, Figure 3.12 and Figure 3.13 in the Appendix. The correlations among traits are displayed in Table 3.6. We see that the correlation between cognitive and social skills is 0.369, the one between the cognitive and the physical traits is 0.172, while maternal social and physical traits are basically uncorrelated.

approach, in which factor scores are computed in a first step and then used as observed covariates in the choice and outcome models in a second step using Iwata's (1992) method to correct for measurement error.[40] As we can see in Table 3.7, the results we obtain are remarkably similar to those obtained using Bayesian estimation.

[40]This procedure is described in details in Section D in the Appendix.

Table 3.2: Summary statistics, outcomes and covariates, whole sample, Cohort 1958

Variables	TOTAL		Education				Smoking			
	ALL		**Post-compulsory**		**Compulsory**		**Smoking in preg**		**Non-smoking in preg**	
Outcome Variable and Treatments										
Child Small for Gestational Age (SGA)	0.140	(0.35)	0.110	(0.31)	0.153	(0.36)	0.221	(0.42)	0.114	(0.32)
Post-compulsory education	0.346	(0.48)	1.000	(0.00)	0.000	(0.00)	0.136	(0.34)	0.354	(0.48)
Smoking past 3rd month of pregnancy	0.247	(0.43)	0.112	(0.32)	0.304	(0.46)	1.000	(0.00)	0.000	(0.00)
Child sex (female=1)	0.484	(0.50)	0.475	(0.50)	0.488	(0.50)	0.497	(0.50)	0.480	(0.50)
Grandmaternal weight (stones)	8.853	(1.47)	8.834	(1.35)	8.863	(1.53)	8.976	(1.67)	8.813	(1.41)
Relative weight, age 11 (% of average)	100.914	(14.79)	100.209	(14.42)	101.288	(14.98)	102.303	(15.56)	100.212	(14.38)
Gestation of mothers in weeks	40.201	(1.73)	40.215	(1.54)	40.194	(1.83)	40.210	(1.95)	40.207	(1.70)
Broken home (age 7)	0.091	(0.29)	0.069	(0.25)	0.103	(0.30)	0.128	(0.33)	0.087	(0.28)
Household size (age 7)	5.073	(1.53)	4.783	(1.35)	5.226	(1.60)	5.602	(1.71)	4.976	(1.46)
Mother firstborn	0.391	(0.49)	0.465	(0.50)	0.352	(0.48)	0.306	(0.46)	0.421	(0.49)
Grandma post-compulsory schooling	0.284	(0.45)	0.460	(0.50)	0.192	(0.39)	0.168	(0.37)	0.282	(0.45)
Grandparents high SES	0.181	(0.39)	0.319	(0.47)	0.108	(0.31)	0.071	(0.26)	0.179	(0.38)
Grandparents medium SES	0.588	(0.49)	0.554	(0.50)	0.606	(0.49)	0.575	(0.49)	0.607	(0.49)
Grandmother age at birth of mother	27.601	(5.57)	28.175	(5.26)	27.298	(5.70)	26.910	(5.67)	27.316	(5.45)
% smoking in region	0.332	(0.04)	0.318	(0.04)	0.338	(0.04)	0.350	(0.04)	0.326	(0.04)
Grandmother smoked during pregnancy	0.317	(0.47)	0.254	(0.44)	0.351	(0.48)	0.413	(0.49)	0.301	(0.46)
Regional change in unemployment	-1.211	(0.38)	-1.209	(0.40)	-1.212	(0.37)	-1.251	(0.41)	-1.195	(0.37)
London (age 11)	0.089	(0.28)	0.081	(0.27)	0.093	(0.29)	0.070	(0.25)	0.093	(0.29)
Wales (age 11)	0.061	(0.24)	0.072	(0.26)	0.055	(0.23)	0.070	(0.25)	0.060	(0.24)
Scotland (age 11)	0.111	(0.31)	0.136	(0.34)	0.098	(0.30)	0.167	(0.37)	0.090	(0.29)
London (age 16)	0.090	(0.29)	0.083	(0.28)	0.094	(0.29)	0.071	(0.26)	0.092	(0.29)
Wales (age 16)	0.062	(0.24)	0.075	(0.26)	0.055	(0.23)	0.068	(0.25)	0.062	(0.24)
Scotland (age 16)	0.112	(0.32)	0.137	(0.34)	0.099	(0.30)	0.163	(0.37)	0.092	(0.29)
N	3217		1113		2104		588		1797	

Source: National Child Development Study (NCDS), Birth cohort 1958.

Notes: The analytical sample on which these estimates are based consists of all female cohort members that have no missings in any of the covariates.
Standard errors in brackets. Own calculations.

Table 3.3: Factor loadings in measurement system

	Cognitive Skills	Social Skills	Body Size
Verbal IQ	1	-	-
Non-verbal IQ	0.912	-	-
	(0.01)		
Reading comprehension	0.795	-	-
	(0.02)		
Mathematics	0.904	-	-
	(0.02)		
Copying desings test	0.345	-	-
	(0.02)		
Inconsequential behavior	-	0.564	-
		(0.04)	
Nervous symptoms	-	0.369	-
		(0.04)	
Anxiety acceptance, adults	-	0.207	-
		(0.02)	
Anxiety acceptance, children	-	0.436	-
		(0.04)	
Hostility towards children	-	0.686	-
		(0.06)	
Writing off adults	-	0.575	-
		(0.04)	
Hostility towards adults	-	0.876	-
		(0.07)	
Miscellaneous symptoms	-	0.574	-
		(0.04)	
Restlessness	-	0.663	-
		(0.06)	
Unforthcomingness	-	0.335	-
		(0.03)	
Depression	-	1	-
Withdrawal	-	0.519	-
		(0.04)	
Maternal height age 11	-	-	1.757
			(0.2)
Materal birthweight	-	-	1
Grandmaternal height	-	-	1.09
			(0.11)

Note: National Child Development Study (NCDS), Birth cohort 1958. The analytical sample on which these estimates are based consists of all female cohort members that have no missings in any of the covariates. Table displays factor loadings in the measurement system. Standard errors in brackets.

Table 3.5: Average marginal effects of a 1 SD change in the latent factors

	Cognitive Skills	Social Skills	Body Size
Education	0.152	0.039	0.016
	(0.046)	(0.014)	(0.013)
Smoking(E=1)	-0.011	-0.063	0.003
	(0.022)	(0.04)	(0.026)
Smoking(E=0)	-0.004	-0.056	0.014
	(0.013)	(0.021)	(0.016)
SGA(E=1, S=1)	-0.008	0.075	-0.149
	(0.1)	(0.127)	(0.206)
SGA(E=0, S=1)	0.027	-0.023	-0.11
	(0.027)	(0.025)	(0.054)
SGA(E=1, S=0)	0.003	-0.012	-0.053
	(0.02)	(0.018)	(0.026)
SGA(E=0, S=0)	-0.015	0.009	-0.046
	(0.012)	(0.015)	(0.021)

Note: National Child Development Study (NCDS), Birth cohort 1958. The analytical sample on which these estimates are based consists of all female cohort members that have no missings in any of the covariates. Table displays the average marginal effects of a 1 sd shift in latent factors. Standard errors in brackets.

Table 3.7: Results of our main specification using the three-step estimation procedure

Variables	**Education**	**Smoking in Pregnancy**		**SGA**			
		E=1	**E=0**	**E=1, S=1**	**E=0, S=1**	**E=1, S=0**	**E=0, S=0**
Cognitive skills factor, bias corrected	0.135	-0.004	0.0003	-0.052	0.018	0.002	-0.021
	(0.00)	(0.76)	(0.98)	(0.52)	(0.41)	(0.92)	(0.09)
Social skills factor, bias corrected	0.027	-0.048	-0.050	0.048	-0.015	-0.015	0.00763
	(0.00)	(0.00)	(0.00)	(0.50)	(0.46)	(0.37)	(0.52)
Physical constitution factor, bias corrected	0.008	-0.003	0.003	-0.081	-0.081	-0.032	-0.0262
	(0.30)	(0.80)	(0.81)	(0.24)	(0.00)	(0.01)	(0.02)
Observations	3209	715	1664	79	480	614	1109
Pseudo R^2	0.211	0.110	0.074	0.100	0.039	0.026	0.017
Covariates included	YES	YES	YES	YES	YES	YES	YES

Note: The coefficients display the average marginal effect of a one standard deviation change in the factor.
The analytical sample on which these estimates are based consists of all female cohort members with no missings in any of the covariates. Own calculations based on the NCDS data.
p-values in brackets. Standard errors are bootstrapped using 100 replications.
The bias correction procedure is the one due to Iwata (1992) and described in Appendix D.

Figure 3.2: Marginal densities of maternal traits by education

Note: National Child Development Study (NCDS), Birth cohort 1958. The analytical sample on which these estimates are based consists of all female cohort members that have no missings in any of the covariates. Marginal distribution of traits by education.

Finally, simulated decisions and outcomes by quantiles of the endowments distributions are displayed in Figure 6.7, Figure 3.5 and Figure 3.6. As compared to what seen before, they allow us to assess the effects of an exogenous shift in each maternal endowment in turn at different parts of their distribution (so to detect nonlinearities in those effects), while holding the other two at their means, rather than the average marginal effect of one standard deviation change (as seen in Table 3.5). Additionally, we compute the effects of maternal endowments on choices and outcomes, rather than the conditional ones.[41] Figure 6.7 shows that women with higher cognitive skills are more likely to obtain post-compulsory education. In fact, an early childhood intervention which moves a girl from the 20^{th} to the 80^{th} percentile of the cognitive skill distribution (holding the other two endowments at the mean), would increase her probability of obtaining post-compulsory education from 20% to around 50%. For social skills and physical constitution the effect is of a smaller magnitude, but still significant.[42]

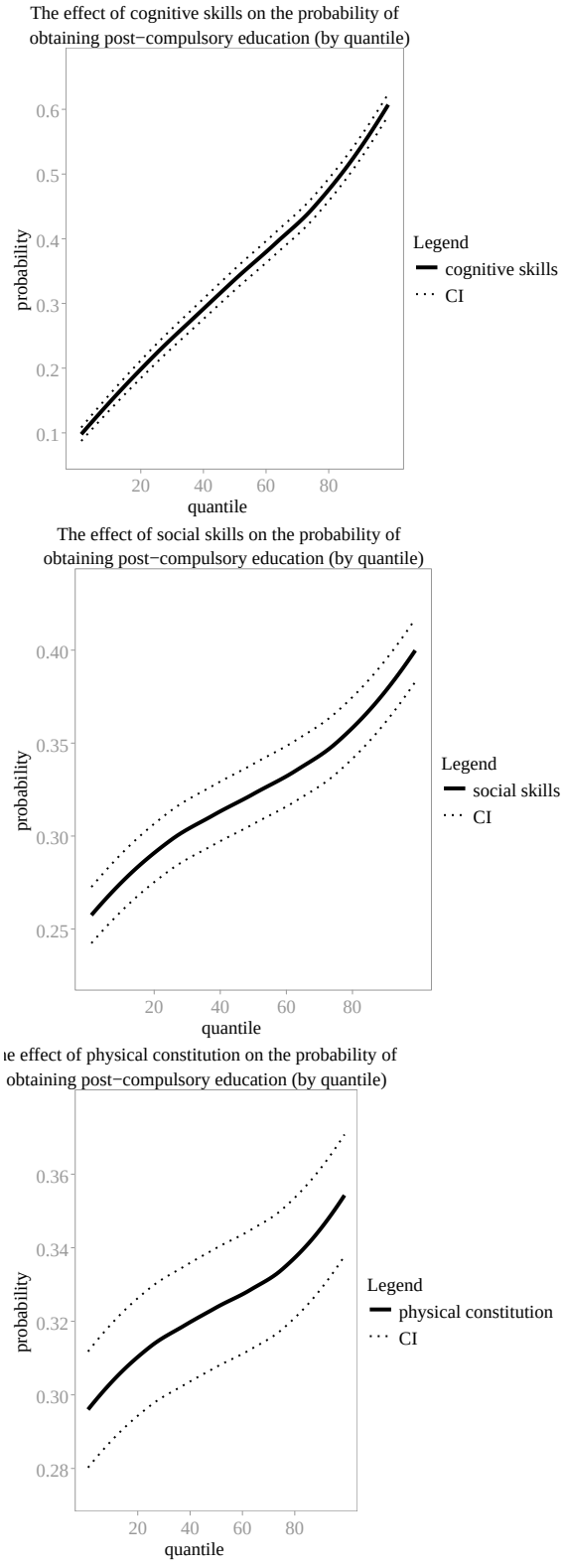
[41] See the next section for a decomposition of the total effect into a direct and an indirect one, i.e. one working through the two channels of education and smoking.

[42] This latter result is consistent with the literature on height and education (Magnusson et al. (2006); Case and Paxson (2010)).

Figure 3.3: Marginal densities of maternal traits by prenatal smoking

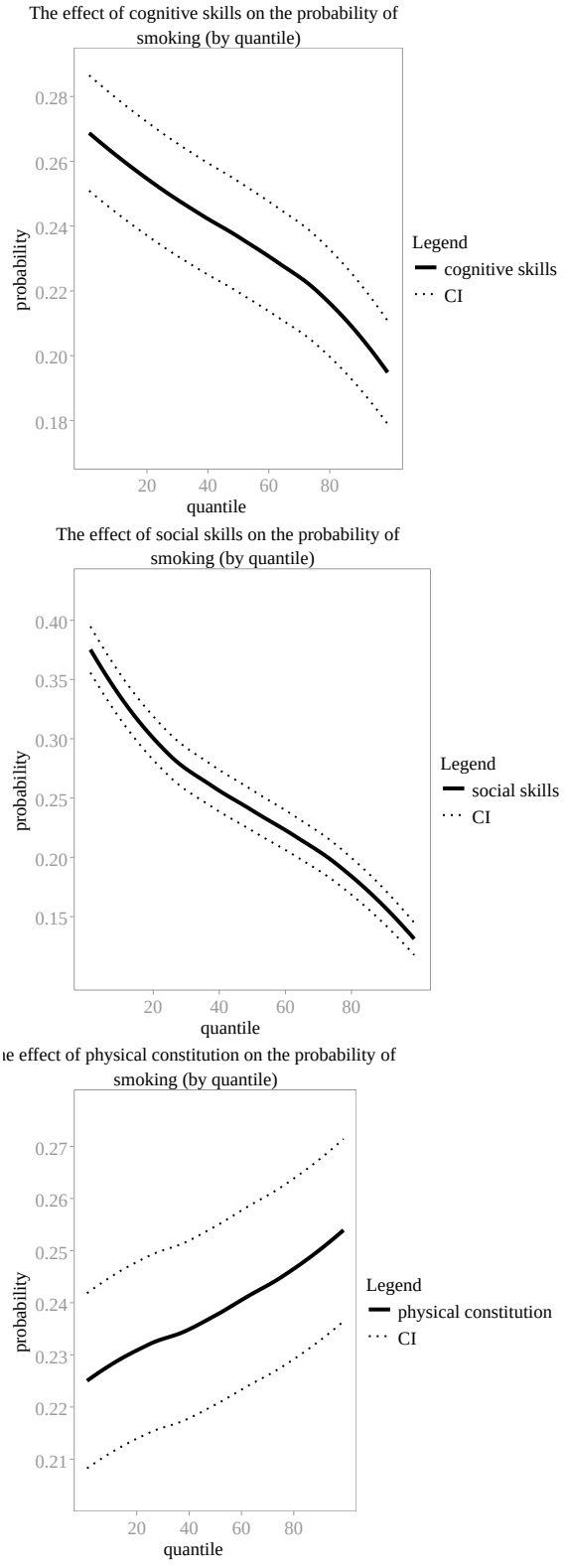
Note: National Child Development Study (NCDS), Birth cohort 1958. The analytical sample on which these estimates are based consists of all female cohort members that have no missings in any of the covariates. Marginal distribution of traits by education.

Figure 3.4: Effect of traits on education



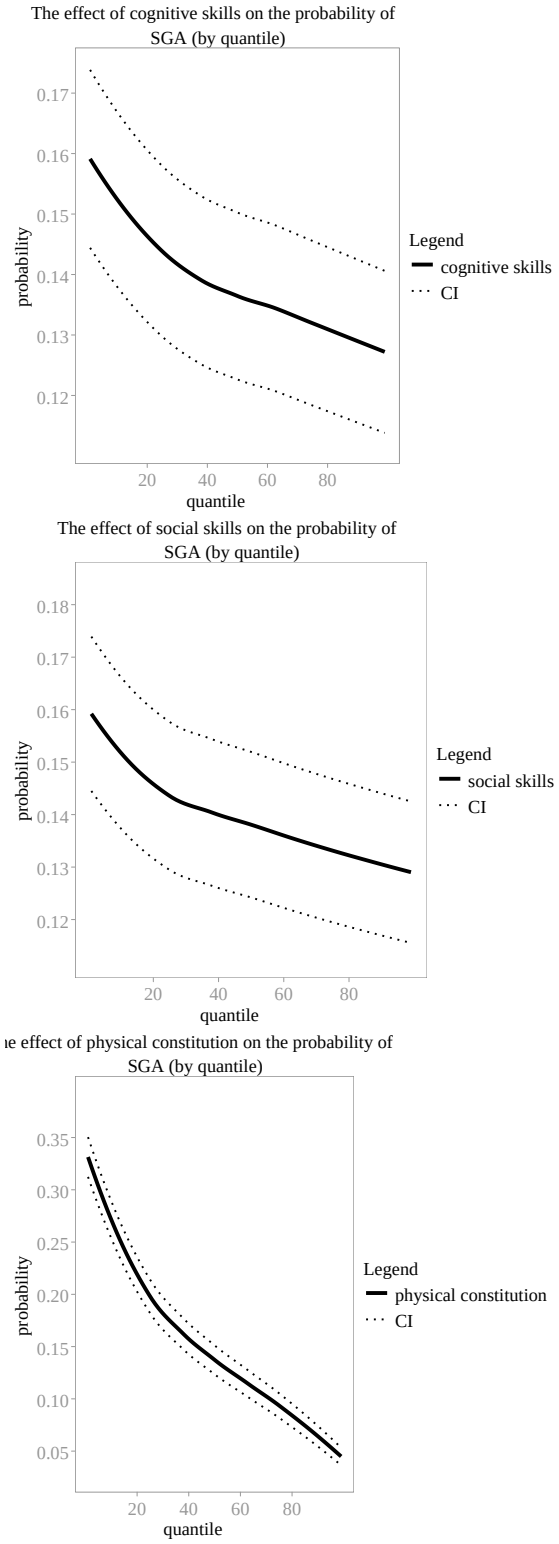
Note: National Child Development Study (NCDS), Birth cohort 1958. The analytical sample on which these estimates are based consists of all female cohort members that have no missings in any of the covariates. Overall effect of traits on outcomes when fixing the other traits at their respective means.

Figure 3.5: Effect of traits on smoking during pregnancy



Note: National Child Development Study (NCDS), Birth cohort 1958. The analytical sample on which these estimates are based consists of all female cohort members that have no missings in any of the covariates. Overall effect of traits on outcomes. Overall effect of traits on outcomes when fixing the other traits at their respective means.

Figure 3.6: Effect of maternal traits on child SGA



Note: National Child Development Study (NCDS), Birth cohort 1958. The analytical sample on which these estimates are based consists of all female cohort members that have no missings in any of the covariates. Overall effect of traits on outcomes. Overall effect of traits on outcomes when fixing the other traits at their respective means.

Second, the decision to smoke during pregnancy is predominantly influenced by social skills, as can be seen from Figure 3.5. We find that an intervention which moves a girl from the 20^{th} to the 80^{th} percentile of the social skills distribution (holding the other two endowments at the mean), would halve her probability to smoke during pregnancy (from 0.30 to 0.15), while a shift of the same magnitude along the distribution of cognitive skills would reduce this probability only by 0.05.[43]

Third, Figure 3.6 displays the overall effect of maternal traits on the probability of giving birth to a small-for-gestational-age baby.[44] We find that an early nutritional intervention which moves a girl from the 20^{th} to the 80^{th} percentile of the distribution of her physical constitution more than halves the probability of delivering a SGA baby (from 0.20 to less than 0.10). Importantly, we find a much smaller effect when using low birthweight as measure of newborn health: a similar intervention would reduce the probability of delivering a LBW baby by only 3 (rather than 10) percentage points (see Figure C-1 in the Appendix). This reassures us about the validity of our measures of SGA as proxying for fetal health, and of maternal physical traits as capturing the fitness of the mother in delivering healthy babies. Then, the corresponding effects of interventions to promote cognitive or social skills are much smaller, with a comparable shift (from the 20^{th} to the 80^{th} percentile) in the respective distributions only yielding a 1 percentage point reduction in the probability of delivering a SGA baby, which is also less precisely estimated.[45] Interestingly, we detect nonlinearities in the effects of maternal endowments, with the bigger gains occurring for women at the bottom of their respective distributions.

In the next subsection we then decompose the effects of maternal endowments on the probability of having a SGA baby into the components which operate through the education and smoking channels, and a residual direct effect.

[43]On the other hand, the effects of an early nutritional intervention arising from a movement of a similar magnitude on the woman's physical constitution are not precisely estimated.

[44]In order to compute the overall effects of skills on child SGA we integrate education and smoking variables as well as covariates out, according to $P(SGA|\Theta=\theta_p) = \int_X \sum_{E=0}^{1} \sum_{S=0}^{1} P(SGA|\Theta=\theta_p, X=x, E=e, S=s)P(S=s|X=x,\Theta=\theta_p, E=e)P(E=e|X=x,\Theta=\theta_p)dF_x$.

[45]The attentive reader might have noticed that the effect of physical traits on the probability of smoking in pregnancy is much smaller than the one estimated in Table 3.5. The reason is the small and positive correlation between the cognitive and physical traits, which have opposite effects on smoking behavior – so that, when they are both allowed to vary (like in Table 3.5), they counterbalance each other out.

3.5.3 Understanding the Mechanisms through which Maternal Endowments Affect Newborn Outcomes

In order to describe our decomposition exercise, we first establish some notation. Let the conditional probabilities of giving birth to a SGA baby, of smoking during pregnancy (S) and of staying on in education beyond the compulsory level (E) be written as, respectively:

$$P(SGA = 1|X = x, \Theta = \theta, E = e, S = s); \tag{3.3}$$

$$P(S = 1|X = x, \Theta = \theta, E = e); \tag{3.4}$$

$$P(E = 1|X = x, \Theta = \theta), \tag{3.5}$$

where X is a vector of predetermined characteristics (which include different subsets of variables, as detailed in Table 2.2, and Θ is the vector of the three maternal endowments.

We first compute these probabilities using simulation (as described in the previous section), and then we apply the product rule to disentangle how a change in maternal endowments affects the newborn SGA through the different channels, as follows:

$$\overbrace{\frac{\partial P\left(SGA = 1|X = x, \Theta = \theta\right)}{\partial \theta_t}}^{\text{Total effect}}$$

$$= \overbrace{\sum_{S=0}^{1}\sum_{E=0}^{1} P(S = s|X = x, \Theta = \theta, E = e)P(SGA = 1|X = x, \Theta = \theta, E = e, S = s)}^{\text{Effect through education}}$$
$$\times \frac{\partial P(E = e|X = x, \Theta = \theta)}{\partial \theta_t}$$

$$+ \overbrace{\sum_{S=0}^{1}\sum_{E=0}^{1} P(E = e|X = x, \Theta = \theta)P(SGA = 1|X = x, \Theta = \theta, E = e, S = s)}^{\text{Effect through smoking}}$$
$$\times \frac{\partial P(S = s|X = x, \Theta = \theta, E = e)}{\partial \theta_t}$$

$$+ \overbrace{\sum_{S=0}^{1}\sum_{E=0}^{1} P(E = e|X = x, \Theta = \theta)P(S = s|X = x, \Theta = \theta, E = e)}^{\text{Residual effect}}$$
$$\times \frac{\partial P(SGA = 1|X = x, \Theta = \theta, E = e, S = s)}{\partial \theta_t}.$$

where $t = C, S, P$, alternatively. This formula shows that maternal traits can affect the newborn SGA in three ways: indirectly through the education (summand 1) and smoking (summand 2) choices, and directly (summand 3).[46]

The results of this decomposition exercise are presented in Table 3.17 and Figure 3.7. We find that 70% and 95% of the overall effect on newborn SGA of cognitive and social endowments, respectively, passes through choices. More specifically, 68% of the overall effect of cognitive skills passes through the channel of education, and only 2% of it works by affecting smoking behavior. Moreover, 30% of the overall effect of cognitive skills on newborn SGA is a residual, i.e. it is likely to work through other behaviors not included in our model, such as prenatal nutrition.[47] Then, 64% of the effect of social skills works through the smoking choice, while 30% through education. On the contrary, maternal physical constitution has a direct, "residual" effect on newborn health, i.e. it is not mediated by education or smoking decisions. Figure 3.7 displays the same information, but in terms of percentage points reduction in the probability of having a SGA baby, and the respective contributions of the three maternal endowments to it. Importantly, when we repeat the decomposition exercise using LBW as outcome (Figure C-2 in the Appendix), we find a significant residual effect of both cognition and social skills on low birth weight. The lack of biological plausibility of this result reassures us again about our choice of outcome as more genuinely capturing the rate of fetal growth.

[46] In practice we compute average partial derivatives, and investigate the impact of a one standard deviation change in the respective factor. Note that we could also investigate a large movement along the factor distribution using finite differences (i.e. neglecting cross-differences), according to the following formula:

$$
\begin{aligned}
\Delta_\theta P\left(SGA=1|X=x,\Theta=\theta\right) = \sum_{S=0}^{1}\sum_{E=0}^{1} & \left[P(S=s|X,\Theta,E=e)\times P(SGA=1|X=x,\Theta=\theta,E=e,S=s)\right. \\
& \left.\times \quad \Delta_\theta P(E=e|X=x,\Theta=\theta)\right] \\
+ \quad & P(E=e|X=x,\Theta=\theta)\times P(SGA=1|X=x,\Theta=\theta,E=e,S=s)\times \Delta_\theta P(S=s|X=x,\Theta=\theta,E=e) \\
+ \quad & P(E=e|X=x,\Theta=\theta)\times P(S=s|X=x,\Theta=\theta,E=e)\times \Delta_\theta P(SGA=1|X=x,\Theta=\theta,E=e,S=s) \\
+ \quad & P(S=s|X=x,\Theta=\theta,E=e)\times \Delta_\theta P(E=e|X=x,\Theta)\times \Delta_\theta P(SGA=1|X=x,\Theta=\theta,E=e,S=s) \\
+ \quad & P(E=e|X=x,\Theta=\theta)\times \Delta_\theta P(S=s|X=x,\Theta=\theta,E=e)\times \Delta_\theta P(SGA=1|X=x,\Theta=\theta,E=e,S=s) \\
+ \quad & P(SGA=1|X=x,\Theta=\theta,E=e,S=s)\times \Delta_\theta P(E=e|X=x,\Theta=\theta)\times \Delta_\theta P(S=s|X=x,\Theta=\theta,E=e) \\
+ \quad & \Delta_\theta P(E=e|X=x,\Theta=\theta)\times \Delta_\theta P(S=s|X=x,\Theta=\theta,E=e)\times \Delta_\theta P(SGA=1|X=x,\Theta=\theta,E=e,S=s)\big]\,.
\end{aligned}
$$

[47] In other words, smarter women might have healthier babies because of better nutrition. Unfortunately, we are unable to model other maternal behaviors due to data limitations.

Table 3.17: Decomposition of the effects of maternal endowments on newborn SGA

	Cognitive Skills	Social Skills	Body Size
Education	0.681	0.302	0.005
Smoking	0.02	0.64	-0.02
Factor residual	0.30	0.06	1.01
TOTAL	1.00	1.00	1.00

Note: Numbers in cell show the percentage of the overall effect of each maternal trait which works through the education and smoking choices, and the residual effect. National Child Development Study (NCDS), Birth cohort 1958. The analytical sample on which these estimates are based consists of all female cohort members that have no missings in any of the covariates.

We now move to the estimation of a variety of treatment effects of education and smoking in the next section.

3.5.4 The Treatment Effects of Education and Smoking

In this section, we investigate the causal impact of education and smoking on the probability of having a baby who is small for gestational age. First, we compute the average treatment effects (ATE) of education and smoking, which can alternatively be interpreted as the average effect of the treatment for a person randomly selected from the population, or as the expected change in the average outcome if the treatment was exogenously given to every individual.[48] Formally, the ATE of maternal education on newborn SGA is defined as:

$$\mathrm{E}\big[SGA^{E=1} - SGA^{E=0}\big] = \int\int \mathrm{E}\big[SGA^{E=1} - SGA^{E=0}|X,\Theta\big]\, dF_{X,\Theta}, \tag{3.6}$$

where $SGA^{E=1}$ and $SGA^{E=0}$ denote the potential outcomes in terms of the probability of delivering a SGA baby, for mothers with and without post-compulsory education.

The well-known fundamental problem of causal inference states that both potential outcomes are never observed for a single individual, but only either $(SGA^{E=1}|X,\Theta,E=1)$ or $(SGA^{E=0}|X,\Theta,E=0)$. Here we make use of the following matching assumption $\mathrm{E}\big[SGA^{E=0}|X,\Theta,E=0\big] = \mathrm{E}\big[SGA^{E=0}|X,\Theta,E=1\big]$ to simulate counterfactual outcomes

[48] In the following we present the formal definitions of treatment effects and we describe the details of their estimation only for the effect of education on newborn SGA. The effects of education on smoking and of smoking on child SGA can be defined analogously.

and estimate treatment effects. In practice, using draws from the posterior distributions of the parameters and of the factors, as well as from the covariates (X) information from our sample, we compute the ATE as follows:

$$\begin{aligned}\mathrm{E}\Big[\widehat{SGA^{E=1}} - \widehat{SGA^{E=0}}\Big] &= \frac{1}{NK}\sum_{k=1}^{K}\sum_{i=1}^{N}\big[\hat{m}^{e=1}(X_i,\Theta_i(k);\Psi(k)) \\ &\quad - \hat{m}^{e=0}(X_i,\Theta_i(k);\Psi(k))\big] \end{aligned} \tag{3.7}$$

where Ψ denotes a vector of all parameters but the factors, and $\hat{m}^e$ denotes the model prediction of $m^e(X,\Psi) = \mathrm{E}[SGA|X,\Theta,E=e]$ and N denotes the number of individuals in the sample.

The second treatment effect we estimate is the Average Treatment effect on the Treated (ATT). This measures the impact of the treatment on an individual drawn at random from the population and actually been treated (in our case, who selected into education). The ATT of maternal education on newborn SGA is thus defined as:[49]

$$\mathrm{E}\big[SGA^{E=1} - SGA^{E=0}|E=1\big] = \int\int \mathrm{E}\big[SGA^{E=1} - SGA^{E=0}|X,\Theta,E=1\big]\, dF_{X,\Theta|E=1},$$

We compute the ATT as follows:

$$\begin{aligned}\mathrm{E}\Big[\widehat{SGA^{E=1}} - \widehat{SGA^{E=0}}|E=1\Big] &= \frac{1}{N_1 K}\sum_{k=1}^{K}\sum_{i:e=1}^{N_1}\big[\hat{m}^{e=1}(X_i,\Theta_i(k);\Psi(k)) \\ &\quad - \hat{m}^{e=0}(X_i,\Theta_i(k);\Psi(k))\big]\,. \end{aligned} \tag{3.8}$$

Lastly, we define and estimate the average marginal treatment effect (AMTE), which is the average effect of the treatment for the marginal person, i.e. the person who is indifferent between participation ($E=1$) and nonparticipation ($E=0$) into the treatment (Carneiro et al., 2010). The AMTE is defined as:

$$\begin{aligned}\mathrm{E}\big[SGA^{E=1} - SGA^{E=0}|\,|Z_E\beta_E + U_E| < \epsilon\big] &= \int\int E\left[SGA^{E=1}\right. \\ &\quad \left. - SGA^{E=0}|X,\Theta,|Z_E\beta_E + U_E| < \epsilon\right] dF_{X,\Theta} \end{aligned} \tag{3.9}$$

[49] The Average Treatment Effect on the Non-Treated (ATNT) can be analogously defined as:

$$\mathrm{E}\big[SGA^{E=1} - SGA^{E=0}|E=0\big] = \int\int \mathrm{E}\big[SGA^{E=1} - SGA^{E=0}|X,\Theta,E=0\big]\, dF_{X,\Theta|E=0},$$

and can be computed as follows: $\frac{1}{N_0 K}\sum_{k=1}^{K}\sum_{i:e=0}^{N_0}\left[\hat{m}^{e=1}(X_i,\Theta_i(k);\Psi(k)) - \hat{m}^{e=0}(X_i,\Theta_i(k);\Psi(k))\right]$

with Z_E and U_E being the observed and unobserved determinants of the educational choice, as defined in Section 3.3, and ϵ close to zero.[50] The AMTE is computed as follows:

$$\mathrm{E}\Big[\widehat{SGA^{E=1}} - \widehat{SGA^{E=0}} \,|\, |Z_E\beta_E + U_E| < \epsilon\Big] = \frac{1}{N_{mar}} \sum_{k=1}^{K} \sum_{i:mar}^{N_{mar}} \big[\hat{m}^{e=1}(X_i, \Theta_i(k); \Psi(k)) - \hat{m}^{e=0}(X_i, \Theta_i(k); \Psi(k))\big]. \quad (3.10)$$

where *mar* denotes the marginal individual.

Table 3.18: Treatment effects of smoking and education

	ATE	ATT	ATNT	AMTE
EDU on SMOKING in Pregnancy	-0.098	-0.094	-0.1	-0.1
	(0.01)	(0.009)	(0.01)	(0.01)
EDU on SGA probability	-0.016	-0.02	-0.014	-0.016
	(0.009)	(0.008)	(0.009)	(0.009)
SMOKING on SGA probability	0.111	0.099	0.115	0.11
	(0.009)	(0.009)	(0.009)	(0.009)

Note: National Child Development Study (NCDS), Birth cohort 1958. The analytical sample on which these estimates are based consists of all female cohort members that have no missings in any of the covariates. The numbers in columns 2-5 are the treatment effects, as specified: ATE=Average Treatment Effect; ATT=Average Treatment Effect on the Treated; ATNT=Average Treatment Effect on the Non-Treated; AMTE=Average Marginal Treatment Effect. The last column displays the average outcome for the untreated group. Standard errors in brackets.

The estimated treatment effects are reported in Table 3.18.[51] The important message that we learn from this table is that the various average treatment effects are not statistically significantly different from each other, suggesting the absence of evidence of sorting on health gains: in other words, when deciding whether to continue education beyond the compulsory level, or whether to smoke in pregnancy, women do not take the unobserved components of health of the baby into account. Additionally, when we use LBW rather than SGA as outcome (see Table C-2 in the Appendix), we find treatment effects of smoking which are half the magnitude, once more reassuring us of our choice of using SGA as a better indicator of health at birth.

[50] The value chosen for ϵ is 0.3, with the net utility taking values from -6 to +6. Using this metric, 15% of individuals in our sample are defined to be "at the margin".

[51] We can also calculate partial treatment effects, for example by fixing the education treatment assignment when calculating the average treatment effect of the smoking treatment: $\mathrm{E}\big[SGA^{S=1}_{E=1}\big] - \mathrm{E}\big[SGA^{S=0}_{E=1}\big]$. For the sake of brevity these results are not presented here.

In order to gain a better understanding of the magnitude of these average treatment effects, we now decompose the difference in means between the outcomes of treated and untreated women, shown in the bottom panel of Figure 3.1. The observed difference in the occurrence of SGA between high and low educated mothers can be decomposed into average treatment effect, sorting gain and selection bias according to:[52]

$$\overbrace{\mathrm{E}[SGA|E=1]-\mathrm{E}[SGA|E=0]}^{\text{observed difference}} = ATE + \overbrace{(ATT-ATE)}^{\text{sorting gain}} + \overbrace{\mathrm{E}\left[SGA^{E=0}|E=1\right]-\mathrm{E}\left[SGA^{E=0}|E=0\right]}^{\text{selection bias}}.$$

The sorting gain is the difference between the ATE and the ATT, and it is informative as to whether individuals sort into a treatment based on their expectations about their idiosyncratic gains or losses. The results of this decomposition exercise are presented in Table 3.19 and graphically displayed in Figure 3.8. We find stark differences between the two treatments. On the one hand, the differences in prenatal smoking behaviors and newborn health that we observe between high- and low-educated women are only partially due to education: selection accounts for between 40% and 60% of this difference. On the other hand, the difference in the probability of delivering a SGA baby observed between women smoking and non-smoking in pregnancy (which amounts to 11 pp) can be entirely attributed to the harmful effects of smoking. Although we do not decompose it, it is worth noticing that the selection bias component reflects differences in the distributions of observables and latent endowments, as well as differences in their support.[53]

3.5.5 Treatment Effects Heterogeneity

We now go beyond mean impacts and investigate whether education and smoking policies have different effects along the distribution of maternal endowments. This is important, since, when allocating public resources under constraints, policy makers will need to know who benefits the most from a given policy, so to allow a more effective targeting. In

[52]The observed difference in smoking between high- and low-educated women, and in the occurrence of SGA between smokers and non-smokers, can be defined analogously.

[53]Support differences in the trait distributions between treated and nontreated individuals become apparent when looking at Figure 3.2 and Figure 3.3. Heckman et al. (1997) show that the selection bias component can be further decomposed into components due to differing supports of X and Θ for the treated and nontreated groups, into differing distributions of X, Θ over the same support in the two groups, and into differences in outcomes that are present even after controlling for observables and unobservables.

Table 3.19: Decomposition of the observed difference in the outcomes into causal and selection components

	Observed Diff	ATE	Sorting Gain	Bias	% Sorting gain + Bias
EDU on SMOKING in Pregnancy	-0.166	-0.098	0.004	-0.073	41.059
EDU on SGA probability	-0.049	-0.016	-0.004	-0.029	66.850
SMOKING on SGA probability	0.102	0.111	-0.012	0.003	8.759

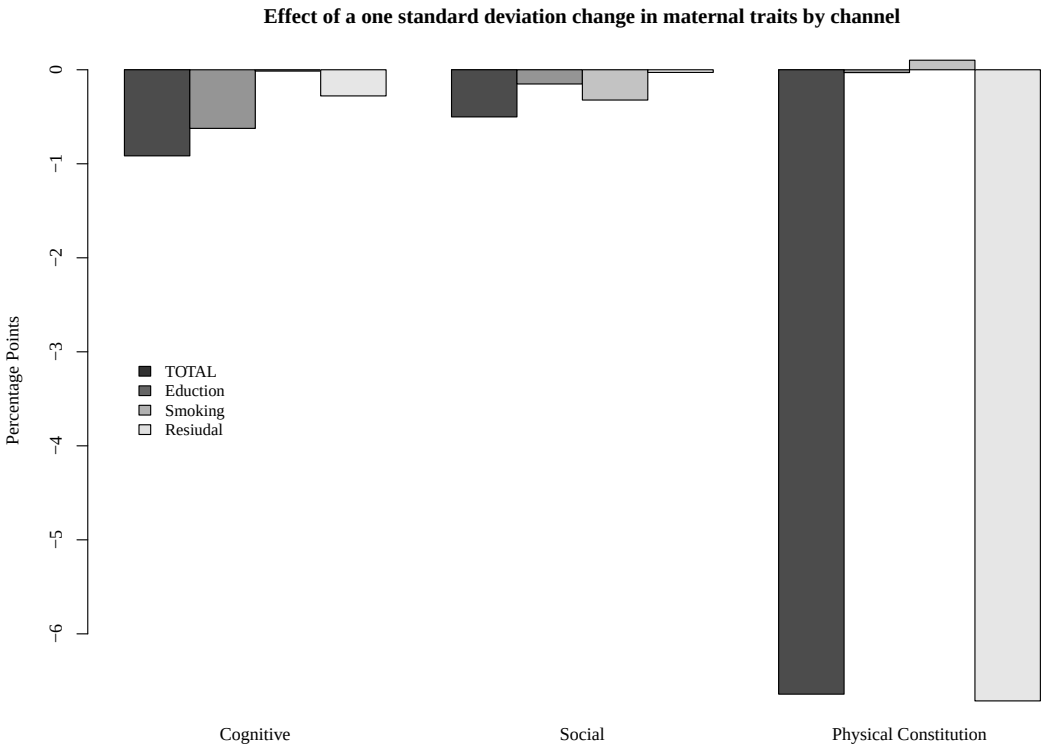
Note: National Child Development Study (NCDS), Birth cohort 1958. The analytical sample on which these estimates are based consists of all female cohort members that have no missings in any of the covariates. The second column shows the difference in the observed outcomes between the treated and the untreated. The third, fourth and fifth columns decompose this difference into the average treatment effect, the sorting gain and the evaluation bias, respectively. The sixth column displays the percentage of the observed difference attributable to the selection component.

practice, we compute average treatment effects for individuals at different quantiles of the distributions of each of their latent traits in turn, while fixing the other two at their mean value. For example, to calculate the effect of education on the probability of delivering a SGA baby for a woman with an average endowment of physical and social traits, but below the 20^{th} percentile of the cognitive skills distribution, we compute:

$$\begin{aligned} \mathrm{E}\Big[\widehat{SGA^{E=1}} - \widehat{SGA^{E=0}}\Big] &= \frac{1}{NK}\sum_{k=1}^{K}\sum_{i=1}^{N}\big[\hat{m}^{e=1}(X_i, \theta_C(p_{0-20})(k), \bar{\theta}_S(k), \bar{\theta}_{BS}(k); \Psi(k)) \\ &\quad -\hat{m}^{e=0}(X_i, \theta_C(p_{0-20})(k), \bar{\theta}_S(k), \bar{\theta}_{BS}(k); \Psi(k))\big]. \end{aligned}$$

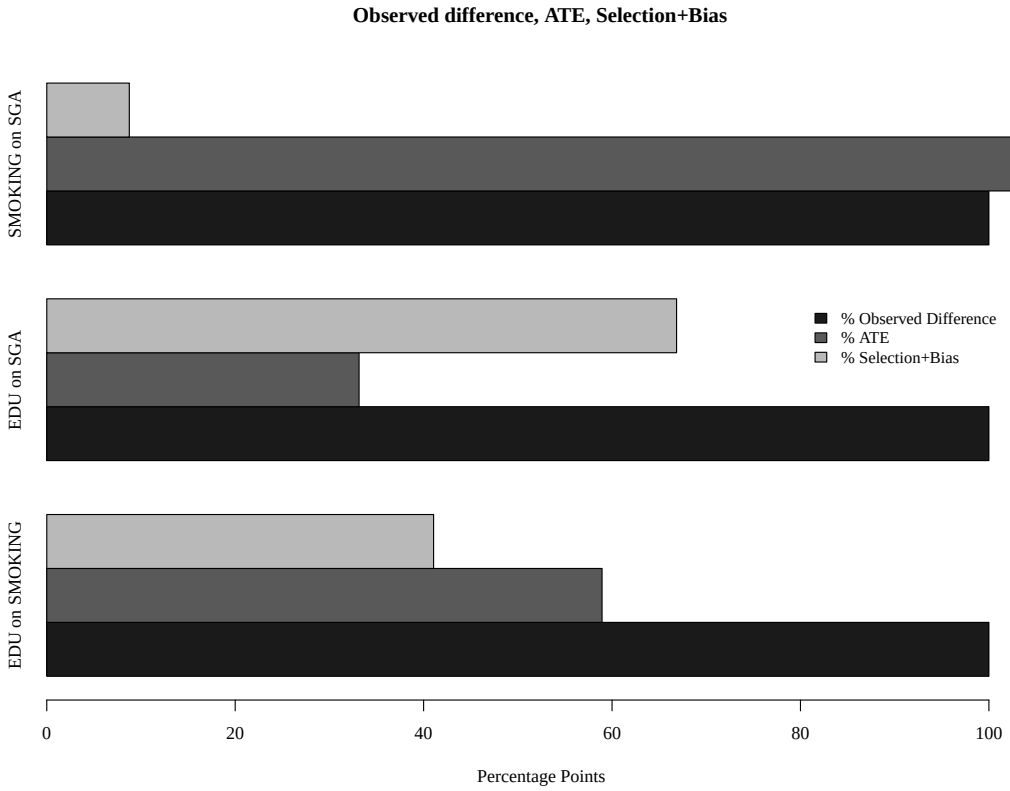
The results of this exercise are presented in Table 3.22 and displayed in Figure 3.9. We make several observations. First, although not by a large magnitude, the treatment effect of education on smoking is bigger at the top of the distribution of each of the maternal endowments. However, we find no significant evidence of heterogeneity in the effect of education and of smoking on the probability of delivering a SGA baby along the distribution of cognitive and social skills. This is a relevant difference with respect to the results we obtain when using low birth weight instead of SGA as measure of newborn health (see Figure C-3 in the Appendix), where we find that the treatment effect of education on the probability of delivering a low birth weight baby is significantly higher at the *bottom* of the maternal traits distribution. Nonetheless, we find significant evidence of treatment effect heterogeneity along the distribution of physical traits, which suggests the presence of complementarities.

Figure 3.7: Decomposing the effects of maternal endowments on newborn SGA



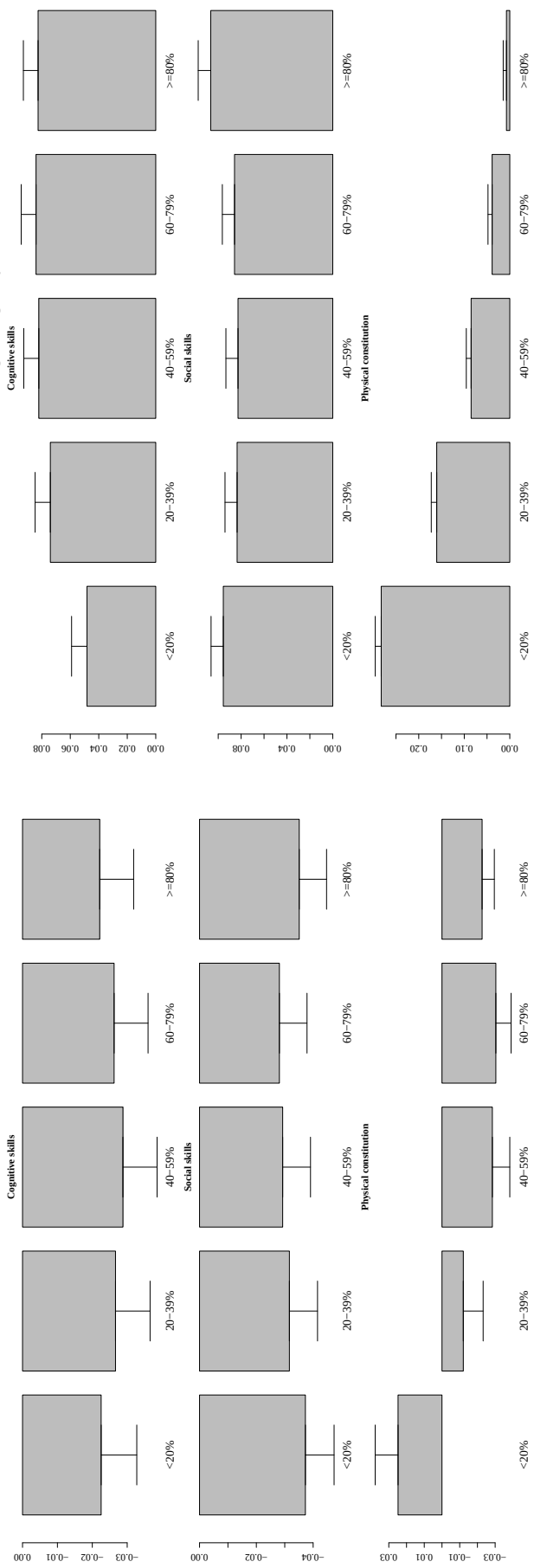
Note: National Child Development Study (NCDS), Birth cohort 1958. The analytical sample on which these estimates are based consists of all female cohort members that have no missings in any of the covariates. The length of the bar "Total" shows the overall reduction in the probability of delivering a SGA baby which is associated with a one standard deviation increase in each of the three maternal endowments. The respective contributions of the various channels are shown in the bars "education", "smoking" and "residual", respectively.

Figure 3.8: Decomposition of the observed differences in the outcomes into causal and selection components



Note: National Child Development Study (NCDS), Birth cohort 1958. The analytical sample on which these estimates are based consists of all female cohort members that have no missings in any of the covariates.

Figure 3.9: Heterogenous treatment effects of education and smoking on SGA along the distribution of maternal traits



Note: National Child Development Study (NCDS), Birth cohort 1958. The analytical sample on which these estimates are based consists of all female cohort members that have no missings in any of the covariates. Whiskers display standard errors.

Table 3.22: Heterogeneity in average treatment effects of smoking and education along the distribution of maternal traits

	<20%	20-39%	40-59%	60-79%	>=80%
EDU on SMOKING in Pregnancy (C)	-0.09	-0.10	-0.11	-0.11	-0.12
EDU on SMOKING in Pregnancy (S)	-0.08	-0.10	-0.11	-0.11	-0.11
EDU on SMOKING in Pregnancy (BS)	-0.09	-0.10	-0.11	-0.11	-0.12
EDU on SGA probability (C)	-0.02	-0.03	-0.03	-0.03	-0.02
EDU on SGA probability (S)	-0.04	-0.03	-0.03	-0.03	-0.04
EDU on SGA probability (BS)	0.02	-0.01	-0.03	-0.03	-0.02
SMOKING on SGA probability (C)	0.05	0.07	0.08	0.08	0.08
SMOKING on SGA probability (S)	0.10	0.08	0.08	0.09	0.11
SMOKING on SGA probability (BS)	0.28	0.16	0.08	0.04	0.01

Note: National Child Development Study (NCDS), Birth cohort 1958. The analytical sample on which these estimates are based consists of all female cohort members that have no missings in any of the covariates. The table shows the average treatment effect of education on smoking (rows 1-3) and on newborn SGA (rows 4-6), and of smoking on SGA (rows 7-9) at different quantiles of the traits distribution (with the other two traits in turn fixed at their mean values).

On the one hand, we find that the effect of education in reducing the probability of having a SGA baby is bigger for mothers with a healthier physical constitution. On the other hand, smoking during pregnancy has a much more detrimental effect for mothers with a poor physical constitution; this effect declines along the distribution, and it only has a small and insignificant impact for females who are physically very fit. Interestingly, these women are also more likely to choose smoking during pregnancy (see Figure 3.5).[54] This finding is consistent with a model where there is heterogeneity in health endowments and women sort into utility-generating risky prenatal behaviors by acting upon knowledge of the consequences of their choices: in other words, those women with a better physical constitution are both more likely to smoke, and at the same time less likely to suffer from the adverse consequences of their behavior, in terms of a reduction in their newborn's health.[55] Our results have important policy implications: they suggest that anti-smoking policies should target women who are physically more prone to have smaller babies, since these are the women whose newborns suffer most from prenatal smoking.

[54]Note that both results have been obtained while fixing the other two endowments at their means, in turn.

[55]See, for example, Rosenzweig and Schultz (1983).

3.6 Conclusions and Policy Implications

This chapter has presented novel evidence on the impact of maternal cognitive, social and physical endowments on newborn health. We have used a sequential selection model with a factor structure to analyze the role of maternal endowments and investments (education and smoking in pregnancy) on the probability of having a baby who is small for gestational age (SGA). We have estimated the total impact of maternal endowments on birth outcomes, and we have also decomposed it into a direct, "biological" effect and a "choice" effect, mediated by maternal behaviors. Then, we have estimated the causal effects of maternal education and smoking in pregnancy, and we have investigated whether women endowed with different traits have different returns. We have found that cognition affects birth outcomes primarily through education, that personality traits mainly operate by changing smoking behavior, and that the physical fitness of the mother has a direct, "biological" effect on SGA. We have also found significant heterogeneity in the effects of education and smoking along the distribution of maternal physical traits, suggesting that women with a less healthy physical constitution should be the primary target of prenatal interventions.

Table 3.23: Treatment effects equivalents in terms of standard deviations of the trait distribution

	ATE	Cognitive (sd)	Social (sd)	Body Size (sd)
Education	-0.016	1.771	3.235	0.244
Smoking	0.111	-12.156	-22.2	-1.676

Note: National Child Development Study (NCDS), Birth cohort 1958. Columns 2-4 display the equivalent of the treatment effect in standard deviations of the endowment distribution. For example, in order to achieve the same effect as a policy that moves everybody from compulsory to post-compulsory education, every individual would need to be endowed with 1.8 sd higher cognitive abilities.

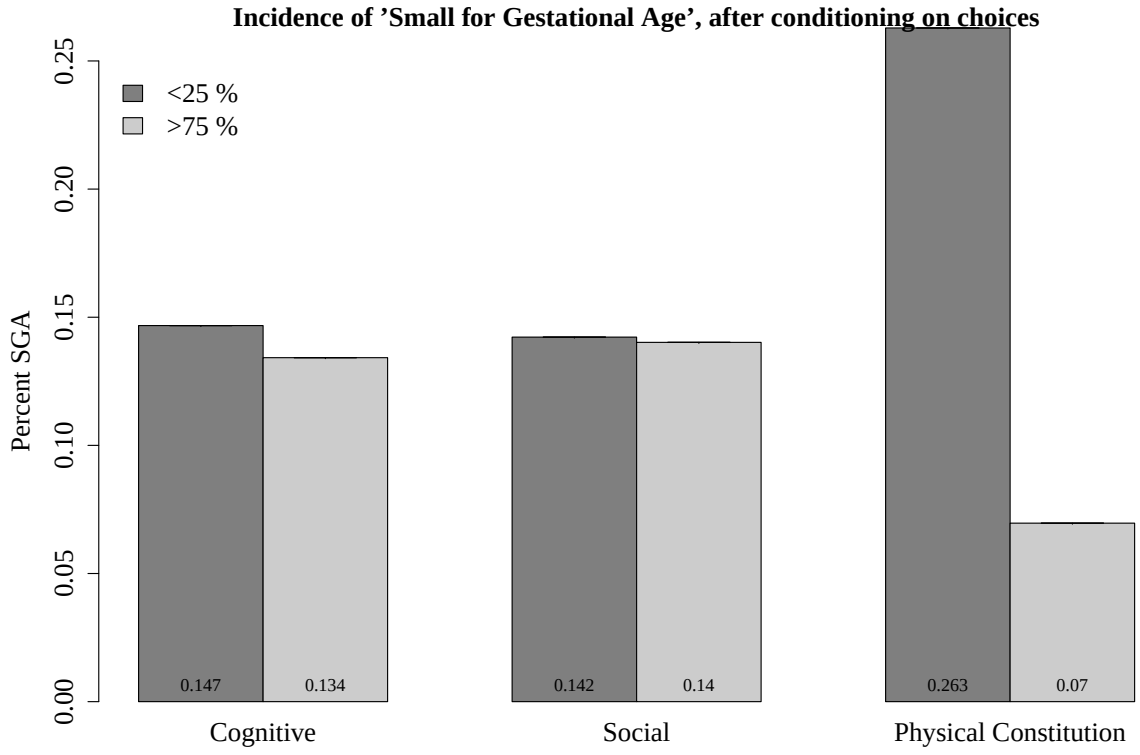
What is then the most effective policy to close the SGA gap by education (4 percentage points) and smoking (11 percentage points)? On the one hand, our results show that a one standard deviation increase in cognitive, social and physical endowments reduce the probability of giving birth to a SGA baby by around 1, 0.5 and 6.5 percentage points, respectively. Additionally, we have found that these endowments operate through different mechanisms: cognitive ability predominantly through education, social skills by affecting the smoking choice, while the physical constitution of the mother directly affects the

health of the newborn. Indeed, our decomposition exercise has shown that 70% and 95% of the gap in SGA between women at the bottom and at the top quartiles of the cognitive and social endowments distributions, respectively, is explained by maternal choices: Figure 3.10 shows that, once we condition on them, no gap in the probability of delivering a SGA baby remains. On the contrary, the SGA differential between women endowed with a poor and with a healthy constitution increases once we condition on the prenatal smoking choice: this occurs because the physical constitution of the mother has both a direct, positive effect on the health of the newborn, and it is also positively associated with her prenatal smoking decisions. On the other hand, we estimate that the average treatment effect of continuing education beyond the minimum compulsory leaving age and of smoking after the third trimester of pregnancy on the probability of delivering a SGA baby are -1.6 and 11 percentage points, respectively.

These results allow us to compare the effectiveness of three different policies in reducing the prevalence of SGA: an early childhood intervention providing both a stimulation and a nutritional component, an educational policy, and a smoking cessation intervention. In Table 3.23, we compare them by computing the change in terms of standard deviations of the endowments distributions that an early childhood intervention would have to achieve in order to cause the same change in the prevalence of SGA as an educational policy or a smoking cessation intervention.[56] What we learn is that a nutritional intervention which improves the physical constitution of the mother appears to be a very effective option. Furthermore, when analyzing heterogeneity in the treatment effects, we find evidence of substitutability: a prenatal antismoking intervention is more effective (in terms of reduction in the probability of delivering a SGA baby) for those women with a poor physical constitution. Of course, a comparison of the overall benefits of the two interventions, and of their respective costs, is necessary before drawing any definite conclusion. However, these results suggest that prenatal interventions, such as home-visiting programs, aimed at (among other things) reducing the prevalence of risky behaviors in pregnant women, and targeting low-income mothers, seem to be an effective way to compensate for maternal endowments differentials, and to guarantee a healthy start of life for the next generation.

[56]So, for example, it shows that an educational policy can achieve the same outcome as an intervention which raises IQ by 1.8 SD.

Figure 3.10: Difference in SGA after conditioning on maternal traits and choices



Note: National Child Development Study (NCDS), Birth cohort 1958. Graph displays outcomes for individuals with endowments in the highest/lowest quartile of the distribution after conditioning on education and smoking. Conditional SGA probabilities are weighted by unconditional smoking and education probabilities according to: $\sum_E \sum_S \left[P(SGA = 1|X = x, \Theta = \theta_{>75}, E = e, S = s) - P(SGA = 1|X = x, \Theta = \theta_{<25}, E = e, S = s)\right] P(S = s|E = e)P(E = e)$

3.A Detailed results

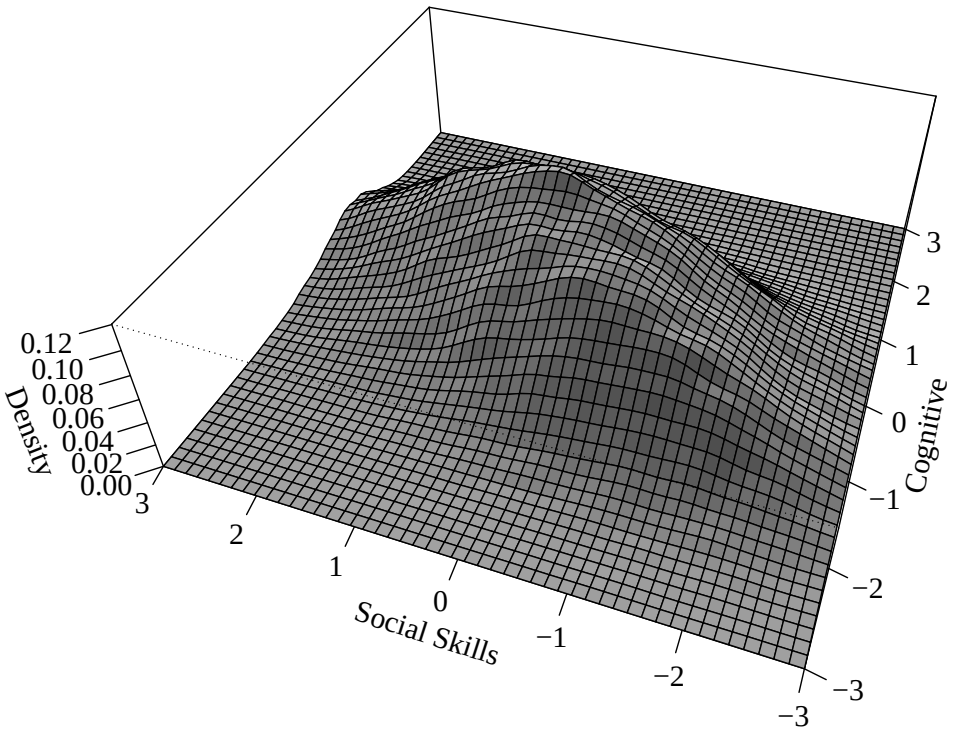
This Appendix shows an extended version of Tables 3.3 and 3.5 which includes the estimated coefficients on all the covariates.

Table 3.24: Full model results (outcome system)

	Education	Smoking(E=1)	Smoking(E=0)	SGA(E=1, S=1)	SGA(E=0, S=1)	SGA(E=1, S=0)	SGA(E=0, S=0)
Cognitive	14.677	-0.616	-0.297	-0.23	1.042	0.153	-1.276
se	(1)	(1)	(1)	(1)	(1)	(1)	(1)
Social	3.727	-3.361	-4.364	0.641	-0.999	-0.803	0.572
se	(1)	(1)	(1)	(1)	(1)	(1)	(1)
Physical constitution	1.271	0.075	0.901	-1.51	-4.083	-2.593	-3.501
se	(1)	(1)	(1)	(1)	(1)	(1)	(1)
Constant	-1.917	-3.061	-1.987	-1.779	-0.907	-1.188	-1.164
se	(0.206)	(0.825)	(0.381)	(1.047)	(0.128)	(0.147)	(0.078)
Child sex (female)				0.07	0.14	0.077	0.111
se				(0.714)	(0.139)	(0.142)	(0.094)
Mother first born	0.439	-0.209	-0.3				
se	(0.062)	(0.157)	(0.077)				
Scotland(age16)	0.541	0.255	0.115	0.739	-0.104	-0.503	-0.119
se	(0.119)	(0.203)	(0.114)	(0.812)	(0.2)	(0.288)	(0.173)
London(age16)	0.069	0.249	-0.125	0.222	-0.061	-0.082	-0.262
se	(0.101)	(0.225)	(0.133)	(1.182)	(0.29)	(0.264)	(0.18)
Wales(age16)	0.398	-0.053	0.135	0.763	0.107	-0.015	-0.186
se	(0.115)	(0.264)	(0.141)	(1.487)	(0.268)	(0.264)	(0.243)
% Smoking in Region		6.572	5.704				
se		(1.952)	(0.919)				
Change in unemployment rate	0.153						
se	(0.098)						
Grandmother smokes		0.008	0.163				
se		(0.148)	(0.072)				
Grandmother Post-Compulsory Education	0.653	-0.43	-0.025	0.573	0.135	-0.183	-0.08
se	(0.064)	(0.16)	(0.094)	(0.748)	(0.18)	(0.148)	(0.125)
Grandparental SES High	0.945	-0.098	-0.559				
se	(0.09)	(0.235)	(0.141)				
Grandparental SES Medium	0.406	-0.021	-0.292				
se	(0.071)	(0.198)	(0.072)				
Grandmother Age at Birth of Mother	0.026	-0.002	-0.01				
se	(0.005)	(0.014)	(0.007)				

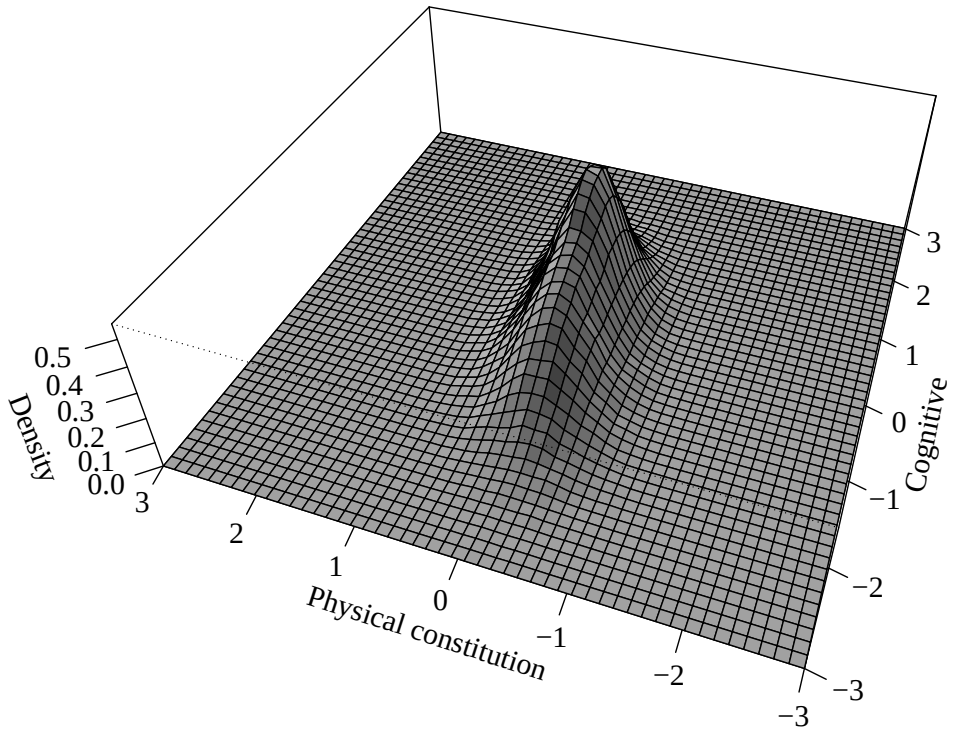
3.B Factor Distributions

Figure 3.11: Joint distribution of cognitive and social skills



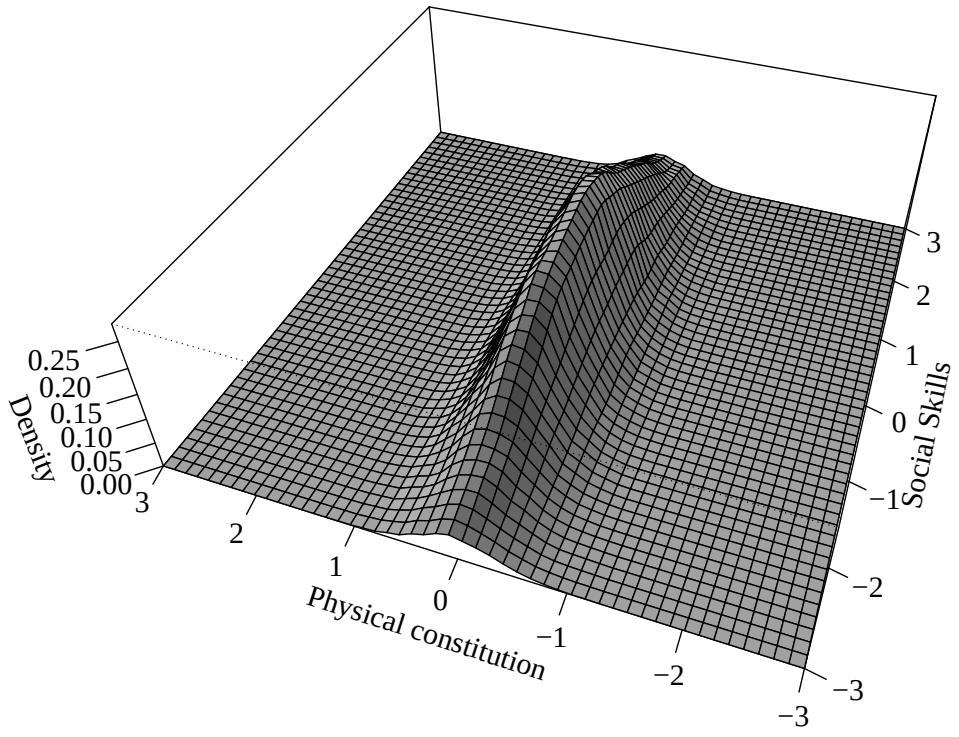
Note: National Child Development Study (NCDS), Birth cohort 1958. The analytical sample on which these estimates are based consists of all female cohort members that have no missings in any of the covariates. **Notes:** Joint mixture distribution of traits.

Figure 3.12: Joint distribution of cognitive skills and physical constitution



Note: National Child Development Study (NCDS), Birth cohort 1958. The analytical sample on which these estimates are based consists of all female cohort members that have no missings in any of the covariates. **Notes:** Joint mixture distribution of traits.

Figure 3.13: Joint distribution of social skills and physical contitution



Note: National Child Development Study (NCDS), Birth cohort 1958. The analytical sample on which these estimates are based consists of all female cohort members that have no missings in any of the covariates. **Notes:** Joint mixture distribution of traits.

3.C Robustness tests

3.C.1 Using low birth weight instead of small for gestational age as measure of newborn health

As already mentioned in Section 3.4.2, size at birth reflects two factors: gestational length and the rate of fetal growth. Hence, if birth weight or low birth weight (less than 2,500 grams at birth) are used as indicators of fetal health as manifested at birth, this leads to confounding effects of growth and maturity. Hence, we use SGA as our main outcome. However, in order to make our results comparable to much of the previous literature, we have re-estimated all our models using low birth weight as an outcome. In our sample, the prevalence of low birth weight is around 7%, therefore lower than the prevalence of SGA. Around 90% of babies who are normal weight are also categorized as normal size for gestational age. However, only a little more than 50% of the babies who are low birth weight are also SGA.

What we learn from this robustness exercise is that, while using LBW as an outcome leaves the main qualitative results unchanged, it does lead to differences in the impact of maternal physical constitution, in the size of the treatment effect of smoking and in the heterogeneity results. We have already discussed these results in the chapter in the different sections where relevant, and we present all of them here now for completeness.

First, Table 3.28 reports the results on the effects of maternal traits on her choices and newborn outcomes (the corresponding SGA Table is 3.5): we notice that, while the effects of the cognitive and social skills are very similar (both quantitatively and qualitatively) to those obtained when using SGA as outcome, the physical constitution of the mother is not a significant determinant of the probability of delivering a low birth weight baby, after conditioning on the education and smoking choices. We then examine the total impact of maternal traits on newborn outcomes: Figure 3.14 shows that the biggest difference with respect to the SGA results is in the effect of maternal physical constitution, while cognition has no effect on the probability of delivering a LBW baby, and social skills a very similar one to the one obtained when using SGA as an outcome. In particular, an early nutritional intervention which would move the mother from the 20^{th} to the 80^{th} percentile of the physical traits distribution would be associated with a reduction in the

Table 3.28: Average marginal effects of a one standard deviation change in maternal traits

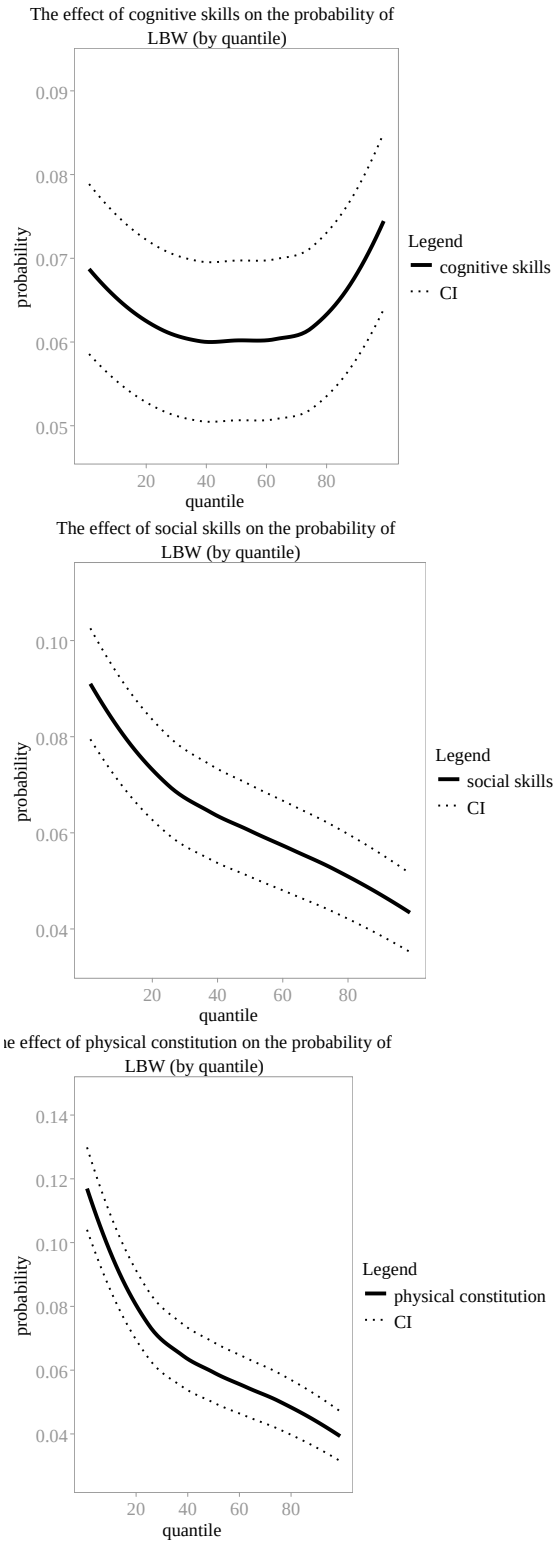
	Cognitive Skills	Social Skills	Physical Constitution
Education	0.153	0.039	0.015
	(0.045)	(0.015)	(0.013)
Smoking(E=1)	-0.012	-0.061	0.003
	(0.022)	(0.038)	(0.026)
Smoking(E=0)	-0.003	-0.056	0.011
	(0.013)	(0.021)	(0.015)
LBW(E=1, S=1)	0.07	-0.005	-0.056
	(0.163)	(0.074)	(0.164)
LBW(E=0, S=1)	0.003	-0.007	-0.039
	(0.017)	(0.017)	(0.026)
LBW(E=1, S=0)	0.01	-0.004	-0.008
	(0.015)	(0.012)	(0.015)
LBW(E=0, S=0)	-0.003	-0.011	-0.007
	(0.009)	(0.009)	(0.01)

Note: National Child Development Study (NCDS), Birth cohort 1958. The analytical sample on which these estimates are based consists of all female cohort members that have no missings in any of the covariates. Standard errors in brackets. E=education; S=smoking; LBW=low birth weight.

probability of delivering a LBW baby of only 3 pp, in contrast to the 10 pp obtained when using SGA as outcome (Figure 3.6).

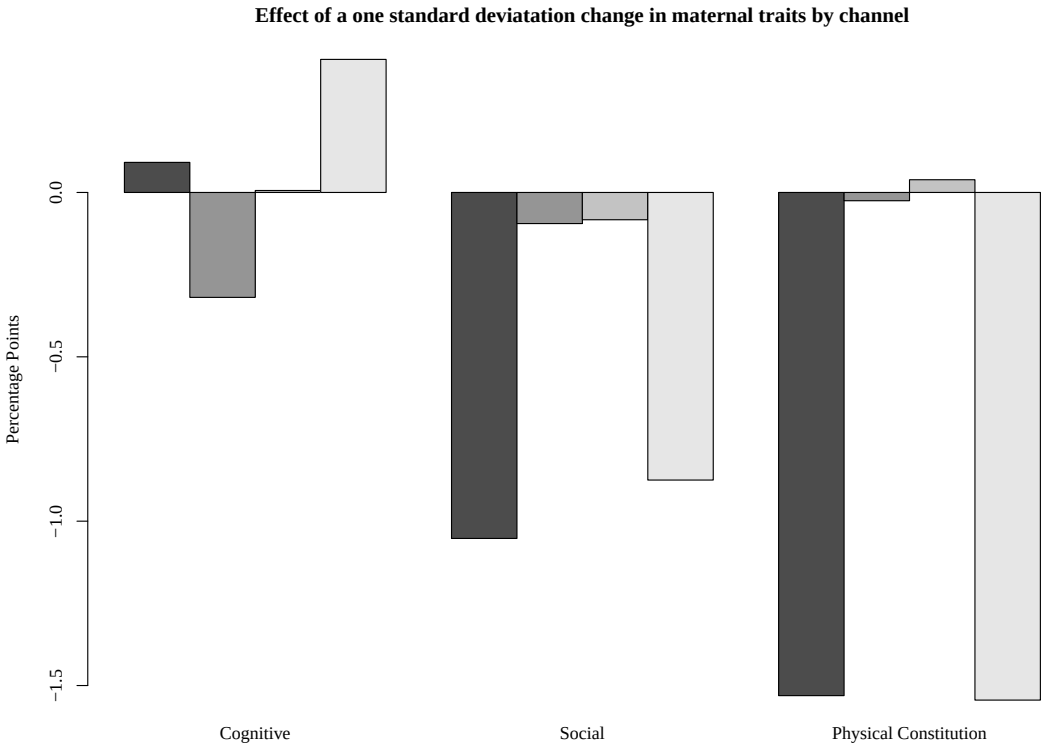
We then decompose the channels through which maternal traits affect LBW, and we report the results in Figure 3.15. As compared to the results obtained using SGA (Figure 3.7), we find a significant residual effect of both cognition and social skills on low birth weight. The lack of biological plausibility of these results reassures us again about our choice of SGA as the main outcome as more genuinely capturing the rate of fetal growth.

Figure 3.14: Effect of maternal traits on newborn LBW



Note: National Child Development Study (NCDS), Birth cohort 1958. The analytical sample on which these estimates are based consists of all female cohort members that have no missings in any of the covariates. When computing the overall effect of each trait in turn on the probability of delivering a LBW newborn, the other two traits are fixed at their respective means.

Figure 3.15: Decomposing the effects of maternal endowments on newborn LBW



Note: National Child Development Study (NCDS), Birth cohort 1958. The analytical sample on which these estimates are based consists of all female cohort members that have no missings in any of the covariates. The length of the bar "Total" shows the overall reduction in the probability of delivering a LBW baby which is associated with a one standard deviation increase in each of the three maternal endowments. The respective contributions of the various channels are shown in the bars "education", "smoking" and "residual", respectively.

Table 3.30: Treatment effect of smoking and education

Treatment effect of	ATE	ATT	ATNT	AMTE
education on smoking in pregnancy	-0.096	-0.092	-0.098	-0.097
	(0.01)	(0.009)	(0.01)	(0.01)
education on the probability of	-0.008	-0.002	-0.011	-0.009
delivering a LBW baby	(0.006)	(0.006)	(0.006)	(0.006)
smoking on the probability of	0.051	0.052	0.051	0.052
delivering a LBW baby	(0.007)	(0.007)	(0.007)	(0.007)

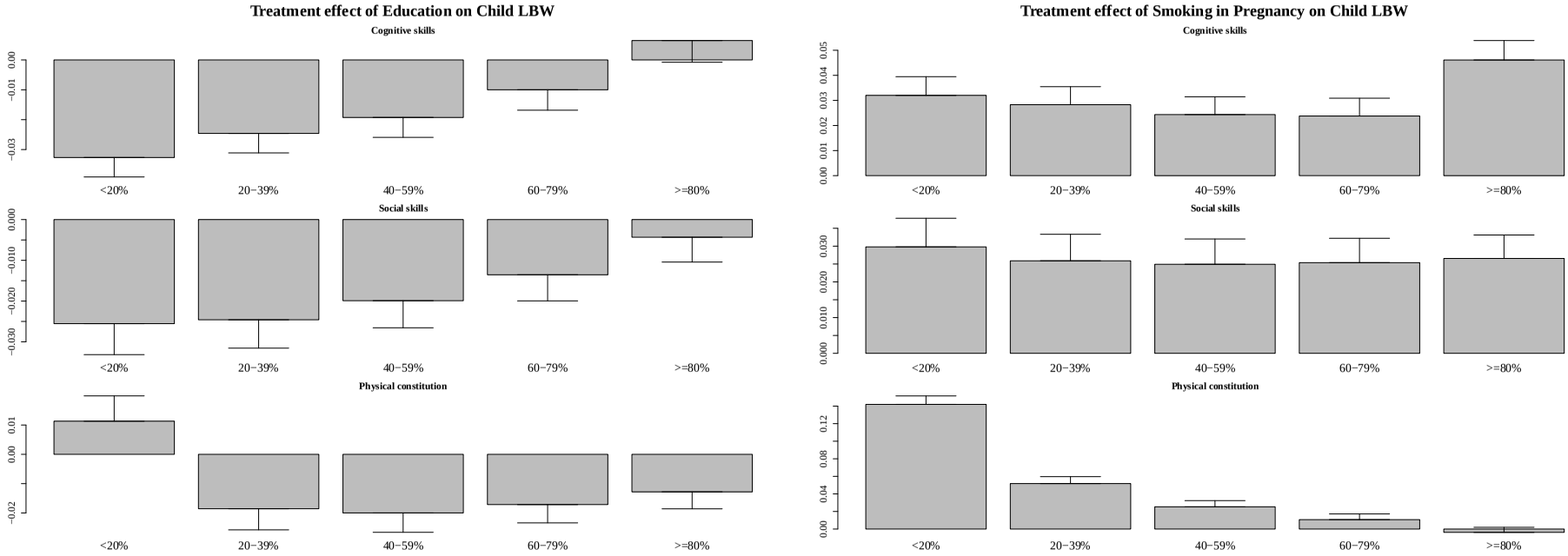
Note: National Child Development Study (NCDS), Birth cohort 1958. The analytical sample on which these estimates are based consists of all female cohort members that have no missings in any of the covariates. The numbers in columns 2-5 are the treatment effects, as specified: ATE=Average Treatment Effect; ATT=Average Treatment Effect on the Treated; ATNT=Average Treatment Effect on the Non-Treated; AMTE=Average Marginal Treatment Effect. Standard errors in brackets.

Table 3.31: Decomposition of the effects of maternal endowments on newborn LBW

	Cognitive Skills	Social Skills	Body Size
Education	354	9	2
Smoking	-10	8	-2.5
Factor residual	-444	83	100.5
TOTAL	100	100	100

Note: Numbers in cell show the percentage of the overall effect of each maternal trait which works through the education and smoking choices, and the residual effect. National Child Development Study (NCDS), Birth cohort 1958. The analytical sample on which these estimates are based consists of all female cohort members that have no missings in any of the covariates.

Figure 3.16: Heterogenous treatment effects of education and smoking on LBW along the distribution of maternal traits



Note: National Child Development Study (NCDS), Birth cohort 1958. The analytical sample on which these estimates are based consists of all female cohort members that have no missings in any of the covariates. Whiskers display standard errors.

We finally compute the treatment effect of education and smoking on the probability of delivering a LBW baby (Table 3.30), and we investigate the presence of heterogeneous effects (Table 3.16). While the treatment effects of education are comparable to those obtained when using SGA as outcome, we find that smoking increases the probability of delivering a LBW baby by around 5pp, which is half of the effect found for SGA (Table 3.18). Lastly, the heterogeneity results reveal that the effect of education on the probability of delivering a LBW baby is significantly higher at the bottom of the cognitive and social skills distribution, while it has basically no impact at the top (instead, we found homogeneity in the effects of maternal education along the cognitive and social skills distributions when using SGA as outcome, see Figure 3.9); comparable to those obtained when using SGA as an outcome are, instead, the treatment effects of smoking on LBW along the distribution of maternal traits (see again Figure 3.9).

3.C.2 Restricting sample to full term deliveries

As mentioned in Section 3.4.2, our main outcome of interest SGA (Small for Gestational Age) can be measured with error. There are two main reasons: 1) the variable is based on the recall of the last menstrual period (LMP) and 2) the charts for SGA are constructed based on the distribution of live birth at any point in gestation. The problem related with the first source of measurement error is minimized for full term deliveries. The intuition is simply that the lesser the expected date (as based on the LMP) and the actual date coincide, the higher the probability that the expected date was wrong. In fact, for deliveries in the 37-41 weeks range, up to 96% of them happen within one week of the expected date (Strauss (2000), Poulsen et al. (2011), Wingate et al. (2007), Kramer et al. (1988), Mustafa and David (2001)). The second problem leads to consistent underestimation of SGA infants in early deliveries, because SGA is defined as being below the 10% of the distribution of birth weight by gestation for live births. Yet, most healthy infants will not be born premature, so problem 2) is comparable to a missing data problem (Hutcheon and Platt, 2008) and would also be minimized by restricting the sample to deliveries in the 37-41 weeks range. Hence, restricting the sample to full term births would minimize the problems discussed above, but at the same time would have the important drawback of loosing important variation from the data, by discarding all the premature births.

Table 3.34: Average marginal effects of a one standard deviation change in maternal traits

	Cognitive Skills	Social Skills	Physical Constitution
Education	0.156	0.038	0.017
	(0.045)	(0.015)	(0.015)
Smoking(E=1)	-0.024	-0.069	0.001
	(0.026)	(0.044)	(0.032)
Smoking(E=0)	-0.009	-0.059	0.029
	(0.015)	(0.024)	(0.02)
SGA(E=1, S=1)	0.009	0.11	-0.17
	(0.121)	(0.173)	(0.253)
SGA(E=0, S=1)	0.02	-0.012	-0.097
	(0.031)	(0.028)	(0.049)
SGA(E=1, S=0)	-0.001	-0.028	-0.061
	(0.024)	(0.025)	(0.056)
SGA(E=0, S=0)	-0.018	0.004	-0.044
	(0.014)	(0.016)	(0.022)

Note: National Child Development Study (NCDS), Birth cohort 1958. The analytical sample on which these estimates are based consists of all female cohort members that have no missings in any of the covariates. Standard errors in brackets. E=education; S=smoking; SGA=small for gestational age. Sample is restricted to full-term babies (37-41 weeks of gestation).

While the results in the chapter are based on a sample which includes all gestational ages, in this appendix we have also re-estimated all the models by restricting the sample to full-term deliveries. As shown in Tables 3.34[57] and 3.35,[58] the results are comparable. This suggests that the presence of measurement error in SGA does not constitute a serious issue in our data.[59]

3.C.3 Using different references for birth weight by gestation

The last robustness test that we perform refers to the use of a difference reference chart, since there is no one unique chart for birth weight by gestational age that is unanimously recognized as the gold standard. We have chosen the one most commonly adopted in the literature: the Babson and Benda's chart, as updated by Fenton. This is also the table featured in standard Neonatology manuals in the USA (e.g, see Gomella et al. (1999)). Another common chart recommended for the USA is the one proposed in Alexander et al. (1996). There exists a debate on whether charts should be population specific or if

[57] This corresponds to Table 3.5 in the main text.
[58] This corresponds to Table 3.18 in the main text.
[59] We thank Heather Royer for suggesting us to perform this robustness test.

Table 3.35: Treatment effect of smoking and education

Treatment effect of	ATE	ATT	ATNT	AMTE
education on smoking in pregnancy	-0.079	-0.086	-0.075	-0.08
	(0.011)	(0.01)	(0.011)	(0.011)
education on the probability of	-0.004	-0.016	0.002	-0.004
delivering a SGA baby	(0.01)	(0.009)	(0.01)	(0.01)
smoking on the probability of	0.118	0.094	0.125	0.114
delivering a SGA baby	(0.01)	(0.01)	(0.01)	(0.01)

Note: National Child Development Study (NCDS), Birth cohort 1958. The analytical sample on which these estimates are based consists of all female cohort members that have no missings in any of the covariates. The numbers in columns 2-5 are the treatment effects, as specified: ATE=Average Treatment Effect; ATT=Average Treatment Effect on the Treated; ATNT=Average Treatment Effect on the Non-Treated; AMTE=Average Marginal Treatment Effect. Sample is restricted to full-term babies (37-41 weeks of gestation). Standard errors in brackets.

instead there should be one, universal chart. The latter approach has been adopted for children, in which case the WHO growth charts are recognized as the standard. Charts for fetal growth, instead, tend to be population specific and even tailored to maternal characteristics (Gardosi, 2006), because of the well-known impact of maternal physical constitution on newborn weight.[60]

Importantly, the difference between the USA and the UK reference charts is minimal. Nonetheless, in order to provide one more check of the robustness of our results, we have also re-estimated all the models using the latest birth weight by gestational age charts for Great Britain, which are constructed on the basis of births from Scotland (Bonellie et al., 2008), and are the latest tables adopted in official publications.[61]

As expected, the results, reported in Tables 3.36 and 3.37, are very similar to those obtained by using the US charts – which reassures us once more of the robustness of our findings.

[60]Notice we do not need to use a "tailored" growth chart since we explicitly include maternal physical constitution in our model.

[61]The classical UK reference chart is outdated (Thomson et al., 1968), and these are constructed on the basis of a much bigger number of births than the original ones.

Table 3.36: Average marginal effects of a one standard deviation change in maternal traits

	Cognitive Skills	Social Skills	Physical Constitution
Education	0.153	0.04	0.016
	(0.045)	(0.014)	(0.013)
Smoking(E=1)	-0.012	-0.062	0.001
	(0.023)	(0.039)	(0.026)
Smoking(E=0)	-0.004	-0.055	0.013
	(0.012)	(0.021)	(0.016)
SGA(E=1, S=1)	-0.045	-0.004	-0.125
	(0.092)	(0.101)	(0.169)
SGA(E=0, S=1)	0.015	-0.022	-0.088
	(0.024)	(0.022)	(0.042)
SGA(E=1, S=0)	-0.005	0.005	-0.043
	(0.018)	(0.019)	(0.032)
SGA(E=0, S=0)	-0.015	0.006	-0.042
	(0.012)	(0.013)	(0.023)

Note: National Child Development Study (NCDS), Birth cohort 1958. The analytical sample on which these estimates are based consists of all female cohort members that have no missings in any of the covariates. Standard errors in brackets. E=education; S=smoking; SGA=small for gestational age. SGA is defined using Bonellie et al. (2008) table.

Table 3.37: Treatment effect of smoking and education

Treatment effect of	ATE	ATT	ATNT	AMTE
education on smoking in pregnancy	-0.097	-0.094	-0.099	-0.098
	(0.01)	(0.009)	(0.01)	(0.01)
education on the probability of	-0.002	-0.013	0.004	-0.004
delivering a SGA baby	(0.009)	(0.008)	(0.009)	(0.008)
smoking on the probability of	0.104	0.105	0.104	0.105
delivering a SGA baby	(0.009)	(0.009)	(0.009)	(0.009)

Note: National Child Development Study (NCDS), Birth cohort 1958. The analytical sample on which these estimates are based consists of all female cohort members that have no missings in any of the covariates. The numbers in columns 2-5 are the treatment effects, as specified: ATE=Average Treatment Effect; ATT=Average Treatment Effect on the Treated; ATNT=Average Treatment Effect on the Non-Treated; AMTE=Average Marginal Treatment Effect. Standard errors in brackets. SGA is defined using Bonellie et al. (2008) table.

3.D Three-Step Estimation Procedure

As mentioned in Section 3.3 we use Bayesian MCMC methods to estimate the parameters of our sequential selection model with a factor structure; however, we also present results from a stepwise procedure which uses factor scores as proxies for factors, similar to the one in Heckman et al. (2013a). The advantage of using factors scores instead of a simple (unweighted) sum of measures is that the weights (loadings) are not required to be uniform across items, but instead reflect the estimated correlation between each item and the latent factor. We perform the following three steps. First, we estimate the parameters of the measurement system. Second, we predict factor scores for each individual, using the estimated parameters obtained in the first step. Third, we include these predicted scores as observed covariates in choice and outcome equations.

While this three-step approach avoids poor convergence, multiple local maxima, and instability of the model with respect to estimated parameters, as compared to full maximum likelihood estimation, the method produces biased coefficients in the outcome equations of the model. If not bias-corrected, coefficients are plagued by attenuation bias by a standard errors-in-variables argument (Croon, 2002). We correct for this bias following a procedure due to Iwata (1992), which is similar to the one described in Lu and Thomas (2008) and employed in Heckman et al. (2013a). However, our method slightly differs from theirs, as we correct the factor scores before using them in the outcome system; additionally, our method has the advantage that it can also be applied to some nonlinear models, like the probit, which we use.

3.D.1 Factor Score Prediction

Here we detail the second step of our three-step estimation procedure. Let's start by assuming a linear relationship between the vector of measurements (M) and the vector of factors (Θ), so that the measurement system for agent i, $i \in \{1, ..., N\}$ can be written as:

$$M_{k,i} = \sum_{s=1}^{S} \lambda_{k,s}\theta_{s,i} + \epsilon_{k,i}, \quad \text{for } k = 1, ..., K \quad \textit{and} \quad s = 1, ..., S$$

or in matrix notation $M_i = \Lambda\Theta_i + \varepsilon_i$. Denote the covariance matrices of ε_i and M_i as $V[\varepsilon_i] = \Sigma$, and $V[M_i] = \Omega$. Furthermore, assume that $\theta_i \perp\!\!\!\perp \epsilon_{k,i}$ and that

$\epsilon_{j,i} \perp\!\!\!\perp \epsilon_{k,i} \quad \forall \quad k \neq j$. Hence, ε_i are mutually independent uniquenesses that capture the stochastic measurement error component and Λ is a matrix of factor loadings. Furthermore, denote $E[\Theta\Theta'] = \Phi$ and assume $E[\Theta] = 0$ and $E[\varepsilon] = 0$ to set the location of the factors. Hence, the covariance matrix of the measurements can be written as $\Omega = \Lambda\Phi\Lambda' + \Sigma$. In practice, we assume that each factor loads on a distinct set of measurements (so-called "dedicated"), so that Λ is also distinct for each system of measurements. Furthermore, we set the scale of the factors by normalizing the factor loading of the first measurement equation for each factor equal to 1. Last, we set $K \geq 3$ for each factor to ensure identification.

After having estimated the covariance matrices and the parameters of the vector Λ in a first step, we now aim to estimate a vector of factor scores $\widehat{\Theta_i}$ that approximates the true vector of skills Θ_i for each individual i. We use Bartlett's estimator which is based on the minimization of the mean squared error (MSE) of the above equation, because of its desirable unbiasedness properties (Saris et al., 1978; Skrondal and Laake, 2001). Bartlett's factor scores are given by:

$$\widehat{\Theta_i}^{B'} = \left(\hat{\Lambda}'\hat{\Sigma}^{-1}\hat{\Lambda}\right)^{-1}\hat{\Lambda}'\hat{\Sigma}^{-1}M_i$$

Hence, Bartlett's estimator is a GLS procedure which provides the maximum likelihood estimates of $\widehat{\Theta_i}^{B'}$, conditional on the data M_i.[62]

3.D.2 Bias Correction

Finally, to correct for the attenuation bias that arises from using estimated instead of true factors, we use a simple estimation method proposed by Iwata (1992), which may be applied to linear models as well as a certain class of nonlinear models. The idea is to replace the unobserved true vector of latent factors by some estimate and then to estimate the regression parameters in the usual way. Consider the following model:

$$Y_i = \Lambda\Theta_i^* + \varepsilon_i$$

[62]Notice that the data are used twice: first to obtain estimates of the factor loadings, then to predict the factor scores, treating the factor loadings as regressors.

where

$$\Theta_i^* = \Theta_i + V_i$$

where Θ_i is the measured Θ_i^*. Furthermore, it is assumed that $\Theta_i \perp\!\!\!\perp V_i$, $\mathrm{E}[[]\, V_i] = 0$ and that $(\Theta_i, \varepsilon_i)$ are iid. Denote the covariance matrix of Θ_i^* as $Cov(\Theta_i^*, \Theta_i^*) = \Sigma_{\Theta_i^*\Theta_i^*}$, the covariance matrix of V as $Cov(V, V) = \Omega$ and $\Sigma_{\Theta_i^*\Theta_i^*} = \Sigma_{\Theta_i\Theta_i} - \Omega$. A consistent estimator of $\Sigma_{\Theta_i\Theta_i}$ can be obtained when estimating the measurement system as explained in the previous section. If we used Θ_i instead of Θ_i^*, the OLS estimator denoted as $\mathbf{L} = (\Theta'\Theta)^{-1}\Theta'\mathbf{Y}$ would be inconsistent. However,

$$\Lambda = (\Sigma_{\Theta^*\Theta^*})^{-1}\Sigma_{\Theta\Theta}\mathbf{L} \tag{3.11}$$

is consistent. Equation 3.11 can be rewritten as:

$$\begin{aligned}
\Lambda &= (\Sigma_{\Theta^*\Theta^*})\,\Sigma_{\Theta\Theta}\,(\Theta'\Theta)^{-1}\Theta'\mathbf{Y} \\
\Lambda &= \left(\Sigma_{\Theta^*\Theta^*}\,(\Sigma_{\Theta\Theta})^{-1}\Theta'\Theta\,(\Sigma_{\Theta\Theta})^{-1}\Sigma_{\Theta^*\Theta^*}\right)^{-1}\Sigma_{\Theta^*\Theta^*}\,(\Sigma_{\Theta\Theta})^{-1}\Theta'\mathbf{Y} \\
\Lambda &= \left(\hat{\Theta}'^*\hat{\Theta}^*\right)^{-1}\hat{\Theta}^{*'}\mathbf{Y},
\end{aligned}$$

where:

$$\hat{\Theta}^* = \Theta\,(\Sigma_{\Theta\Theta})^{-1}\Sigma_{\Theta^*\Theta^*}.$$

Hence, $\hat{\Theta}^*$ is a consistent estimator of Θ^*. Furthermore, Iwata (1992) shows that this estimator retains consistency in certain classes of nonlinear models, such as the probit model.

Chapter 4

A Validation Study of Intergenerational Effects of Early-Life Conditions on Offspring's Economic and Health Outcomes Potentially Driven by Epigenetic Imprinting[1]

4.1 Introduction

Recently, epigenetic imprinting has become a focal point in medical, biological and epidemiological research on intergenerational effects of nutrition, behaviors and life circumstance. This chapter validates and extends recent findings from this literature according to which nutritious shocks in one generation of individuals can transmit to later generations via epigenetic imprinting. *Epigenetics* is defined as the process by which patterns of gene expression are modified in a relatively stable and heritable manner through methylation of the chromatin. Methylation involves the addition of a methyl group to the DNA base, which can turn down a gene's activity or switch it off entirely. *Epigenetic imprinting* implies that shortly after conception, when stem cells are formed, some of the methyl tags from previous generations remain, causing heritable changes in gene functioning that are not caused by changes in the DNA sequence. Methyl markers are passed on through the germ line, with potentially different expressions of the maternal and paternal alleles in the offspring. Epigenetic modifications may depend on the sex of

[1]This chapter is joint work with Gerard van den Berg.

the parent who transmits it and can lead to intergenerational non-genetic inheritance of lifetime experiences across generations (Hochberg et al., 2011).

So far, almost all empirical evidence on epigenetic transmission stems from experiments on mice, while research on humans is extremely rare.[2] The reason is that evidence on humans has to rely on non-experimental data, which leads to identification problems if nutrition in childhood is endogenously related to unobserved characteristics, that also influence outcomes more directly. Hence, to identify the causal effect of food deprivation on subsequent generations, an exogenous shock in nutrition on one generation is needed, as well as outcomes for later generations. Using famine variation from the Dutch Hunger Winter, there exists evidence that nutrition is an important driver of epigenetic modifications in a single generation of humans (Tobi et al., 2009). Intergenerational epigenetic transmission for three generations has only been studied in a single line of research papers using historical harvest data and church registers from a remote are in Northern Sweden(Bygren et al., 2001; Kaati et al., 2002, 2007; Pembrey et al., 2006). The authors find that low paternal grandfather's food supply in pre-adolescence is associated with a lower mortality risk ratio of sons and grandsons, while low paternal grandmother's food supply is linked to a lower mortality risk of their granddaughters. Besides, the authors find low food supply during the paternal grandfather's pre-puberty phase to be associated with lower third generation mortality from cardiovascular diseases, and higher diabetes mortality with a surfeit of food. The authors postulate that these effects are potentially triggered by methylation of epigenetic marks in the sperm during the ancestor's slow growth period (SGP), which took place at ages 8-10 for girls and at ages 9-12 for boys. They hypothesize that the SGP is a sensitive period for the methylation of male sperm, with methyl tags being transmitted to subsequent generations via epigenetic imprinting (Pembrey et al., 2006).

The above-named sequence of research papers has evoked a lot of discussion and controversy: It is explorative in a sense that the authors look at the effect of either a surfeit of food or poor food availability for six ancestors (4 grandparents and 2 parents), during several age periods, and on several outcomes. Furthermore, samples are rather small, ranging from ca 100-300 individuals. For such a large number of parameters and

[2]Thus far, intergenerational inheritance of epigenetic states has been demonstrated in agouti-mice and rats through paternal and maternal transmission (Anway, 2005; Rakyan et al., 2003).

given the small sample size, it is possible to find effects that prove unimportant in other samples.

Hence, in this chapter we assess the external validity of these findings and extend the analysis to schooling and mental health outcomes. Specifically, we examine adult outcomes of subsequent generations following first-generation exposure to the German famine of 1916-1918 during the slow growth period. First of all, we investigate second generation longevity. Secondly, because the third generation individuals in our sample are too young to have died of a natural cause, we use first and second generation height as a universally accepted proxy for health outcomes, which is positively and almost linearly related to life expectancy at birth (Waaler, 1984; Steckel, 2008). In addition, we look at mental health and secondary schooling. We choose these measures, because they are influenced by early childhood circumstance and nutrition (Neugebauer et al., 1999; St Clair et al., 2005; Neelsen and Stratmann, 2011) and related to epigenetic modifications (Gräff and Mansuy, 2008; McGowan et al., 2008; Radtke et al., 2011). Furthermore, schooling and mental health are closely linked with cognitive and noncognitive skills, which are developed during childhood and have an important impact on economic outcomes later in life (Cunha and Heckman, 2008; Heckman et al., 2006).

In regard of the previous literature, we expect an individual's famine exposure during the slow growth period to be positively associated with second and third generation outcomes. Furthermore, if famine exposure during SGP affects methylation of the male gametes but not the female ones[3], we expect only male SGP famine exposure to affect offspring results. Moreover, any causal impact of the famine, epigenetic or not, should be stronger for individuals who suffered from the famine for a longer period of time during their SGP.

Using the German famine as an exogenous shifter in nutrition has several advantages and disadvantages. First and foremost, it provides us with a large exogenous shock to a generation of individuals whose children and grandchildren are living today and for whom a large number of outcomes is available. Furthermore, probands were affected by the famine at different ages, such that sensitive periods can be separately identified. We also have to address a number of problems. Following the famine, Germany was hit by the Spanish influenza, such that famine and influenza effects cannot be disentangled. Besides,

[3]The female ovum fully develops during fetal development.

the second world war influenced individuals, and later generations possibly survived the war at different rates. Historical factors, related to environmental influences or upbringing, may have affected the SGP famine cohort differently from adjacent cohorts. We control for that by including a large number of controls on population growth and GDP at time of birth, birth year dummies and background controls. Last, because we lack information on methylation patterns, epigenetic imprinting cannot be pinned down as the unique cause of our findings and other mechanisms or explanations are possible, biological and non-biological. Non-biologically, a famine at pre-pubertal age may lead to more mature behavior. Elder (1999) for example, investigating the impact of the Great Depression on children born in 1920-1921, finds that experiencing economic hardship around the age of 10 led to more resilience and psychological strength. Biologically, if the paternal grandmother (PGM) was affected by the famine at ages 8-10, her behavior could more easily be transmitted to granddaughters than grandsons due different degrees of X-chromosome relatedness along the paternal line. Granddaughters share more genetic material with their PGMs than grandsons do (Fox et al., 2010).

We use data from the German Socioeconomic Panel (SOEP), a large household panel that is representative for the German population. The data are well-suited for our analysis because they allow us to identify whether a first generation of individuals (usually the parents of SOEP respondents) was affected by the famine during the slow growth period. Furthermore, they contain information on a wide range of health information, longevity and economic outcomes for the second and third generation.

Our findings suggest that among second generation individuals, maternal famine affectedness during SGP reduces mortality among males, and is associated with lower adult height. Besides, second generation males tend to be taller if their father has been affected. Paternal grandfather SGP-famine seems to increase mental health of third generation sons while maternal grandmother SGP-famine has a positive effect on granddaughter mental health. We do not find robust effects for schooling.

The chapter proceeds as follows. Section Section 4.2 gives an overview over the most recent developments in the epigenetic literature and explains their relevance for economic research. Section 4.3 describes the famine we use as an exogenous shock and summarizes the evidence obtained so far for it. Section 4.4 describes the empirical approach. In Section 5.4 we describe our data. Section 5.6 presents our main findings for generations

2 and 3. We investigate how the results change if we move from controlling only for a basic set of variables to controlling for additional background variables, which potentially account for non-biological channels. Besides, we present robustness checks where we vary famine intensity and period. Section 4.7 concludes.

4.2 Epigenetics and Economics

Biological circumstances during early life have been shown to have important effects on longevity, health and economic outcomes. So far, the economics literature has focused almost exclusively on first-generation effects that emerge from *in utero* or early-life exposure to famines, influenza or even rainfall (Maccini and Yang, 2009; Doblhammer et al., 2013; Almond and Mazumder, 2005; Lumey and Stein, 1997). Most of these studies find detrimental effects of adverse shocks on adult outcomes. In biology, such effects, termed fetal-programming, are well-known and can be produced in mammals by exposing offspring in utero to food restrictions on the pregnant female (Barker, 1995; Nathanielsz, 2003; Whitelaw, 2006).

Much fewer studies have investigated whether there exists an association of first generation exposure to reduced food supply with second and third generation outcomes. Painter et al. (2008) find that gestational famine exposure was associated with reduced offspring length in the next generation and that children of famine-exposed mothers were more likely to be in poor health from acquired neurological, auto-immune, respiratory, infectious, neoplastic, or dermatological problems. In other studies, ancestral food supply was found to affect birth weight, risk of stillbirth, perinatal death and longevity of the second or even third generation (Bygren et al., 2001; Lumey and Stein, 1997; Kaati et al., 2007). Moreover, Pembrey et al. (2006) analyze the effect of smoking during the SGP by exploring UK data on parental interviews of newborn children. They found that the sons of fathers who smoked during their SGP had higher body mass index as 9 year olds. Last, Kaati et al. (2002) and Kaati et al. (2007) study an exogenous variation in nutrition triggered by a food shortage in northern Sweden by collecting records of harvests and food prices during the 19th Century. They show that individuals experiencing food shortages in the SGP, had descendants with lower risks of mortality from cardiovascular disease and diabetes. In particular, the authors find that the mortality risk ratio of grandsons is adversely affected by their paternal grandfather's SGP-exposure to rich food supply while

the paternal grandmother's food supply during her SGP adversely affected the mortality risk of her granddaughters.

The molecular basis for environmentally-triggered nongenetic intergenerational effects is not fully known to date, but the most prominent hypothesis is that it involves epigenetics.[4] Fraga et al. (2005) show that epigenetic patterns are formed over the entire life-course: while 3-year-old monozygotic twins have almost identical methylation patterns, their DNA methylation differs markedly at the age of 50. Besides, findings by Heijmans et al. (2008) and Tobi et al. (2009) support the presumption that epigenetic modifications in humans are related to nutrition. The authors find sex-specific differences in methylation patterns between individuals who have been prenatally exposed to the famine and their same-sex siblings.

Epigenetic inheritance implies that methylation patterns in one generation influence gene expression in the next. How such epigenetic transmissions or inheritance in humans works biologically is not fully resolved (Harper, 2005). Shortly after conception, when the first cell divisions are taking place, the stem cells are generally cleared of all methylation (Ahmed, 2010; Mayer et al., 2000). However, if epigenetic modifications take place on the part of the genome that is genetically imprinted, this could explain sex-specific epigenetic inheritance. 'Imprinted genes' keep their methyl tags (about 1% of genes), which function as a biological marker to flag up their maternal or paternal origin (Masterpasqua, 2009). Economists and social scientists are merely interested in whether adverse experiences can be transmitted non-genetically from one generation to the next rather than in the exact mechanism.[5] Any non-genetic transmission of experiences, would revolutionize economic thinking about intergenerational transmission and human capital accumulation in at least two ways. First, if life experiences were transmitted, not only via upbringing and social circumstance but also biologically, from one generation to the next, this would imply that that the costs and benefits of any policy measure would have to be re-evaluated to include such biological effects on subsequent generations. Second, nature (genetic predisposition) and nurture (upbringing) were found to be inseparable and the long-fought nature-nurture debate would become obsolete. In the future, models of human capital investment would

[4]Other potential mechanisms are DNA amplification or changes in telomere length (Kaati et al., 2007).

[5]For an overview of molecular genetics and economics see Lundborg and Stenberg (2010). For Epigenetics and Psychology see Harper (2005).

need to account for gene expressions and gene-environment interactions as well as for critical and sensitive periods of epigenetic transmission.

Kaati and co-authors argue that the slow growth period of a child may be such a sensitive period. It takes place at ages 9-12 for boys and at ages 8-10 for girls and is thus the developmental period just before onset of puberty. In this period, the first male gametes mature, which may make it a sensitive period for reprogramming of methylation imprints (Pembrey, 2002). This period of childhood is also known as the 'fat spurt': growth is low and the body is accumulating reserves for in anticipation of the puberty-related development spurt (Marshall and Tanner, 1968; Gasser et al., 1994; Gasser, 1996). It is plausible that limited food availability during the 'fat spurt' leads to worse pubertal development and imprinting on the sperm or egg. Indeed, this growth period has previously been found to be critical for development. Sparén et al. (2004) for example find that a famine at this age increases cardiovascular problems later in life. Similarly, Van den Berg and Gupta (2007) and Van den Berg, G.J. and Lundborg, P. and Nystedt, P. and Rooth, D.O. (2010) find this age period to be critical for life expectancy and adult height, respectively.

4.3 The Famine

The World War 1 famine in the German empire is said to be the severest famine experienced in Europe outside of Russia since Ireland's travail in the 1840s (Raico, 1989). At the end of the war, the German 'Reichsgesundheitsamt' (Health Office) calculated that 763,000 German civilians had died from starvation.[6]

Four factors had led to the extreme shortage of food. First, by mid-1916 the Allied Powers had successfully enacted a complete naval blockade of Germany restricting the maritime supply of raw materials and foodstuffs. Before the war Germany had imported one third of its food, but after the blockade Germany was cut from foodstuff imports of all sorts: fodder for livestock, grain and potatoes. Importantly, the blockade continued even after the Armistice and until June 1919 to force Germany to sign the Treaty of Versailles. In

[6]The overall population of the German empire at that time was about 65 million. In addition there were about 2 million military deaths, who in a conventional ground-based war like WW1, were almost exclusively men of age 17-60.

fact, throughout 1919 rationing was maintained in many parts of the country at a rate of 1000-1300 calories per day (Vincent, 1985).[7]

Second, due to the general war mobilization, around 40% of the male agricultural labor force was absent and a similar fraction of horses and cattle. This reduction in the male work force was not adequately compensated by employment of prisoners of war, women, adolescents and children (Huber and Fogel, 1920). As a consequence, between 1913 and 1919 annual production of crops, potatoes and milk decreased to about 50% between 1913 and 1919 when compared to pre-war levels (Blum, 2011).

Third, in the summer 1916, the root crop and grain harvest were particulary bad and the potato crop failed almost completely. The latter was particularly detrimental, because much of the German food supply was based on potatoes and during the war more agricultural crop land had been shifted away from turnip cultivation and towards potatoes (Klein, 1968). The Winter of 1916-1917 thus marked the climax of the famine and is today remembered as the 'turnip winter' (Steckrübenwinter), because the only food in sufficient supply during that winter were turnips.

Last, food storage was a concern. Before the war most of the potato crop was stored in the countryside and only supplied to the cities on demand. After the start of the war, when transportation and dislocation became more difficult, and potatoes had to be stored in larger quantities by individuals unschooled in the proper techniques of storage, which led to spoilage and waste (Vincent, 1985).

The period of food scarcity started in June 1915 when bread began to be rationed, but only in early 1916 food rationing became severe. From 1916 to mid-1919, the German population had to live on less than 1500 calories (Starling et al., 1919). Yet, because the portion of bran in the bread was very large, the calorie value was further reduced by about 15 to 20 percent.[8] Most Germans had to live on a meagre diet of dark bread, slices of sausage without fat, three points of potatoes per week and turnips (Vincent, 1985). Table 4.1 displays an overview over the amount of food consumed during the famine as compared to prewar times. While these amounts are well below subsistence to begin with, the situation was aggravated by the mere length of the famine which started in 1916 and extended into 1919. At the height of the famine, purchasing foodstuffs on the black market

[7]The reason for continued food rationing was that even after the end of the blockade in June 1919 Germany could not import freely, since all funds had to be saved for war reparations.

[8]In comparison, a man needs about 2500 calories a day and a women needs about 2000.

Table 4.1: Food consumption before and during the famine

Item	Before War	During Famine
Calories	2280 per day	1313 per day
Protein	70g per day	30-40g per day
Fat	70g per day	15-20g per day
Bread	225g a) per day	160g per day
Meat	1050g per week	135g per week
Potatoes	100%	71%
Grain	100%	53%
Sugar	100%	49%
Vegetable oil	100%	39%
Meat	100%	31%
Butter	100%	22%
Eggs	100%	18%
Pulse	100%	14%
Cheese	100%	3%

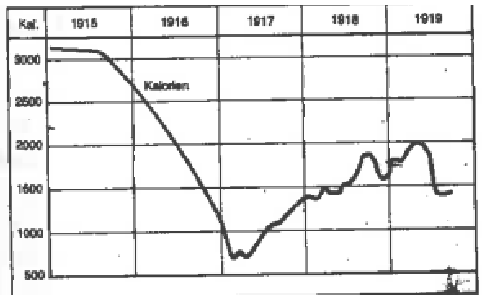
Notes:
Adult quantities reported.
Lower part of the table (percentages) indicate official rations and not actual amounts which were often lower.
Products that vanished almost entirely: cheese, fruit, leather.
a) in 1915.
Sources:
Ernest H. Starling, 1919, Report on Food Conditions in Germany. (London: H.M. Stationary Office) pp. 7-16.
Paul C. Vincent, 1985, The politics of Hunger.
Klein, 1968, Deutschland im ersten Weltkrieg.
For more details on food ratios and average caloric consumption see Table 4.2 and Figure 4.1.

was the only way to prevent starvation. Black market prices in cities quickly skyrocked (see Table 4.3) and many families had to rely on excursions to the countryside, where blackmarket prices where somewhat lower.

The effects on the population were detrimental. On average the German population has lost about 15-25 percent of their weight between 1916 and 1919.[9] Mortality rates for adults rose substantially during this time, but not for children aged 1-15 (Roesle, 1928). Nevertheless, many children suffered from edema, tuberculosis, rickets, influenza, scurvy, and keratomalacia.[10] There even exist studies claiming that the famine had such a damaging effect on German youth that it impaired adult rational thinking and laid ground for later adherence to National Socialism (Loewenberg, 1971). Blum (2011) shows that the average height in the male population born in the period between 1914 and 1917

[9]Individuals who had lost 30 percent or more mostly died.
[10]The number of occurrences of epidemic diseases such as typhoid, rabies, trichiniasis and dysentery stayed roughly constant in the population.

Figure 4.1: Caloric consumption, German Empire



Sources:
reprinted from Arnulf Huegel, 2003, Kriegsernährungswirtschaft Deutschlands während des Ersten und Zweiten Weltkrieges im Vergleich (pp 180).

is around 1.5cm less than for adjacent cohorts, indicating a sharp decline in biological living standards. Moreover, the decline is sharper for individuals of lower socioeconomic classes.

Table 4.2: Food consumption before and during the famine (percent of pre-war level)

Item	July 1916 until June 1917	July 1917 until June 1918	July 1918 until December 1918
Meat	31	20	12
Fish	51	-	5
Eggs	18	13	15
Lard	14	11	7
Butter	22	21	28
Cheese	3	4	15
Rice	4	-	-
Pulse	14	1	7
Sugar	49	56-67	80
Vegetable oil	39	41	17
Potatoes	71	94	94
Flour	53	47	48

Sources:
reprinted from Arnulf Huegel, 2003, Kriegsernährungswirtschaft Deutschlands während des Ersten und Zweiten Weltkrieges im Vergleich (pp 180).

4.3.1 Spanish Influenza

In 1918/1919, the Spanish Influenza hit many countries all over the world. In Germany, about 150,000 individuals have died as a result of the disease (Vincent, 1985). This number is low when compared to the overall number of deaths that resulted from starvation, but still considerable. The first wave of the pandemic hit Germany in June 1918, the second

Table 4.3: Prices for foodstuffs in Bonn (in Mark), 1914 to 1918

Product	Unit	Official 1914	Official 1917/18	Black market 1917/1918	Increase (%)
Beef	1 lb	1	2.8	4.75	375
Veal	1 lb	1	2.8	5	400
Pork	1 lb	0.8	-	6	650
Gammon	1 lb	1.2	-	13	983
Bacon	1 lb	0.7	2.75	15.5	2.114
Suet	1 lb	0.4	2	14.5	3.525
Lard	1 lb	0.8	5	18	2.15
Concentrated milk	1 can	0.5	1.7	4.5	800
Butter	1 lb	1.3	3.4	14	977
Curd	1 lb	0.2	2.3	3.5	1.65
Eggs	1 piece	0.06	0.4	0.65	983
Colza oil	1 l	0.6	5	21.5	3.483
Salad oil	1 l	1.4	-	24	1.614
Olive oil	1 l	3	-	50	1.567
Rye flour	1 lb	0.15	1.85	4	2.567
Wheat flour	1 lb	0.2	-	4	1.9
Rice	1 lb	0.25	-	8	3.1

Source:
reprinted from Blum, 2011 (original source Roerkohl, Anne, 1991), Government decisions before and during the First World War and the living standards in Germany during a drastic natural experiment (p. 558).

one in the fall and the third one in January 1919 (Witte, 2008). In our study, it is thus not possible to separate the effects of the famine from the effects of the influenza pandemic. However, what characterized the Spanish Influenza was that, for the most part, it was lethal only for individuals of ages 20-40, such that concerns of selective survival can be neglected for our cohorts of interest. Mamelund (2003) using data on Norway, which was neutral during WWI, confirms this presumption. He finds that the cohorts born 1890-1910 experienced the lowest immediate death rates of the Spanish Influenza. Nevertheless, since the continuation of the blockade and the third wave of the Spanish influenza extended well into 1919, we conduct robustness checks with the year 1919 included in our famine period.

4.4 Identification and Outcome Models

We use common coefficient models and matching to identify the effect of SGP famine exposure of first generation (G1) ancestors on second (G2) and third generation (G3) individuals. We thus calculate a famine effect that compares G2 and G3 individuals with

the same background and and birth year, but who differ with respect to exogenous first generation SGP famine exposure. Because, first generation famine-SGP exposure is a historical incident that is exogenous at the individual level, our approach allows us to identify the true impact of having ancestors of a certain age during the famine on the second and third generation.[11]

4.4.1 Famine exposure

To investigate systematically how adult outcomes of G2 and G2 vary with first generation SGP exposure to the famine, we focus on individuals who have at least one ancestor born during the years 1902-1913. This implies that the G1 males of our core sample were too young to have been drafted and G1 females were too young to have conceived a child during the war.[12] Also, none of the individuals in our core sample were born during the war, such that we neither need to account for selective fertility during war times nor for in utero exposure to the famine.[13]

Table 4.4 displays the number of years the famine affected different birth cohorts of first generation individuals during their SGP, defining 1916-1918 as the famine period. Note that all first-generation individuals have been affected by the famine, despite at different ages.[14] We thus identify the effect of being famine exposed during SGP as compared to being affected by the famine at a different point in time. Our data thus allow us to separate the effect of 'being affected by a famine at some point during life' from 'being affected by a famine during SGP'.

Our analysis relies on the assumption that there are no differences in famine survival between individuals affected by the famine during the SGP and the control groups. Hence, we assume that children in their slow growth period have been about equally likely to die from the famine than children that were slightly younger or older at the time. Historical sources seem to back this claim: death rates of children between the ages of one and five

[11] Note that our measure of famine exposure is whether someone had reached a certain age by the time of the famine. This measure thus reflects the intention to treat and not the actual treatment effect. We cannot identify whether someone actually went hungry. It is however save to assume that the intention to treat is similar to the actual treatment effect, because hunger during the German famine was so widespread that compliance was very high.

[12] Only men of ages 17-60 were drafted into the military (Foerster, 1994).

[13] Note that this does not have to be true for all other ancestors. E.g. for third-generation individuals only ONE out of four grandparents has to be born in the period 1902-1913.

[14] A rare exception for a historical event where there is a control group which is largely unaffected is the Dutch Hunger Winter (see e.g. Scholte et al. (2012)).

had risen by fifty percent during the famine, while for children from five to fifteen were only slightly higher (fifty-five percent) (Vincent, 1985). Selection into fertility would be a problem if parents from differen social classes had been more or less likely to conceive children in the periods 1902-1903 (1902-1904) or 1910-1913 (1911-1913) than during the years 1904-1909 (1906-1910). Figure 4.3 however shows that for the time period of births we are analyzing (1901-1914), overall birth rates do not show any systematic pattern. In fact, it was impossible to anticipate a decade earlier.[15]

Table 4.4: Number of years of famine exposure during the slow growth period, first generation males and females.

Birth year	1902	1903	1904	1905	1906	1907	1908	1909	1910	1911	1912	1913
Males	0	0	1	2	3	3	2	1	0	0	0	0
Females	0	0	0	0	1	2	3	2	1	0	0	0
Age during famine	14-16	13-15	12-14	11-13	10-12	9-11	8-10	7-9	6-8	5-7	4-6	3-5

Famine years: 1916-1918. The slow growth period ranges from ages 8-10 and 9-12 for females and males, respectively.

The exogenous shifter in nutrition, $Z_{G1} \in \{0, 1\}$, is defined as an indicator variable of whether G1 ancestors were SGP-exposed to the famine, i.e. of ages 8-10 (females) and of ages 9-12 (males) in 1916-1918. We distinguish between three first-generation cohorts. a) males [females] whose SGP lies in the famine period ($Z_{G1} = 1$): birth years 1904-1909 [1906-1910] b) males [females] born ante 1903 [1904] and in SGP before the famine hit ($Z_{G1} = 0$) and c) males [females] born after 1910 [1911], i.e. who were too young for the famine to have affected them in their SGP ($Z_{G1} = 0$).

Each generation is affected differently by the famine. For G2, $Z_{G2,i}$ is a 2x1 vector with the first entry indicating whether the mother was affected by the famine during her SGP and the second indicating whether the father was affected during that same period. Following the same logic, $Z_{G3,i}$ has four entries: whether paternal grandfather (PGF) was SGP famine affected, whether paternal grandmother (PGM) was SGP famine affected, whether maternal grandfather (MGF) was SGP famine affected and whether maternal grandmother (MGM) was SGP famine affected.

[15] During WW1, on the other hand, the birthrate was falling from thirty per thousand to fifteen per thousand (Vincent, 1985).

4.4.2 Outcome Models

We estimate three different types of outcome models to account for the different distributional properties of the respective outcome variables: A duration model for individual mortality, a discrete choice probit model for the decision to obtain a higher secondary school degree and a linear regression model for the continuous outcomes height and mental health. In all models, x denotes an individual-specific vector of observable characteristics, which always comprises some basic control variables and in some specifications an additional set of background controls. f is a vector of birth year fixed effects (for G2 and G3) that captures any variation that may be cohort or birth year specific. For the third generation (G3), this vector also comprises parental birth year fixed effects to capture for example business cycle effects.[16]

Duration Model

First, to estimate the impact of the famine on longevity of the second generation, we model the hazard of mortality at any given point in time as being multiplicatively separable in a (nonfrailty) hazard function μ_0 and a frailty term (α):

$$h(t) = \alpha\mu_0 \tag{4.1}$$

Following the standard biological literature on modeling mortality, the baseline hazard has the shape of a Gompertz distribution with ancillary parameter γ:

$$\mu_0 = exp(\gamma t)exp(\sum_{i=1}^{N} Z_i'\delta + x_i'\beta + f_t'\eta). \tag{4.2}$$

The frailty distribution α follows a gamma distribution with:

$$\alpha \sim \Gamma(1, \theta) \tag{4.3}$$

Our sample is special in a sense that selection into the sample is conditional on ever having responded to a the household questionnaire. We thus account for left truncation by adjusting the likelihood for the fact that individuals only enter in adult life, using age at first interview as truncation point.

[16]For the effect of business cycle variation on outcomes see e.g. van den Berg et al. (2011).

Probit Model

We model binary outcomes such as disability status or upper secondary schooling as a binary outcome latent index model with $Y_{it} = 1_{[Y_{it}^* > 0]}$, where Y_{it}^* denotes the latent continuous variable. The latent variable in turn is determined by famine exposure, birth year fixed effects and observable control variables. We assume a linear structure and additive separability in the error term:

$$Y_{it}^* = Z_i'\delta + x_i'\beta + f_t'\eta + \epsilon_{it}.$$

The observed binary variable Y_{it} is an indicator variable that is assumed to equal one if the latent variable crosses zero as a threshold $Y_{it}^* > 0$. We estimate a probit model, assuming that $P(Y_{it} = 1|x_i, f_t, Z_i) = \Phi(Z_i'\delta + x_i'\beta + f_t'\eta)$ where Φ denotes the normal cdf.

Linear Regression Model

For continuous outcomes, we estimate the following linear model between outcomes Y_{it}, famine effects and covariates for adult i born in year t:

$$Y_{it} = Z_i'\delta + x_i'\beta + f_t'\eta + \epsilon_{it}.$$

again x_i denotes a vector of control variables and the equation comprises a vector of own birth-year fixed effects (f_t) to capture any variation that may be cohort or birth year specific.

4.5 Data

We use data taken from the German Socioeconomic Panel (SOEP), a representative longitudinal micro-dataset for Germany (Wagner et al., 2007). The data are well-suited for our analysis in that they allow us to identify whether a first generation of individuals (the parents of SOEP respondents) were affected by the famine during their slow growth period. Moreover, the data contain information on a wide range of health information, longevity and economic outcomes for the children and grandchildren of famine affected individuals.

4.5.1 Sample

Our sample comprises 7233 first generation individuals who were born 1902-1913 and had children who later became SOEP respondents. Primary SOEP respondents and constitute the second generation in our sample. G2 children enter the SOEP upon the age of 17 and form the third generation. All G2 and G3 individuals are sampled if at least one parent (G2) or one grandparent (G3) has been born in 1902-1913. Table 4.5 contains the sample sizes and birth years of G2 and G3 individuals, as well as the fraction of SGP famine affected parents (for G2) and grandparents (for G3). The advantage of using the SOEP is that it gives us a second-generation sample that is representative of the German population. However, as with any study on intergenerational transmission, first-generation individuals are sampled only if they have reached reproductive age and conceived children. In our case, this implies that they need to have survived the famine, the Spanish influenza, and World War 2, or had children early. Yet, our sample is representative of Germans living today who, by definition, have direct ancestors that survived these events and had children. We do however account for first-generation individuals being more likely to be sampled if the they had more children, e.g. by conditioning on sibship size.

Table 4.5: Sample size and ancestor famine affectedness.

Variables	Second generation				Third generation			
	Males		**Females**		**Males**		**Females**	
Father famine in SGP	0.40	(0.49)	0.42	(0.49)	-	-	-	-
Mother famine in SGP	0.35	(0.48)	0.33	(0.47)	-	-	-	-
PGF famine in SGP	-	-	-	-	0.29	(0.45)	0.34	(0.48)
PGM famine in SGP	-	-	-	-	0.23	(0.42)	0.24	(0.43)
MGF famine in SGP	-	-	-	-	0.26	(0.44)	0.22	(0.42)
MGM famine in SGP	-	-	-	-	0.19	(0.39)	0.16	(0.37)
Birth year	1938.86	(6.37)	1938.76	(6.63)	1973.17	(7.83)	1974.38	(7.13)
N	2063		2083		715		575	

Source: German Socioeconomic Panel (SOEP).
Own calculations.

4.5.2 Outcome and Control Variables

Our outcome measures (Y_{it}) are height, longevity, mental health and whether an individual has obtained an upper secondary school degree. In the SOEP, age at death can be obtained for individuals who have participated in the survey at least once and who dropped out of the survey because they died. The death year is provided by the SOEP in the person-related metafile (ppfad). Mental health and height measures are obtained

using the most recent information from the biannual SOEP health module. Height is self-reported repeatedly and mental health is measured by the Mental Component Summary Scale (MCS), one of the two sub-dimensions of the SF-12 questionnaire. The MCS is measured on a scale that ranges from 0 to 100 with mean 50 and standard deviation 10. It results from a factor analysis comprising the dimensions 'general mental health', 'emotional functioning', 'social functioning' and 'vitality' each measured on separate scales (for details see Andersen et al. (2007)). We define whether an individual has obtained the German university or technical college entrance diploma (German "Abitur" or "Fachhochschulreife") on hands of the international Comparative Analysis of Social Mobility in Industrial Nations (CASMIN) classification. Table 4.6 comprises summary statistics of all outcome variables.[17]

Table 4.6: Descriptive statistics: Age at death, height, mental health and schooling outcomes by sex and generation.

Variables	Second generation				Third generation			
	Males		**Females**		**Males**		**Females**	
Age at death	71.20	(6.35)	72.36	(6.05)				
Height	175.54	(6.63)	163.74	(6.02)	180.52	(7.04)	167.74	(6.07)
Mental health	52.48	(9.78)	50.43	(10.69)	49.45	(9.18)	47.60	(9.97)
Upper secondary school degree	0.36	(0.48)	0.19	(0.39)	0.49	(0.50)	0.52	(0.50)
N	2063		2083		715		575	

Source: German Socioeconomic Panel (SOEP).
Own calculations.

We define two sets of control variables: basic controls and background controls. Basic controls comprise all variables that account for any bias, that arises because G1 famine affected individuals and their offspring were born in different years from adjacent cohorts. For example, recent literature demonstrates long-run mortality effects of economic and health conditions at birth and during infancy (van den Berg et al., 2011). Such effects would bias our findings if the famine affected generation was born in years with systematically higher or lower growth rates. Hence, we are controlling for business cycle trends and population growth during the year of birth of G1 (see Figure 4.2) Moreover, because the data are a cross section with outcomes measured at different ages of the individuals, we are controlling for age at the time of measurement whenever relevant to the outcome variable. Last, we include parent and individual birth year fixed effects (for

[17] We are not looking at wages, because our data are a cross section of individuals who are sampled at different points in their lives, which makes the computation of permanent income very unreliable.

G2 and G3) to captures cohort or birth year specific variation. For G3, we also include parental birth year fixed effects to capture business cycle fluctuations at parental birth and to capture some of the variation in the probability for an individual to be part of the sample.

We include background controls as additional covariates in some specifications to detect behavioral, non-biological pathways. One of the most important pathway is parental education, reflecting parental cognitive ability, parenting skill, social class and family earnings potential all of which are essential for health and schooling outcomes. We define parental education dummies for different educational degrees in Germany. Besides, we include parental age at birth and the number of siblings, as a proxy for parental resources. Descriptive statistics of all included background variables can be found in Table 4.7.

Table 4.7: Descriptive statistics: Background control variables

	Second generation		**Third generation**	
Father upper secondary school	0.10	(0.30)	0.33	(0.47)
Father intermediate school	0.09	(0.28)	0.17	(0.37)
Mother upper secondary school	0.03	(0.17)	0.21	(0.41)
Mother intermediate school	0.11	(0.31)	0.27	(0.44)
Number of brothers	0.98	(1.06)	0.72	(0.77)
Number of sisters	0.99	(1.10)	0.77	(0.87)
Father's age at individual's birth	33.04	(6.58)	31.46	(5.92)
Mother's age at individual's birth	29.52	(5.71)	28.04	(5.30)
Mother's birth year	1909.29	(4.65)	1945.70	(7.76)
Father's birth year	1905.77	(5.16)	1942.28	(7.51)
N	4146		1290	

Source: German Socioeconomic Panel (SOEP).
Own calculations.

4.6 Empirical Results

The results are discussed in two stages. First, we provide a description of the main findings for G2 and G3 in Section 4.6.1. In particular, we investigate how the results change if we move from controlling only for a basic set of variables to controlling for additional background variables, which potentially account for non-biological channels. Second, we present robustness checks where we vary famine intensity and period, by excluding individuals who experienced the famine for only one year and by extending the famine period to 1919. If famine exposure has a causal effect on outcomes, we expect this effect to become larger and more significant if individuals with weak famine exposure are excluded from the analysis.

4.6.1 Main Results

The coefficients displayed in Tables 4.8 and 4.9 display the effect sizes of parental and grandparental SGP-famine exposure on the outcomes under consideration, conditional on the basic set of control variables and on the combined set of basic and background controls. Since epigenetic inheritance is likely to be sex specific, we perform all analyzes separately for males (left panel) and females (right panel) (Pembrey et al., 2006). We report robust standard errors for height, mental health and schooling models. For G3, standard errors are clustered on the household level, because siblings in that sample mostly have the same history of ancestral famine exposure.

In Table 4.8, the left columns of coefficients for each outcome are coefficients from estimated models with basic controls only. We find that body height is significantly higher for males whose fathers have been SGP exposed to the famine and lower for males with SGP exposed mothers. Mortality of second generation males on the other hand is much lower if mothers have been exposed during SGP.[18] Thus, male individuals turn out shorter with maternal SGP exposure but mortality is lower. Furthermore, the coefficient on parental SGP famine exposure on schooling is negative and significant for males. The right columns display coefficients for models with additional controls for parental education, sibship size and parental age at birth. We conduct a likelihood ratio test for the improvement in model fit due to the additional coefficients and find that controlling for parental background improves the fit to the data. Famine coefficients however hardly change with this additional set of controls. Only the coefficient from the regression of sons' higher secondary schooling on parental SGP famine exposure decreases in size and significance.

[18]Results of a Likelihood ratio test for the presence of unobserved heterogeneity indicates that it is important in the sample of females only.

Table 4.8: Main results: Second generation, parental SGP famine effects on longevity, height, mental health and schooling.

Variables	Males								Females							
	Mortality		**Height**		**Mental health**		**Schooling**		**Mortality**		**Height**		**Mental health**		**Schooling**	
main																
Father famine in SGP	1.013	0.972	0.599*	0.698**	0.0523	0.112	-0.0618**	-0.0442**	1.203	1.156	-0.120	-0.0374	0.974*	0.948*	0.0187	0.0337*
	(0.95)	(0.89)	(0.09)	(0.04)	(0.92)	(0.83)	(0.01)	(0.05)	(0.67)	(0.79)	(0.69)	(0.90)	(0.08)	(0.08)	(0.36)	(0.07)
Mother famine in SGP	0.633**	0.644**	-0.549*	-0.640*	0.135	0.185	0.0305	0.0281	1.110	1.881	-0.276	-0.410	0.770	0.597	0.0223	0.00903
	(0.02)	(0.03)	(0.09)	(0.05)	(0.79)	(0.72)	(0.21)	(0.22)	(0.80)	(0.36)	(0.37)	(0.18)	(0.16)	(0.29)	(0.27)	(0.63)
Observations	2063	2063	2063	2063	2063	2063	2063	2063	2083	2083	2083	2083	2083	2083	2078	2078
pval LLR-test		0.01		0.00		0.00		0.00		0.01		0.00		0.00		0.00
γ	0.406	0.387							0.696	0.893						
θ	2.211	1.210							19.144	29.719						
pval σ^2 frailty	0.102	0.261							0.001	0.001						
Basic controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Background controls	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES
Log lik.	-109.775	-104.859	-6.7e+03	-6.7e+03	-7.6e+03	-7.6e+03	-1.3e+03	-1.1e+03	-108.499	-104.257	-6.7e+03	-6.6e+03	-7.8e+03	-7.8e+03	-960.605	-836.178
R-squared			0.075	0.115	0.043	0.056					0.041	0.065	0.054	0.076		

p-values in parentheses
Source: SOEP Data. Own calculations.
Note: For mortality models coefficients are hazard rates. For schooling models, coefficients reported are average marginal effects.
pval LLR-test is the p-value of a likelihood ratio test for the improvement in model fit due to the background controls. γ is an ancillary parameter that parametrizes the Gompertz baseline hazard function. θ parameterizes the gamma frailty distribution. *pval* σ^2 *frailty* results from a LLR-test for H0='frailty variance is zero'. Standard errors are robust.
Basic controls: GDP and population growth at first generation birth, age at time of measurement (mental health and height), parent and individual birth year fixed effects, indicators for whether first generation individuals were born before 1902 or after 1913.
Background controls: Parental education, parental age at birth of children, number of siblings.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.9: Main results: Third generation, parental SGP famine effects on height, mental health and schooling.

Variables	Males						Females					
	Height		Mental health		Schooling		Height		Mental health		Schooling	
Paternal grandfather famine in SGP	-0.152	-0.00277	1.717*	1.992*	0.00704	0.0487	0.526	0.528	0.266	0.411	0.0326	0.0701
	(0.86)	(1.00)	(0.09)	(0.06)	(0.90)	(0.33)	(0.54)	(0.53)	(0.84)	(0.76)	(0.55)	(0.15)
Paternal grandmother famine in SGP	-0.540	-0.771	-0.324	-0.239	0.0532	-0.0246	-1.484*	-1.420	-0.322	-0.251	0.0448	0.0599
	(0.52)	(0.37)	(0.77)	(0.83)	(0.38)	(0.64)	(0.10)	(0.12)	(0.82)	(0.86)	(0.44)	(0.24)
Maternal grandfather famine in SGP	-0.952	-0.924	-0.113	-0.213	-0.0371	-0.0263	-0.120	0.0918	0.133	0.233	0.0927	0.0632
	(0.29)	(0.31)	(0.92)	(0.84)	(0.54)	(0.61)	(0.90)	(0.92)	(0.93)	(0.88)	(0.12)	(0.24)
Maternal grandmother famine in SGP	0.425	0.544	0.498	0.651	0.0208	0.0601	-0.814	-0.761	3.742**	3.542**	-0.0351	-0.0408
	(0.68)	(0.59)	(0.69)	(0.60)	(0.76)	(0.29)	(0.46)	(0.49)	(0.03)	(0.04)	(0.62)	(0.50)
Observations	715	715	715	715	715	715	575	575	575	575	575	575
pval LLR-test		0.00		0.00		0.00		0.03		0.05		0.00
Basic controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Background controls	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES
Log lik.	-2343.4	-2335.2	-2545.8	-2540.2	-443.4	-378.4	-1801.8	-1798.2	-2082.8	-2079.7	-319.2	-257.5
R-squared	0.170	0.189	0.139	0.152			0.161	0.172	0.174	0.182		

p-values in parentheses
Source: SOEP Data. Own calculations.
Note: Standard errors are robust and clustered on the household level. For schooling models, coefficients reported are average marginal effects. *pval LLR-test* is the p-value of a likelihood ratio test for the improvement in model fit due to the background controls.
Basic controls: GDP and population growth at first generation birth, age at time of measurement (mental health and height), parent and individual birth year fixed effects, indicators for whether first generation individuals were born before 1902 or after 1913.
Background controls: Parental education, parental age at birth of children, number of siblings.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.9 gives a summary of results for G3. In this sample, there are four ancestors who have potentially been affected by the famine during their SGP: the paternal grandfather, the paternal grandmother, the maternal grandfather and the maternal grandmother. Again we estimated all models separately for males and females. Kaati et al. (2007) and Pembrey et al. (2006) argue that the SGP may be a sensitive period for the methylation of male gametes, leading to male imprinting of nutritious shocks. The only results that points in that direction is that paternal grandfather SGP exposure has positive effects on mental health of grandson, while maternal grandmother SGP exposure positively affects granddaughters' mental well-being.[19] Again, background controls improve model fit, but do not reduce the size and significance of mental health coefficients. The results indicate that having a paternal grandfather or a maternal grandmother who has been affected by a famine during SGP improves mental health by about 1.8 points, or 18% of a standard deviation, for males and by about 3 points, or 30% of a standard deviation, for females.

4.6.2 Robustness Checks

We repeat the previous analysis twice, including the full set of basic and background controls in the model. First, we exclude all individuals who were only mildly affected by the famine, i.e. who were only affected by the famine for a single year during the SGP. Second, we extend the famine period to the year 1919, because the blockade and the third wave of the Spanish influenza only ceased in the middle of that year.

[19]Note that in our sample, almost none of the third-generation individuals have died, such that we cannot investigate mortality effects.

Table 4.10: Robustness checks: Second generation, parental SGP famine effects on longevity, height, mental health and schooling.

Variables	Males								Females							
	Mortality		**Height**		**Mental health**		**Schooling**		**Mortality**		**Height**		**Mental health**		**Schooling**	
main																
Father famine in SGP ($\geq$ 2 yrs)	0.828		0.818**		0.636		-0.0234		1.832		0.122		0.138		0.0120	
	(0.45)		(0.05)		(0.30)		(0.40)		(0.37)		(0.74)		(0.84)		(0.59)	
Mother famine in SGP ($\geq$ 2 yrs)	0.654		-0.967**		-0.571		0.00893		1.783		-0.0964		0.608		0.0197	
	(0.10)		(0.02)		(0.37)		(0.76)		(0.41)		(0.80)		(0.39)		(0.42)	
Father famine in SGP (incl. 1919)		1.151		0.873**		-0.0746		-0.0378*		0.924		0.0743		0.196		0.0136
		(0.50)		(0.01)		(0.89)		(0.10)		(0.88)		(0.81)		(0.72)		(0.46)
Mother famine in SGP (incl. 1919)		0.675*		-0.273		-0.222		0.0392*		1.539		-0.187		1.089*		0.0210
		(0.06)		(0.42)		(0.67)		(0.09)		(0.41)		(0.55)		(0.06)		(0.28)
Observations	1539	2063	1539	2063	1539	2063	1539	2063	1558	2083	1558	2083	1558	2083	1558	2083
γ	0.399	0.395							0.939	0.810						
θ	1.203	1.645							36.48	25.04						
pval σ^2 frailty	0.250	0.178							0.000115	0.000879						
Hunger period	$\geq$ 2 yrs	incl 1919	$\geq$ 2 yrs	incl 1919	$\geq$ 2 yrs	incl 1919	$\geq$ 2 yrs	incl 1919	$\geq$ 2 yrs	incl 1919	$\geq$ 2 yrs	incl 1919	$\geq$ 2 yrs	incl 1919	$\geq$ 2 yrs	incl 1919
Basic controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Background controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Log lik.	-67.29	-105.4	-4981.7	-6702.7	-5623.1	-7572.4	-876.9	-1148.5	-85.03	-104.7	-4920.3	-6625.5	-5848.3	-7809.5	-610.3	-838.9
R-squared			0.120	0.115	0.0512	0.0560					0.0829	0.0642	0.0800	0.0756		

p-values in parentheses
Source: SOEP Data. Own calculations.
Note: For mortality models coefficients are hazard rates. For schooling models, coefficients reported are average marginal effects. γ is an ancillary parameter that parametrizes the Gompertz baseline hazard function. θ parameterizes the gamma frailty distribution. *pval* σ^2 *frailty* results from a LLR-test for H0='frailty variance is zero'. Standard errors are robust.
Basic controls: GDP and population growth at first generation birth, age at time of measurement (mental health and height), parent and individual birth year fixed effects, indicators for whether first generation individuals were born before 1902 or after 1913.
Background controls: Parental education, parental age at birth of children, number of siblings.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4.11: Robustness checks: Third generation, parental SGP famine effects on height, mental health and schooling.

Variables	Males						Females					
	Height		**Mental health**		**Schooling**		**Height**		**Mental health**		**Schooling**	
Paternal grandfather famine in SGP ($\geq$ 2 yrs)	-0.400		2.897**		0.000991		-0.540		1.682		0.102*	
	(0.69)		(0.02)		(0.99)		(0.59)		(0.32)		(0.08)	
Paternal grandmother famine in SGP ($\geq$ 2 yrs)	-0.273		-2.980*		-0.0417		-0.527		-0.992		0.115**	
	(0.80)		(0.05)		(0.55)		(0.63)		(0.58)		(0.05)	
Maternal grandfather famine in SGP ($\geq$ 2 yrs)	-2.139**		-0.775		-0.00264		-0.619		0.0258		-0.0163	
	(0.03)		(0.59)		(0.97)		(0.63)		(0.99)		(0.80)	
Maternal grandmother famine in SGP ($\geq$ 2 yrs)	0.806		1.661		-0.0477		-1.005		4.389*		-0.143*	
	(0.55)		(0.33)		(0.53)		(0.54)		(0.08)		(0.08)	
Paternal grandfather famine in SGP (incl. 1919)		-0.340		2.134**		0.0214		0.275		0.519		0.0653
		(0.69)		(0.04)		(0.68)		(0.75)		(0.72)		(0.19)
Paternal grandmother famine in SGP (incl. 1919)		-0.693		0.297		0.0329		-1.353		1.819		0.0120
		(0.42)		(0.79)		(0.54)		(0.15)		(0.18)		(0.82)
Maternal grandfather famine in SGP (incl. 1919)		-1.265		-0.215		-0.0116		0.673		1.816		0.0441
		(0.14)		(0.84)		(0.82)		(0.54)		(0.24)		(0.42)
Maternal grandmother famine in SGP (incl. 1919)		1.006		0.234		-0.0139		0.419		3.402**		-0.00867
		(0.32)		(0.85)		(0.81)		(0.71)		(0.04)		(0.89)
Observations	516	715	516	715	511	715	415	575	415	575	415	575
Hunger period	$\geq$ 2 yrs	incl 1919	$\geq$ 2 yrs	incl 1919	$\geq$ 2 yrs	incl 1919	$\geq$ 2 yrs	incl 1919	$\geq$ 2 yrs	incl 1919	$\geq$ 2 yrs	incl 1919
Basic controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Background controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Log lik.	-1638.9	-2334.3	-1803.7	-2539.9	-244.2	-380.9	-1257.8	-1798.6	-1474.0	-2077.2	-156.3	-260.2
R-squared	0.253	0.191	0.217	0.153			0.274	0.170	0.243	0.190		

p-values in parentheses
Source: SOEP Data. Own calculations.
Note: Standard errors are robust and clustered on the household level. For schooling models, coefficients reported are average marginal effects.
Basic controls: GDP and population growth at first generation birth, age at time of measurement (mental health and height), parent and individual birth year fixed effects, indicators for whether first generation individuals were born before 1902 or after 1913.
Background controls: Parental education, parental age at birth of children, number of siblings.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Tables 4.10 and 4.11 display the coefficients of the robustness checks where the left column shows the famine effect of at least two years of famine during SGP and the right column shows the SGP famine effect with the extended famine period. We expect coefficients to increase if only the more highly affected individuals are included in the analysis. Contrary to this, the maternal mortality effects for males become smaller and less significant. This is also the case for the schooling effect. When the famine effect is reduced to the most highly affected individuals, the effect of paternal SPG famine exposure on male schooling becomes insignificant at the 10% level. However, the positive effect of paternal SGP famine exposure on son's height and the negative effect of maternal SGP famine exposure remain. As expected, the coefficients become larger in absolute terms and more significant, indicating that if the father (mother) was affected by the famine during his SGP, this leads to an increase (a reduction) in son's height by 0.8cm (0.9cm). Note that the positive effect of paternal SGP famine exposure on male height also persists if we include the year 1919 in the famine period.

As expected, for G3 individuals, point estimates for mental health increase if we only include the most highly affected individuals. If the paternal grandfather is affected by the famine during his SGP, grandsons' mental health tends to be 28% of a standard deviation higher and if the maternal grandmother is affected, the increase in mental health for granddaughters is more than 40% of a standard deviation. Effects also remain large and significant when the famine period is extended to the year 1919.

4.6.3 Discussion

In the present study, we find patterns in the data that hint towards the intergenerational transmission of famine effects triggered by a reduction in food during the ancestors slow growth period. In this section we try to answer: [1] whether the results we find are indeed likely to be triggered by the famine. [2] Whether they can be related back to epigenetic inheritance and [3] which other explanations are possible, behavioral or otherwise.

First, to assess whether our findings are indeed related to a food shortage in SGP, we have excluded all individuals from the analysis, who have only been affected by the famine for a single year during their SGP. Any causal impact of the famine during SGP, should be stronger for individuals who suffered from the famine for a longer period of time. We find that some of the effects, such as the positive effect of maternal SGP exposure on sons' mortality risk, or the negative effect of paternal SGP exposure on sons' probability

to obtain a higher secondary school degree do not persist if we exclude mildly affected individuals. Other effects, such as the positive (negative) effect of paternal (maternal) SGP exposure on sons height or the third generation sex specific effects become stronger and more significant. We interpret this as evidence that the latter are more likely related to SGP famine effects than the former.

Concerning the channels through which these intergenerational famine effects operate, we cannot be certain that our estimated effects are of epigenetic origin. Yet, in line with some of the biological literature (Pembrey, 2002; Pembrey et al., 2006), we would expect an epigenetic transmission of SGP famine effects to take place along the male line, because during this age period the sperm is still in development, while the female ovum is fully developed upon birth. Furthermore, if epigenetics was the cause of our findings, we would expect strong effects on the second generation and a fading out of effects in the third generation. Last, we would expect SGP famine exposure to be positively associated with height and longevity. Our results only confirm few of these hypotheses. First, we do not find much stronger effects of paternal and grandpaternal SGP famine exposure than of maternal one. Furthermore, none of our results are strong among the second generation and significant but lower in magnitude for the third generation. Thus paternal SGP famine exposure has a positive effect on sons height among the second generation, but the coefficient of paternal grandfather SGP famine exposure is negative and insignificant. Similarly, paternal grandfather SGP famine exposure is positively associated with third generation mental health among males, but paternal SGP famine exposure does not have a significant impact on males among the second generation. However, we find that males are taller if their fathers have been affected by a famine during their SGP, a finding that may indicate positive effects of paternal SGP exposure on male mortality. A reduction in mortality is in line with the findings of Kaati et al. (2002), who find that paternal SGP famine exposure reduces cardiovascular mortality, but somewhat contrary to the findings in Kaati et al. (2007), where the authors find that a surfeit in parental food supply during the SGP leads to an increase in longevity.

Epigenetic imprinting is only one possible channel through which intergenerational transmission of famine exposure may operate. Other explanations for potential findings may be that SGP exposure affects height, fertility or cognitive and noncognitive skills of the first generation, which then influence later generation outcomes. For example the

experience economic hardship around the age of 10 can lead to more resilience and psychological strength (Elder, 1999).[20]. If our results were driven by these mediating factors, we would expect the famine to have a positive impact on first generation schooling probabilities, family size, age at birth or socioeconomic status, which would then affect later generations indirectly. Yet, we find that controlling these variables, does significantly improve model fit, but does not have an impact on the estimated famine coefficients. We thus think it is unlikely that a behavioral impact on the first generation is driving our results.

4.7 Conclusion

This chapter investigates how a reduction in food supply due to the German World War 1 famine of 1916-1918 affects descendants of the second and third generation. A highly debated line of literature in biology has found that low food availability during the slow growth period of male individuals positively affects health outcomes of subsequent generations. That literature argues that such effects are potentially triggered by methylation of epigenetic marks in the sperm, with methyl tags being transmitted to subsequent generations via epigenetic imprinting.

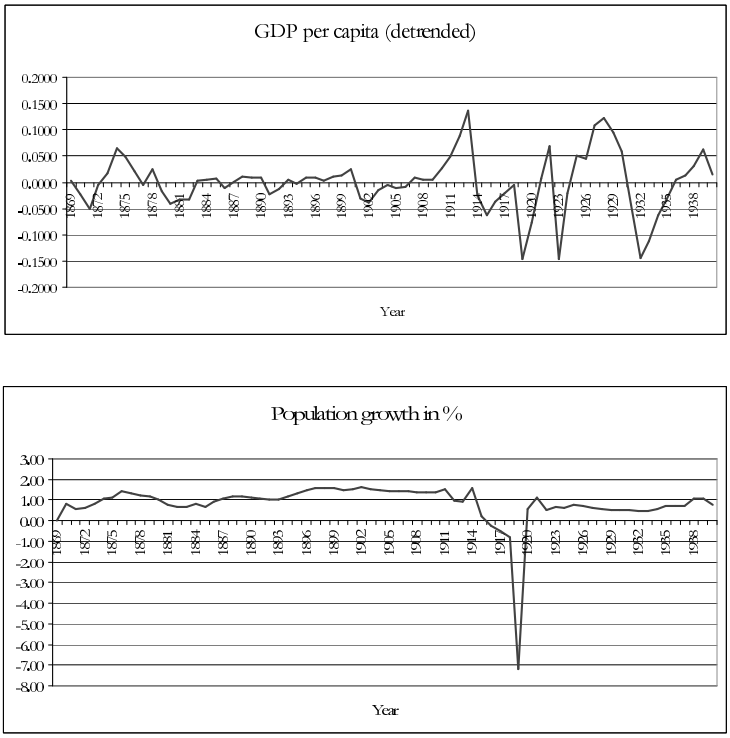
We find that males are indeed taller if their fathers have been affected by a famine during their SGP and shorter if their mothers have been famine exposed. Furthermore, paternal grandfather SGP-famine exposure is associate with higher mental health of third generation sons, while maternal grandmother SGP-famine exposure has a positive effect on her granddaughters mental health. Besides, our findings suggest that among second-generation individuals maternal SGP-famine exposure is associated with lower height. Both, second generation height and third generation mental health effects are likely to be causally related to the famine, as they persist and become stronger if only the most severely affected individuals are made part of the analysis. However, we do not find the same strong male line effects on second generation mortality and third generation height (as a proxy of mortality) as in (Bygren et al., 2001; Kaati et al., 2002; Pembrey et al., 2006).

[20]Effects along the female line could be driven by maternal nutrition during SGP. Recently, it has been shown that nutrition during pregnancy matters less for birth weight than nutritional status at the time of conception or even before (Kuzawa, 2008).

Research on intergenerational famine effects and epigenetic inheritance is only starting. Further research on epigenetic markers for famine affected individuals is badly needed and similar analyses to this one, but on different datasets would be highly valuable to confirm and extend our understanding on the intergenerational transmission of nutritious effects.

4.A Additional Figures and Tables

Figure 4.2: GDP per capita and population size, German Empire



Notes: GDP per capita (logarithm) is measured in 1990 International Geary-Khamis dollars. Population growth measures the yearly population growth in percent. GDP was detrended using the Hodrick-Prescott-Filter (λ=100).
Source: A. Maddison (2006). The World Economy: Historical Statistics, OECD Development Centre, Paris 2006.

Table 4.12: G2, differences in control variables

Variables	Males			Females		
	Father(Z=1)	**Father(Z=0)**	**Pval**	**Father(Z=1)**	**Father(Z=0)**	**Pval**
Father's age at death	65.36	66.07	0.32	65.02	65.42	0.61
Mother's age at death	73.34	73.56	0.75	73.06	73.56	0.49
Father died before age 13	0.17	0.16	0.64	0.18	0.18	0.84
Mother died before age 13	0.04	0.05	0.90	0.04	0.05	0.21
Father upper sec school	0.10	0.08	0.18	0.09	0.07	0.32
Mother upper sec school	0.03	0.03	1.00	0.03	0.03	0.78
Father intermediate school	0.08	0.07	0.51	0.09	0.08	0.52
Mother intermediate school	0.11	0.10	0.26	0.10	0.09	0.36
Number of brothers	0.90	0.97	0.16	0.94	1.00	0.19
Number of sisters	0.96	1.01	0.33	0.91	0.99	0.08
Father's age at birth	32.68	33.23	0.07	32.68	32.88	0.51
Mothers's age at birth	29.11	29.47	0.15	29.11	29.49	0.14
Year of birth	1941.08	1939.81	0.00	1941.11	1939.33	0.00
N	2103			1977		

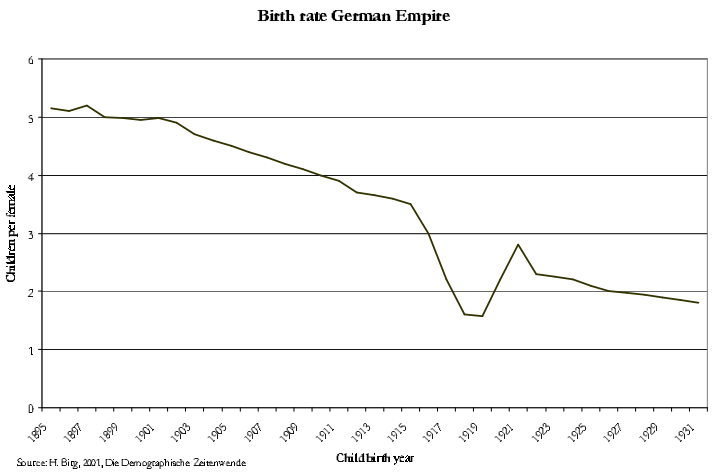
Source: SOEP, waves 1982-2009. Own calculations.
Notes: p-values of a two-sided t-test for differences in means are reported.
'Age at death' is conditional on having died. Standard deviation in brackets.

Table 4.13: G3, differences in control variables

Variables	Males			Females		
	PGF(Z=1)	**PGF(Z=0)**	**Pval**	**PGF(Z=1)**	**PGF(Z=0)**	**Pval**
Father upper secondary school	0.27	0.26	0.81	0.29	0.23	0.09
Mother upper secondary school	0.23	0.16	0.05	0.21	0.18	0.50
Father intermediate school	0.15	0.16	0.90	0.13	0.14	0.60
Mother intermediate school	0.19	0.25	0.07	0.26	0.16	0.00
Number of brothers	0.66	0.50	0.02	0.55	0.70	0.02
Number of sisters	0.62	0.62	0.93	0.63	0.67	0.63
Father's age at birth	31.25	32.24	0.05	31.20	31.46	0.57
Mother's age at birth	28.66	28.37	0.56	28.33	28.26	0.88
Year of birth	1974.04	1973.48	0.38	1972.75	1971.98	0.23
N	625			811		

Source: SOEP, waves 1982-2009. Own calculations.
Notes: p-values of a two-sided t-test for differences in means are reported.
'Age at death' is conditional on having died. Standard deviation in brackets.

Figure 4.3: Birth rate, German Empire



Source: H. Birg, 2001, Die Demographische Zeitenwende.

Chapter 5

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

5.1 Introduction

During the past decade, the effect of early-life conditions on health outcomes late in life has become a focal point of research in economics, demography, epidemiology, biology, and related fields.[2] Empirical studies use non-experimental data of real-life individuals and typically relate indicators of early-life conditions to outcomes of the same individuals at high ages. A major concern is that (a) observed conditions in the parents' household early in life, including conditions in utero as captured by birthweight, and (b) 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, candidate 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

[1] This chapter is joint work with Gerard van den Berg and Johannes Schoch. An earlier version of this chapter is published as IZA working paper number 6110.

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

a famine involves nutritional shortages, in terms of quality and/or quantity of nutrition. Birth outside of a famine is expected to involve better nutritional conditions. A negative association between birth during a famine and a favorable health outcome later in life is interpreted as evidence for a causal effect of early-life conditions on that health outcome.[3] Studies have detected associations with adult height, schizophrenia, mortality, fertility, hospitalization, next generation birthweight and even gene expression. Interestingly, a few studies have also reported associations between exposure to a famine at the onset of puberty and late-life health outcomes. For example, Sparén et al. (2004) find that boys who had been exposed around age 9 to the Leningrad siege famine of 1941 have higher rates of cardiovascular morbidity and mortality much later in life. Oppers (1963) finds an effect of the Dutch Hunger Winter famine exposure at ages 7-14 on adult height.[4]

However, exposure to a famine is not equivalent to exposure to a nutritional shortage. During a famine, a fraction of all households does not face food shortages, for example because the household belongs to the ruling or wealthy class, or because it is self-sufficient in terms of food, or, in a war context, because it is allied to those responsible for the cause of the famine. Similarly, in non-famine eras, a fraction of households faces food shortages because of poverty. This means that the comparison of famine-born individuals to non-famine-born individuals does not provide a quantitative 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. But this is almost impossible given that the study outcomes concern health at high ages. To observe these 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. Moreover,

[3]Note that also this requires that the composition of newborns is not systematically different between famine-born and non-famine born, in terms of unobserved characteristics of the newborns.

[4]Many qualitative results from the famine-based studies are in agreement to those in studies using other contextual indicators of conditions around birth, such as business cycles and seasons. Van den Berg, G.J. and Lindeboom, M. (see e.g. 2012) for an overview. Van den Berg, G.J. and Lundborg, P. and Nystedt, P. and Rooth, D.O. (2010) use immigrant siblings to study the effect of economic conditions at later childhood ages, and they find a significant effect on adult height. They also survey other studies concerning later childhood ages.

for some of the sample, the critical period must have occurred during a famine. During famines, data are typically not collected, as societies are in a state of disruption.

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

Our 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. Notice that IV is generally not applicable in studies of long-run effects of early-life conditions, because of non-observability of household conditions early in life. This 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 a famine and an actual hunger episode. In terms of the IV treatment evaluation literature, we look at "compliance" to the instrument.

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 most recent available (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.[5] Their origins are well-established in the literature, and they have been used to study long-run effects (see the literature discussion in Section 5.2). Notice

[5]Barring famines that affected only a few respondents, like the German 1916/18 famine; see Van den Berg and Pinger (2011). Our analysis restricts attention to births in 1920-1955.

that evidence based on multiple famines is 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 early in life. Even with a perfect recollection of past periods of hunger (e.g. if the individual obtained this information from his 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 more 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 estimating the probabilities in the denominator of the nonparametric Wald estimator from a sample of older children. Intuitively, when we consider long-run effects of nutritional shortages for newborns, we relate famines around birth to health later in life, but we may use a sample of older children to estimate the connection between famine exposure and nutritional shortage. This requires the assumption that the latter connection is the same for all children. This is not innocuous. As we shall see, there is 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 since we only use men, this number is halved. However, 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 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. In fact, given the relatively small sample size, moderate changes in the definitions of these intervals only affect the status of few respondents, and the results are often 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. Child hunger is not only prevalent in many parts of the developing world, but also in industrialized countries. "Feeding America" reports that even in the US 11.9 million citizens regularly suffer from hunger. From these, around a third are children under the age of 18.[6] Undernutrition is essentially an economic problem that can be mitigated by public policy. From an economic point of view, it matters to find out at which age of the children exposure to nutritional shortages is most detrimental in the long run. This helps to address which policy measures are most efficient and cost-effective in preventing adult health problems.

The chapter proceeds as follows. In Section 5.2 we review the explanatory frameworks to understand the long-run effects. Section 5.3 describes the three famines in our observation window and summarizes the evidence obtained so far for those famines. In Section 5.4 we describe our data. Section 5.5 formally presents the econometric methods. Here we also examine selectivity issues associated with the famines. Section 5.6 presents our results. We do not only study effects of adverse conditions around birth but we also identify whether periods in early adolescence are critical with respect to nutritional shocks. We also consider the strength of the association between famines and actual nutritional shortages. This is of importance for the empirical literature in which the three famines have been used as indicators of early-life conditions. In Section 5.6 we also carry out placebo estimations using cohorts from countries that were not affected by famines. Moreover, we use cohorts

[6]The data come from the United States Department of Agriculture which characterizes households in which one or more people were hungry at times during the year as households with "very low food security" or as "food insecure with hunger" because the household could not afford enough food. "Hunger" in that description referred to "the uneasy or painful sensation caused by lack of food."

from other countries (Belgium, France and Italy) as control cohorts, as an additional way to verify that the effects of hunger do not reflect the effects due to exposure to World War II and its aftermath. Section 4.7 concludes.

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

5.2.1 Conditions around birth

Most explanations for long-run effects of nutritional conditions around birth build on Barker's fetal origins or fetal programming hypothesis (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, 2004a). Underlying this model is the idea that several critical periods early in life influence the development of humans. During these periods, developing systems modify their settings in response to social and biological cues (Kuzawa and Quinn, 2009). This includes durable epigenetic changes that modify gene expressions.

Along this way, adverse conditions are known to influence inflammation, measured in terms of interleukin-6 production, in adolescence and adulthood, plausibly through changes in gene expressions (see e.g. Morozink et al., 2010). An episode of hunger early in life may thus engender a proinflammatory phenotype. Over time, this takes an allostatic toll on the body, resulting in a higher risk of chronic diseases later in life (Morozink et al., 2010; Miller and Chen, 2010; Miller et al., 2009; Zhang et al., 2006; Cole et al., 2011), notably cardiovascular diseases, diabetes and hypertension. Yet, such mechanisms should be seen as a predictive adaptive response to the future environment (Gluckman and Hanson, 2004a; Cole et al., 2011). The long-run effects of reduced nutrition in utero are stronger if the affected individuals are exposed to a much more favorable environment in childhood (Schulz, 2010).

Of course, severely adverse nutritional conditions may also directly affect the build-up of organs and other body parts. This involves the postneonatal period but also puberty (see e.g. Cameron and Demerath, 2002).

The above causal pathways are all biological after the initial nutritional shortage. However, non-biological mechanisms are also possible. An episode of hunger may lead

to a permanent change in the role of food in the household. After such an episode, adults may cook more greasy food or force their children to finish their plates by any means (See e.g. Hamelin et al. (1999), for changes in household behaviors in response to food insecurity). This may cause adverse long-run health outcomes of the children. Also, an episode of hunger may induce fights for resources and thus a higher level of stress, changing within-household relationships and leading to a prolonged exposure to a higher stress level (Hadley and Patil, 2006; Whitaker et al., 2006).

5.2.2 Conditions in later childhood

Recently, interest has increased in long-run effects of nutritional conditions after birth. 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 Marshall and Tanner (1968), 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. Other studies have related the calcium intake around age 9 to adult height.

A nutritional shortage 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). Last, adverse conditions may affect children positively by inducing more responsible behavior. For example, Elder (1999) investigates the impact of the Great Depression on children born in 1920-1921 and finds that the experience of economic hardship around the age of 10 led to more resilience and psychological strength.

5.3 The Famines in European Countries in the 1940s

5.3.1 The Dutch famine

The Dutch famine has been studied for decades as a cause of adverse living conditions. Therefore, the following account can be brief. The relevant literature starts with contemporaneous studies, notably Dols and van Arcken (1946), who provide a detailed

description of the famine and report data on rations and agricultural production, and Banning (1946), who focuses on public health issues.

Prior to World War II, food standards had been high in the Netherlands, both in terms of caloric value as well as composition of the diet. There were no notable disruptions in food availability during the first years of the occupation of the Netherlands, which started in May, 1940. In September 1944, parts of the South of the country were liberated, and the London-based Dutch Government in exile called out a railroad strike in the occupied parts of the Netherlands in order to support Operation Market Garden and in order to display its authority over the occupied nation. As a reaction, the occupying forces initiated an embargo that prohibited any food transports to the densely populated western part of the country, i.e. the provinces of North and South Holland and Utrecht. This sanction, 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 effectively closed off the western part of the country from any imports of food, fuel, medication etc. This triggered the Dutch "hungerwinter". 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 percent for men and 173 percent for women (Dols and van Arcken, 1946). Banning (1946) reports a higher incidence of tuberculosis and hunger oedema and an increased infant mortality rate. Inhabitants of large cities were struck hardest by the famine. However, 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 additional food rations but in addition, effort was taken to send adversely affected children to districts where the food situation was somewhat

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

better. Reports on the activities of the council yield information about the situation of the children in the famine-struck areas: 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 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). Susser and Stein (1994) find that adult stature is susceptible to the postnatal but not the prenatal environment. 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 Section 5.2.1 above.

5.3.2 The Greek famine

At the end of April 1941, Greece surrendered to Axis forces and was subsequently divided into 13 different zones occupied by Germany, Italy, and Bulgaria.[9] These areas were isolated from each other and the transfer of goods and individuals was often close to impossible. The famine was triggered by a naval blockade of the Allies which made it impossible to supply foodstuffs to Greece. Despite being an agricultural country, Greece heavily relied on food imports (Hionidou (2006), mentions that in 1939, over 20 percent of the wheat consumption had been imported). In addition, the blockade prohibited fishing

[8] Recall that in this section we only cite studies based on the famines we consider in our empirical analysis. Results based on other famines are discussed in the overview studies listed in Section 5.2.1.

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

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 amounted to only 320 kcal in November 1941 Neelsen and Stratmann (2011). The regions of Greater Athens and Piraeus were affected most severely. Mortality rates suggest that there was a general positive correlation between the famine's severity and the degree of urbanization. Furthermore, Hionidou (2006) notes that society was very unequally affected by the famine during this period; the low social classes suffered the most from deprivation, which can be attributed to the fact that soon after the rationing system had been superimposed by the occupiers, black market activity flourished with prices beyond the levels an ordinary worker could ever afford. The winter 1941/42 marked the maximum of the famine period in terms of fatalities.

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 (1946b) reports death rates for Athens to have been six times higher than under usual circumstances. Both Valaoras (1946b) and Hionidou (2006) state that the largest fraction of the excess mortality during the famine is attributable to starvation, whereas epidemics and infectious diseases are of minor importance. Of special interest for our present study are the figures cited in Valaoras (1946b) 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.

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

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

5.3.3 The German famine

After World War II, Germany faced various structural problems that led to the catastrophic situation we label the German Famine. Specifically, (i) many cities and much of the industry and transportation infrastructure were destroyed by Allied bombardments, (ii) millions of citizens and military had died, (iii) factories were dismantled and shipped off to the occupying powers, (iv) the agricultural land in the East which had been a major provider of food had been lost to Poland and the Soviet Union, (v) about 10 million refugees from the lost lands in the East had fled to the new German mainland, and (vi) Germany was divided into 4 occupation zones that were administered separately. Inter-zone trade was difficult because of political and bureaucratic barriers and because of a lack of transportation means. See Klatt (1950), Farquharson (1985), Trittel (1990), Häusser and Maugg (1985) and Reichardt and Zierenberg (2008) for overviews.

Before World War II, estimated daily caloric consumption amounted to about 3000 kcal, while during the war, civilians used foodstuffs of about 2500 kcal per day (Klatt, 1950). At the end of the war (European Theater) in May 1945, nutritional conditions deteriorated to levels far below those before or during the war. For May 1945, Trittel (1990) reports an average caloric ration of only 1200 kcal. After May 1945, these values decreased further. For example, in the British-administered Ruhrgebiet, only approximately 600 kcal per person was assigned in June 1945. The situation further deteriorated during the so-called "Hungerwinter" of 1946-47, which was an unusually cold winter. From mid-December 1946, the distribution of food collapsed in many areas. Until March 1947, real average daily caloric intake was around 800 kcal per day for some cities in the Ruhrgebiet. The crop of 1947 failed to meet the demand of the starving population, so the period of hunger lasted until the spring of 1948, when a combination of foreign help and political reforms managed to improve nutritional conditions. Conditions improved spectacularly with the currency reform in June 1948. We therefore define the famine to last from June 1945 until June 1948.

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

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

Jürges (2012) analyzes the effect of birth during the German famine on educational attainment and occupation in the labor market in 1970. He finds strong evidence that these outcomes are particularly adverse for those born in the winter of 1945/46. This suggests that an indirect causal pathway from early-life conditions through education and occupation to health later in life may be possible. Some epidemiological studies consider cohort-specific adult health outcomes without focusing a priori on famine cohorts. Notably, Hermanussen et al. (2001) use data on male conscripts; they observe that the 1946 cohort stands out in terms of low obesity at age 19 even though the average weight is not much lower than for the surrounding cohorts. 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 Section 5.2.1). In this sense, individuals born in World War II may face less adverse long-run effects than those born in the 1946/47 "Hungerwinter". Onland-Moret et al. (2005) use large samples of women from a large range of European countries. They report the average height by country and by 5-year birth cohorts. The average height among women born in Germany in 1945-49 is much lower than that in the surrounding 5-year cohorts. For Greece and the Netherlands no

such patterns are found, which is not surprising in the light of the fact that the famine spells in those countries spanned only a small part of a 5-year interval.

5.4 Data

For the empirical analysis we use the Survey of Health, Aging, and Retirement in Europe (SHARE), a pan-European multidisciplinary and longitudinal data set on individuals aged 50 and over. By 2011, three waves of SHARE have been made 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] To date, this is the most recently available wave. Additional waves are under construction.

Sample construction We use 3100 men 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 in 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] Our working sample consists of about 2700 men.

We restrict attention to male individuals. This is because the literature on the effects of early-life conditions on CV morbidity and mortality has found the strongest effects on men (see the surveys: Poulter et al. (1999), Rasmussen (2001), Lawlor et al. (2004), and Huxley et al. (2007)). For the Finnish famine of 1866-1868, Doblhammer et al. (2013) find that the long-term effects of early life circumstances differ by gender, and that the results for women are less conclusive than for men. Also, the effect of the business cycle

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

[12] We drop the full proxy interviews but keep assisted interview information. Moreover, we employ this criterion only for the interviews yielding the outcome measures and the undernutrition indicator.

[13] Conversely, there are men for whom we observe height but not the hunger variable since they participated in the first two waves of the survey but not in SHARELIFE. Sample attrition is high in SHARE. We treat these observations as missing at random, keeping in mind that even if attrition was somehow related to health performance, this would make our estimates more conservative.

at birth on late life mortality is stronger for men than for women in the Danish twin data (Van den Berg et al., 2009), and in Dutch data (Van den Berg et al., 2006). This is also true for the effect of the Dutch Potato Famine on late-life mortality (Lindeboom et al., 2010).

A medical study on the relation between birth weight and mortality from heart disease finds stronger effects for men (Leon et al., 1998). 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). Pitkaïnen (1993) found large sex differences in the Finnish famine mortality figures, with males being far more likely than females to die. Although this sex differential is particularly large for ages 10 to 40 in the high-impact areas of the famine, it also exists for younger ages.[14]

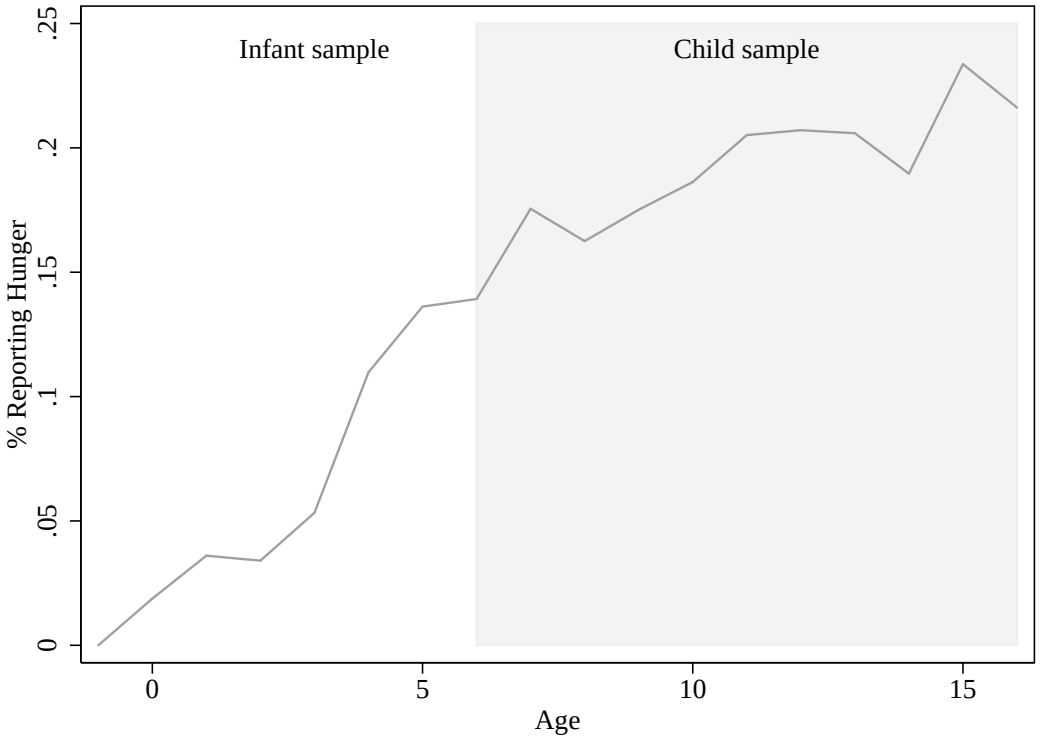
One exceptional feature of SHARE is that respondents were asked whether they had ever experienced periods of hunger in their lives and if so, in which year of their life this period began and in which year it ended. The survey question reads "Looking back at your life, was there a distinct period during 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, since it is self-reported and dates back many years.[15] Moreover, recall bias is likely to by systematically higher 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.

To deal with systematic, age-related recall bias, we distinguish between two samples of different ages when we define the undernutrition indicator: an *"infant sample"* and a *"child sample"*. The infant sample is composed of individuals who potentially experienced hunger/famine 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 from ages 6-16. Hunger is defined as a binary indicator (or treatment variable) which equals one if a person has

[14]It has been hypothesized (see e.g. Low, 2001) that the smaller impact for females may be explained by a so-called male vulnerability because males are the heterogametic sex: they have an unprotected Y chromosome and, therefore, they may be more vulnerable to adverse environmental conditions in the short as well as in the long term. See also Eriksson et al. (2010) for a biological explanation of these findings.

[15]Also, only one period of hunger can be reported, such that individuals are likely to choose the period where nutrition was shortest in supply.

Figure 5.1: Probability to report hunger conditional on famine experience at respective age

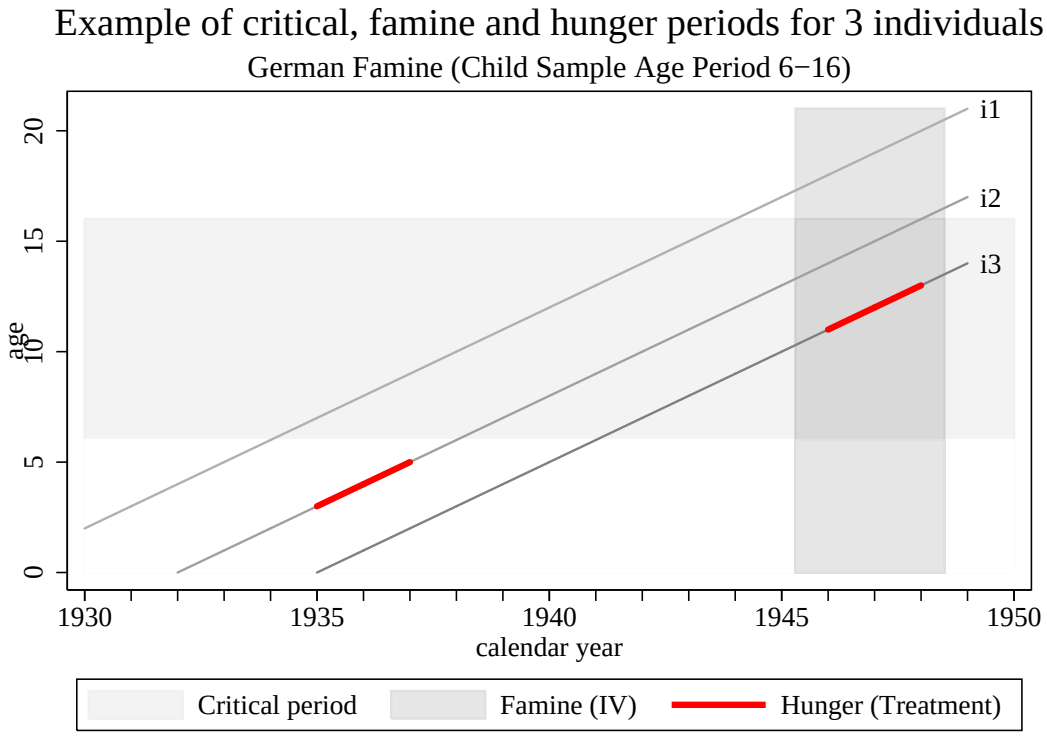


experienced a period of undernutrition during age interval $[age_{-1}, age_4)$ for the infant sample and during $[age_6, age_{16})$ for the child sample. A respondent is considered to have suffered undernutrition if he reported that he experienced an episode of severe hunger and if this period either started before age_{start} and ended thereafter, or if the period started within the specified interval. The intervals are defined with an eye on the small sample sizes. In sensitivity analyses we assess the robustness of the results with respect to these definitions.

Graph Figure 5.1 provides evidence on the extent of recall among the infant and child samples. The graph displays the probability of reporting hunger only for those individuals who lived during a famine period. Individuals in the infant sample often do not report to have lived in a period of hunger, while individuals in the child sample report to have suffered from undernutrition at a stable rate of just under 20%. If an individual has lived during a famine, his probability to report hunger increases if the famine happened until age 6, and it remains stable for increasing ages at famine.

Famine instrument The famine periods are defined in Section 5.3,

Figure 5.2: Example treatment definition



1. Greece: May 1941 to end of June 1942

2. The Netherlands: November 1944 to end of April 1945, only the Western part[16]

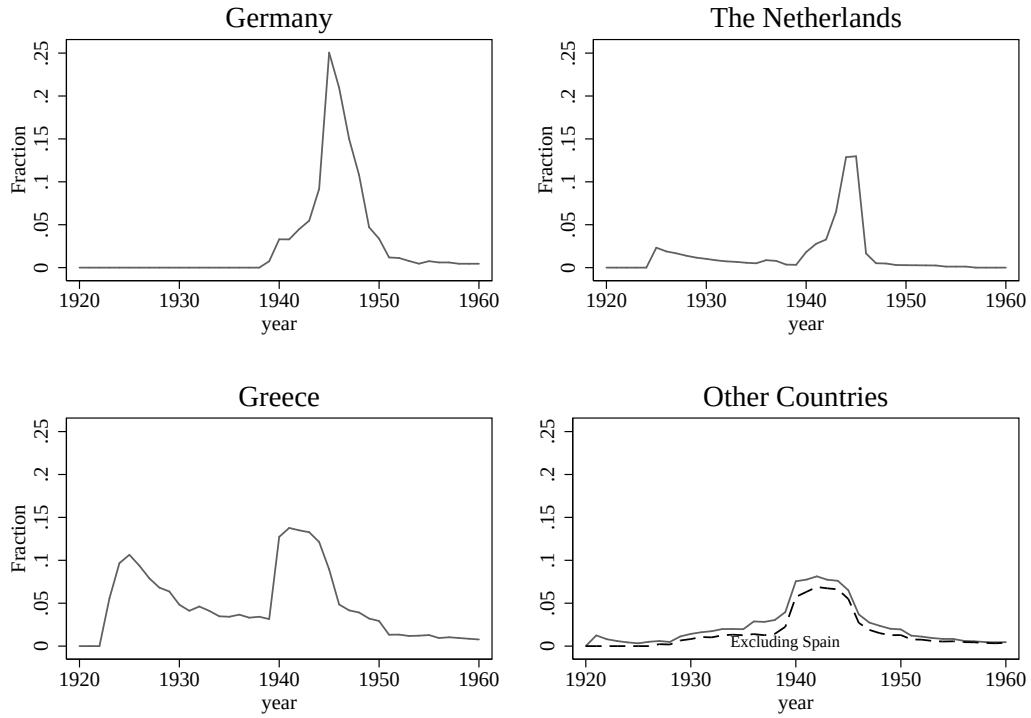
3. Germany: June 1945 to end of June 1948

We construct a binary instrumental variable that takes the value one if a famine affected the individual in utero/at ages 0-4 or 6-16, respectively. This uses the information on individual's month of birth. Notice that changing the start or end months of the famine affects our results only very little, since 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.

Figure 5.2 provides a graphical illustration of an example of our hunger (treatment) and instrumental variable definition for the child sample of German individuals. In this

[16]To select the Dutch sample into those exposed to a famine or not, we use information on the accommodation a respondent lived in the year the famine started, i.e. 1944; if a respondent moved house in the very same year, we use the previous accommodation in order to rule out endogeneity due to internal migration. The SHARELIFE variable distinguishes five possible regions for the Netherlands: Noord-Nederland, Zuid-Nederland, Oost-Nederland, West-Nederland, and Midden-Nederland. The famine instrument is set to one only for West-Nederland.

Figure 5.3: Probability for episode of hunger by calendar year (given observations of sample already alive)



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.

Figure 5.3 and Figure 5.4 provide graphical assessments of the hunger periods we find in the data. Figure 5.3 shows 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 toward the end of the war, which marks the beginning of the period we define as the famine; even though the fraction of those reporting hunger for this period declines somewhat after this peak, it stays high until the end of the famine. For the Netherlands, we can distinguish one single peak for the famine period while propensities for hunger amount to basically zero before and afterwards. The fraction of observations reporting hunger during the famine years is lower for the Netherlands, because only the Western

Table 5.1: Reaction of hunger at age 0 – 16 to famine exposure at the same time (cohorts 1920 – 1955)

	Greece		Germany		The Netherlands		Sum	
	treated	non-treated	treated	non-treated	treated	non-treated	treated	non-treated
Famine-exposed	66	406	73	334	30	79	169	819
Not Famine-exposed	22	670	11	246	20	719	53	1635

part of the Netherlands was affected by the famine. For Greece we also find a spike around the famine period but the level remains relatively high for the 1920s (given the small number of individuals born in the early cohorts, it is likely that a few outliers drive these high rates). The impact of the famines becomes even more obvious when taking a look at the other SHARE countries: Here, the overall fraction of those reporting episodes of hunger is comparably small and increasing only somewhat during the second World War. 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 graph implies 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.[17] This provides a rationale to exclude the other countries from the baseline analyses. We return to this in Section 5.6 when we discuss sensitivity analyses.

Figure 5.4 shows the average duration of a hunger period for hunger periods starting in different years. The graph shows that hunger periods get shorter when coinciding with a famine. This supports our presumption that some individuals experience short but intense periods of severe undernutrition because of the famine. Table 5.1 shows that a fraction of approximately 17 percent of those having experienced a famine before age 16, report hunger for the same period. For non-famine periods, this is only 4.6 percent. Hence, famine is a powerful instrument for periods of undernutrition, although the overall propensity to report hunger remains rather low.

Adult height As an outcome and proxy for late life adult health we use adult height, measured in centimeters. Height is frequently used in the literature as a marker of early

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

Figure 5.4: Average length of hunger periods by calendar year

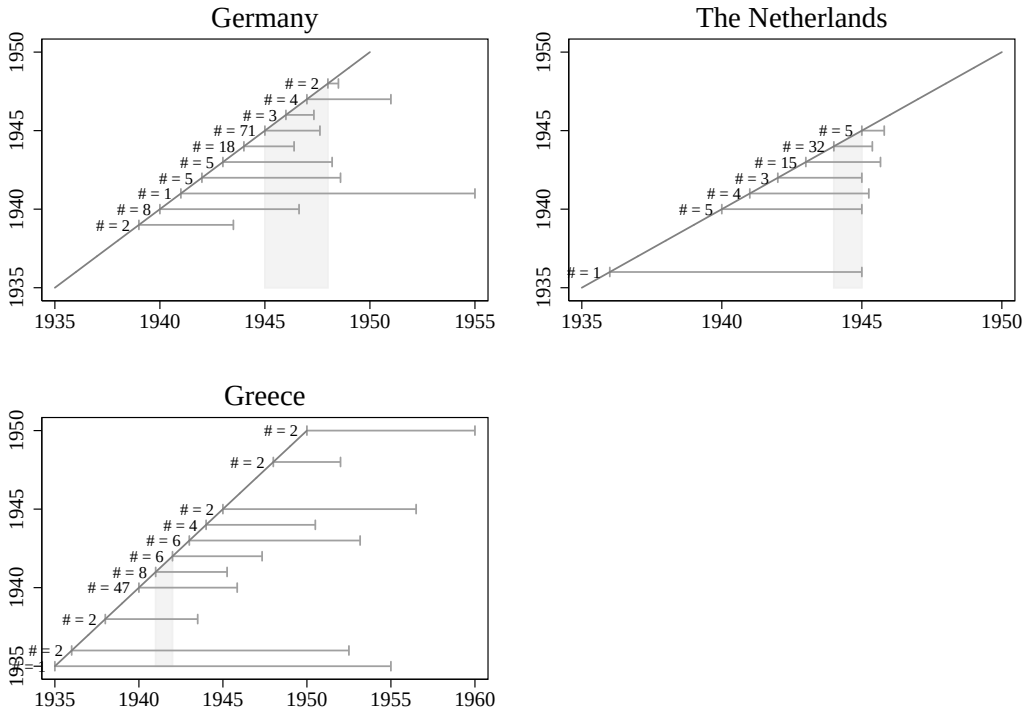


Table 5.2: Descriptive statistics for famine countries (cohorts 1920 to 1955)

	All (N= 2676)		Control (N=2454)		Treated before age 16 (N=222)		Treated before age 4 (N=54)	
Variable	mean	(s.d.)	mean	(s.d.)	mean	(s.d.)	mean	(s.d.)
Year of Birth	1942	(8.501)	1942	(8.490)	1936	(5.609)	1938	(6.541)
Episode of Hunger being 0-16 (1 = yes)	0.083	(0.276)	0.000	(0.000)	1.000	(0.000)	1.000	(0.000)
Experienced famine being 0-16 (1 = yes)	0.369	(0.483)	0.334	(0.472)	0.761	(0.427)	0.722	(0.452)
Born in rural area (1 = yes)	0.409	(0.492)	0.414	(0.493)	0.353	(0.479)	0.340	(0.478)
Adult Height in cm	175.547	(7.248)	175.669	(7.235)	174.203	(7.269)	173.426	(7.895)

life health and has the advantage of being constant during adult life (Costa, 1993).[18] It is universally 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, there exists evidence that height it is almost linearly related to life expectancy (Steckel, 2008). Height is measured in the second wave of the survey. Table 5.2 compares relative magnitudes of adult height between those who experienced hunger before age 16 and those who did not.

[18]The second wave of SHARE contains information on a large number of other health conditions and chronic diseases. However, many of these conditions, such as hypertension or diabetes only manifest relatively late in life. Hence, the use of these outcomes generates problems of cohort and age effects.

Covariates In the analysis, we control for country and gender and for a linear time trend.[19] Besides, we include controls for the degree of urbanization of the place of birth or of the place where an individual lived at the time of our reference age.[20] We consider this information as a proxy for socio-economic background. Table 5.2 provides basic summary statistics for the overall sample, for the subsample of respondents who reported an incidence of hunger sometime before age 16, and for the subgroup which never reported an incidence of hunger.

5.5 Empirical Strategy

5.5.1 Instrumental variable methods

Our model framework is as follows,

$$Y_i = \psi(D_i, Z_i, X_i, \epsilon_i) \tag{5.1}$$

$$D_i = \phi(Z_i, X_i, \varepsilon_i), \tag{5.2}$$

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

If effects of nutrition on adult height is 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 the presence of a famine. The size of the complier group and the estimated effect may depend on the severity of the famine (Angrist et al., 1996). In this respect, an advantage of our approach is that the famines are regarded to

[19]The SHARE data, especially the third wave, provides a great deal of background information. However, the 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.

[20]If an individual changed the accommodation in the year he turned six, information on the accommodation inhabited the year before was used in order to prevent bias from selective internal migration.

be severe. Formally, the LATE is defined as:

$$\begin{aligned} LATE &= \mathrm{E}[Y_{D=1} - Y_{D=0} | D_{Z=1} > D_{Z=0}] \\ &= \int (\psi(1, Z, X, \epsilon) - \psi(0, Z, X, \epsilon))\, dF_{X,Z,\epsilon|D_{Z=1}>D_{Z=0}} \end{aligned} \tag{5.3}$$

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 that allows for conditioning on covariates,

$$LATE = \frac{\int \mathrm{E}[Y|X = x, Z = 1] - \mathrm{E}[Y|X = x, Z = 0]\, f(x)dx}{\int \mathrm{E}[D|X = x, Z = 1] - \mathrm{E}[D|X = x, Z = 0]\, f(x)dx}. \tag{5.4}$$

Equation (Equation (5.4)) 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.

5.5.2 Two-sample estimation

As explained above, imperfect recall in the infant sample may cause the incidence of severe hunger for that group to be misreported. We tackle this problem by replacing the estimate of $\mathrm{E}[D|X = x, Z = 1] - \mathrm{E}[D|X = x, Z = 0]$ in the denominator of the nonparametric Wald estimator (Equation (5.4)) by the estimate of $[\mathrm{E}[D|X = x, Z = 1] - \mathrm{E}[D|X = x, Z = 0]]_{child}$ from the child sample but integrating over the marginal distribution of covariates in the

infant sample.[21] Note that this assumes that the true distributional effect of a famine on the probability of malnutrition is the same among infants as among older children.[22] This assumption 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 e.g. Dols and van Arcken, 1946; De Rooij et al., 2010; Klatt, 1950; Valaoras, 1946a; 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 (Hutchinson et al., 1951). At the same time, as we have seen in Section 5.3, special food aid programs were available for children in school-going ages. However, this was a response to the severity of nutritional shortages among those children, instead of an indication that society favored food allocation to those children over the allocation to infants.

More in general, parents may redistribute food supply across children with different ages within the household. For parents it may be more efficient to allocate resources towards stronger and older children, as these are more likely to survive a famine. Conversely, they may prefer to invest more into their younger children if they favor more equal outcomes, or if they prefer all of their children to survive even if this comes at a cost for the older children's development. Whether decisions are driven by equity or efficiency concerns depends on the number of children, their probability of survival, parental preferences, and the parental budget constraint (Becker and Tomes, 1976; Behrman, 1997; Behrman et al., 1982). In developing countries, parents' child investment decisions are generally driven by efficiency concerns (see e.g. Rosenzweig and K., 1988). For the developed economies, most research indicates that parents follow equity considerations (Griliches, 1979; Del Bono et al., 2008), although Datar et al. (2010) point in the opposite direction. The famine-stricken countries in our data would perhaps classify in-between current developed and developing countries, although parents living through the famines

[21] We bootstrap standard errors using 500 bootstrap iterations.

[22] 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. We use the two-sample two-stage least squares estimator (2S2SLS) proposed by Inoue and Solon (2010), which adjusts the original two sample IV (2SIV) estimator developed by Arellano and Meghir (1992) for use in small samples.

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.

5.5.3 Selection issues in famine studies

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

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

Our data are not suitable to study the extent of selective mortality, since they only contain the exposed cohort members who are still alive at ages around 50 to 55. Selective fertility can only be controlled for to the extent that it is captured by the region of birth. Some of the studies listed in Section 5.3 that examine long-run effects of the Dutch, Greek and German famines argue that, with these famines, selective fertility does not create a major source of bias. Retrospective information on the parents' social class is typically unavailable. However, Susser and Stein (1994), using military conscription data on men, report that the higher the occupational category of the father, the lower the reduction of

Table 5.3: Effects of hunger at age 6 – 16

	Reduced Form	Instrumental Variables Models		
	Famine at age 6 – 16	2SLS	cond. Wald	cond. Wald – Trend corrected
Effect	-0.011	-0.065	-6.999	0.698
(S.E.)	(0.339)	(1.949)	(1.885)	(1.630)
t-stat.	-0.033	-0.033	-3.712	0.428

Note: Size of treated population = 203; control group = 2470. 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 year of birth.

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. We do not find statistically significant differences (results available upon request) for any of the 10 occupational categories. This suggests that selective fertility is not likely to dominate the estimated effects of interest in the remainder of the chapter.

5.6 Estimation Results

5.6.1 Causal long-run effects of hunger in later childhood

In this section we estimate models in order to reveal whether hunger spells in later childhood causally affect 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 using 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, and see whether these studies yield accurate estimates of the causal effects. After this, we discuss the IV estimation results.

Table 5.3 shows the first set of results. The set of covariates is listed in the notes underneath the table; it is kept very limited because of the reasons detailed in the data

section. The first column concerns the reduced-form estimates of famine exposure. We find that famine exposure in the age window 6 to 16 does not significantly affect adult height.

Table 5.4: First-stage coefficients for probability of experiencing hunger at ages 6–16

Experienced famine being 6-16 (1 = yes)	0.205***	0.173***
	(0.021)	(0.023)
Germany		0.017
		(0.016)
The Netherlands		0.004
		(0.012)
Lived in rural area at age 6		-0.028***
		(0.011)
Year of Birth		-0.003***
		(0.001)
Constant	0.031***	6.217***
	(0.005)	(1.228)
R^2	0.103	0.114
F-Stat.	96.667	27.107
N	2673	2669

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

Next, we verify that the famine instrument is informative, by presenting first-stage results. Table 5.4 shows the results 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. The models yield F-statistics beyond 30, i.e. values exceeding the typically recommended value of 10. We conclude that the instrumental variable is informative.

We use IV analyses to assess the causal effects of hunger at ages 6 to 16. 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 time trend. The residuals are then used as the dependent variable in regressions using the Wald estimator. The results of this last estimation approach are reported in the last column of

Table 5.5: Effects of hunger at age 6 – 16, instrument is famine *duration* in age interval

	Reduced Form Estimate	Two Stage Least Squares
Effect	-0.001	-0.013
(S.E.)	(0.215)	(2.879)
t-stat.	-0.005	-0.005

Note: See table 5.3 for details.

Table 5.3. The IV results suggest that famine induced hunger at age 6 to 16 does not have an impact on adult height. The 2SLS estimates are very small in magnitude. The fully nonparametric Wald estimate is large and significant but loses both size and precision when first detrending height.

We test the robustness of the above results by replacing the binary famine instrument with the length of famine exposure within the defined age interval in months; we rescale this variable by dividing by twelve.[23] We should note that we use this reformulation of the instrument to account for a further dimension of the famine exposure, namely severity. The validity of this proxy is, however, debatable. The length of exposure in our opinion is only an inferior way of including this aspect in our analysis, but it is probably the only one. Estimating the same models as before with the new instrument produces effects as depicted in Table 5.5. The results remain unchanged.

5.6.2 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 and pool this information with potential exposure in utero. Gestation is defined to begin nine months before birth. We require gestation and the famine period to overlap more than two months in order to preclude additional noise. For Dutch individuals we take the region of the accommodation at birth

[23] To be more clear on this, this variable is just the overlab between the treatment window age $[6, 16)$ and the famine interval defined by the calender dates given in the data section.

Table 5.6: Reduced form estimates and treatment effects of hunger at age 0 – 4 or in utero

	Famine at age 0 – 4	Cond. Wald	cond. Wald – Trend corrected
Effect	-0.856	-3.249	-3.413
(S.E.)	(0.342)	(1.991)	(1.722)
t-stat.	-2.499	-1.632	-1.983

Note: Size of population = 2673; thereof famine-exposed at age 0 – 4: = 423; thereof treated / famine-exposed at age 6 – 16 = 203 / 619. Regressions include control country fixed effects and a dummy for urbanization of birthplace and year of birth (last variable 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.

as a criterion for famine exposure and set it to one when this accommodation was located in the Western part of the Netherlands.

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 column of Table 5.6 and show that famine exposure and accompanying potential suffering from hunger reduces adult height by about 0.85 centimeters. If we now take a crude estimate for the effect of famine exposure on the marginal propensity to report hunger to amount to 20 percent, we would expect the true underlying causal effect of hunger on adult height to be around -4.25 centimeters (applying the Wald estimator's formula).

In what follows now, the reports for ages 6 to 16 are taken as the reference point. Hence, we predict hunger propensities for the early years by using conditional expectations for hunger at this later age. This analysis allows to directly compare the size of the commonly computed reduced form effects we presented in the first column of Table 5.6 with quantitatively more reliable causal effects of actual hunger. To this end, we use the 2-sample IV methods outlined above. The last columns of Table 5.6 shows 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. We find that a famine-caused hunger experience early in life has a negative impact of 3 centimeters when using the Wald estimator both with and without correcting for a secular trend in height.[24] The results show that the reduced form estimate presented above underestimates the

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

Table 5.7: Robustness of effects at ages 0–4 or in utero: placebo famine exposure and sensitivity and inclusion of other countries

	Placebo Famine Exposure	Inclusion of other countries
Effect	0.032	-0.851
(S.E.)	(0.032)	(0.343)
t-stat.	0.111	-2.485

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; thereof pseudo-famine exposed = 640. For column 2, original sample has been extended to include France, Belgium, and Italy. Sample Size = 2864; thereof famine exposed = 0. 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 year of birth.

underlying causal effect of hunger on height by a factor 4 amounting to more than 2 centimeters.[25]

Adding 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 5.4. We first estimate a first stage, i.e. the effect of pseudo famine exposure on the probability to report hunger for ages 6 to 16. These turn out to be statistically significant; however, the size of the effects is less than half the size we find in the original analysis. As noted in Section 5.4, this may reflect the effect of World War II and its aftermath in Belgium, France and Italy. Next, we estimate placebo reduced-form relationships of pseudo famine exposure on health outcomes for both ages 6 to 16 and in utero to age 4. These results are displayed in the first column of Table 5.7 and provide no significant evidence along the lines of the patterns we found in the main analyses. This confirms that we identify 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

[25]We also used a 2S-2SLS estimator for this exercise. However, this restricts both the relationship of height on hunger and the first stage to be linear. Using this alternative estimation technique produces causal effects in the same order of magnitude and significance.

Table 5.8: Reduced form estimates and treatment effects using varying treatment windows

	Reduced Form Effect	Cond. Wald	cond. Wald – Trend corrected
	Treatment Window: Age 0–6 and/or in utero		
Effect	-0.683	-3.616	-2.897
(S.E.)	(0.337)	(1.871)	(1.580)
t-stat.	-2.028	-1.933	-1.833
	Treatment Window: Age 0–3 and/or in utero		
Effect	-0.657	-1.903	-2.551
(S.E.)	(0.356)	(2.048)	(1.829)
t-stat.	-1.848	-0.929	-1.395

Note: Size of population = 2673; thereof famine-exposed at age 0 – 6 (0 – 3): = 563 (341); thereof treated / famine-exposed at age 6 – 16 = 203 / 614; Regressions include control country fixed effects and a dummy for urbanization of birthplace and year of birth (last variable 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.

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 from this exercise is displayed in the second column of Table 5.7. 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 adverse non-famine conditions.

Additional sensitivity analyses In addition to the sensitivity analyses discussed so far, we performed a range of estimations to assess the robustness of the results with respect to a number of assumptions and decisions concerning the operationalization of the key variables. 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 included in the famine period.

Further, we may vary the age intervals within which early-life conditions are assumed to extern their influence. For example, we may exclude in utero from the age interval that merges in utero with age 0-4, or we may exclude age 4 from these intervals, or consider a wider treatment window. Table 5.8 gives two examples where we define the treatment window to be in utero or at ages 0 to 6 or 0 to 3, respectively. For the extended treatment window, we can see that the effects remain relatively stable even though the reduced form estimate somewhat shrinks in absolute value; this may be suggestive for the

very early years of childhood being the most critical age driving our results. Defining the the treatment window as running from 0 to 3 shows that we here run into sample size problems: even though the reduced form estimate is significant and negative, the results from using the Wald estimators are not despite being negative and relatively large in absolute value. We thus conclude that our original choice of ages 0 to 4 or in uetro exposure as the treatment window is the best solution to the trade off between effect relevance and statistical precision. Additionally, we note that changing the correctly recalled reference age window from 6 to 16 in a similar way does also not change the results (results can be made available upon request).

Last, we may drop our measure of urbanization at the reference age from the covariates in order to see whether results are sensitive to this change. Concerning the outcome measures, we may trim extreme values. It turns out that all results are insensitive with respect to such changes. 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. Finally, we perform separate estimations by country. It turns out that the current samples by country are too small for meaningful analyses. Furthermore, results by country are more sensitive to effects of cohort-specific events.

5.7 Conclusions

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

For men aged 6-16, we find average adult-height effect of undernutrition of about 6 cm, which is however levels out after introducing a secular time trend in height. If restricted in nutritional supply in the interval from in utero until age 4, men who suffered from low nutrition in infancy, display a negative height effect of more than 3 cm. This effect remains after controlling for trends in height over time.

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 5.3). However, our estimated causal effects are about four times larger than our reduced-form famine effects. This emphasizes the importance of nutrition in early childhood - over and above the findings and statements in the famine literature.

The difference between contextual and causal effects is related to the rate of "compliance" to the famine. In our study, the fraction of "compliers" is close to 20%. To the extent that reduced-form studies implicitly assume perfect compliance, this assumption seems untenable even for shocks as severe as the famines used in this study.[26]

The analyses in this chapter are restricted by some notable data limitations. First, the samples are small. Admittedly, the full SHARE data cover many European countries and include over 20,000 elderly respondents, but we only use men in the 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 say 4 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 health outcomes later in life.

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

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

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

Chapter 6

Intergenerational Effects of Economic Distress: Paternal Unemployment and Child Secondary Schooling Decisions

6.1 Introduction

Does an economic crisis have adverse effects on the next generation? A large body of literature shows that job loss reduces future earnings, future employment prospects, marital stability and (mental) health of the unemployed. However, if the children of the unemployed are equally affected, a crisis may have long run consequences even on next-generation human capital. This chapter investigates how unemployment resulting from temporary shocks in the paternal labor market affects child schooling decisions.

There are at least two potential mechanisms that link paternal unemployment with child human capital. First, paternal joblessness may reduce parental monetary and non-monetary investments into child skills and competencies. Second, paternal unemployment may act as a temporary shock to a child's confidence in being able to graduate successfully if it occurs during a critical decision period. This chapter shows that paternal unemployment adversely affects children's educational choices but not immediate school performance.

To understand the impact of labor market fluctuations and paternal unemployment on upper secondary schooling decisions, I estimate a latent variable model for the joint probability of paternal unemployment and child upper secondary school choice using the cyclical component in adult male unemployment as an exogenous shifter for paternal unemployment. I focus on paternal instead of maternal unemployment because the father

tends to be the main breadwinner and because psychological effects of unemployment tend to be higher for men than for women (Theodossiou, 1998). Moreover, I show that the association between paternal unemployment and child schooling is much stronger for paternal than for maternal unemployment (for similar findings on the effects of paternal and maternal unmeployment, see Kalil and Ziol-Guest, 2008; Rege et al., 2011).

The decision whether to complete upper secondary schooling is vital in the German context. An upper secondary schooling certificate entitles individuals to a large range of white-collar vocational training positions and is a prerequisite for university attendance (Jenkins and Schluter, 2002). Children make this choice approximately at age 16, a time where they are still highly dependent and influenced by familial distress factors, such as parental unemployment.

The relationship between paternal unemployment and child education decisions cannot be investigated using experimental methods. Hence, a major concern is that paternal unemployment and child schooling may be jointly dependent on unobserved confounders. I address this concern by matching German household panel data with macro data on 97 regional economic centers for the years 1998-2009, including the years of the most recent crisis. I construct the cyclical component of regional labor market fluctuations in the labor market of the father and use this as an exogenous shifter for paternal unemployment. The identification strategy relies on the assumption that temporary unemployment shocks in the paternal labor market only influence child schooling decisions through paternal employment. This assumption is not innocuous if regional unemployment is correlated with youth unemployment. Therefore it is important to control for apprenticeship vacancies, youth unemployment, changes in the tax base as well as permanent unemployment in the region of residence. The underlying model also comprises indicators for urbanization, parental age and parental education and a number of other background variables because regions with different characteristics may be affected by recessions. I find that the null hypothesis of child schooling and paternal unemployment being independently determined by unobservables can no longer be rejected. Therefore, part of my results build on a matching assumption.

The data come from the German Socioeconomic Panel (GSOEP), a representative longitudinal micro-dataset that contains a wide range of socio-economic information on individuals in Germany, comprising yearly follow-ups during 1984-2010 (Wagner et al.,

2007). Information on the children stem from a special youth survey comprising information on 17-year-old children of the responding households, collected in the years 2000-2010. The data are well-suited to my analysis because they can be linked to a large number of regional economic indicators and contain a vast number of parental and child characteristics. These allow me to study heterogeneity in the effect of paternal unemployment with respect to paternal cognitive ability, education, age, school tracks and gender. Moreover, detailed information on child characteristics and economic preferences allow me analyze several potential channels through which the paternal unemployment effect operates, such as the effect of paternal unemployment on expected school success.

The main finding of this chapter is that paternal unemployment causally reduces offspring educational attainment and that a child's subjective probability of school success is an important mechanism. Detailed results are (1) the reduced form effect of a one percentage point increase in the cyclical component of regional unemployment amounts to a decrease in the probability of child upper secondary school choice by 2 percentage points, of a base level of 52%. (2) paternal unemployment lowers the probability of upper secondary school completion by 18 percentage points. (3) paternal unemployment reduces the subjective probability of successful school completion by 11 percentage points and if an individual believes that school completion is rather unlikely, the probability of upper secondary school choice decreases by 6.5 percentage points. After controlling for parental background variables, this subjective probability of school success accounts for 2 - 7.5 percent of the overall unemployment effect. (4) my findings can be explained by a theoretical framework that allows paternal unemployment to affect the assessment of the return to education through expected school success.

The contribution of this chapter is fourfold. First, it is the first chapter that uses variation in the cyclical component of regional adult male labor market fluctuations as an exogenous shifter for paternal unemployment.[1] In a linear IV-setting with heterogenous effects, the effect I identify is thus a weighted local average treatment effect for children of individuals that suffer from unemployment due to a regional labor market downturn. This is the relevant effect for policy makers who want to obviate second order effects of an economic crisis (Carneiro et al., 2011). In this regard the chapter is related to literature on other

[1]To my knowledge it is also the first chapter to use cyclical macro variation for the most recent crisis years to investigate second order effects on next generation outcomes.

topics using regional unemployment shocks as exogenous shifters for endogenous variables (e.g., Ham and Jacobs, 2000; Ginja, 2010).

The second contribution of this chapter is related to the use of household data. While most studies in this literature use administrative data with limited background information, household data allow me to investigate the impact of paternal unemployment on behavioral traits of the child and heterogeneity in the effect of paternal job loss for different groups of individuals. Whereas most of the existing literature focuses on income or parental investments as potential mechanisms (see e.g. Dahl and Lochner (2012); Blau (1999); Rege et al. (2011)) this chapter innovates by laying the main focus on the psychological channels.

The third contribution is to explain my findings within a simple theoretical framework where paternal unemployment affects the return to education through the subjective probability of successful school completion. In this model children make upper secondary schooling decisions by comparing discounted wage flows for each schooling choice, where the high-education wage stream is weighted by the probability of actually achieving the higher education level. This study thus takes the standpoint of the child and does not focus primarily on parental investments.

The chapter is organized as follows. Section 6.2 reviews the related literature and describes the institutional setting. Section 6.3 describes a simple framework that relates paternal unemployment with the subjective school success probability and child upper secondary education decisions. Section 6.4 describes the data, and Section 6.5 discusses the empirical estimation strategy and econometric models used. Section 6.6 presents the results of the analyses. Section 6.7 concludes.

6.2 Literature and Institutional Background

It is well-established that job loss has large adverse consequences on adult individuals. Fewer studies investigate the impact of parental job loss on child outcomes. Hence, Section 6.2.2 reviews the related literature whereas Section 6.2.3 and Section 6.2.4 describe the institutional labor market context as well as relevant parts of the German education system.

6.2.1 Literature on paternal unemployment and child outcomes

It is well-established that job loss has large adverse consequences on adult individuals. Recently, a limited number of studies have also investigated impacts of parental job loss on various child outcomes. I use household data that are ideally suited to answer the question at hand. Since these data were collected in Germany, it is important to understand the institutional context in which school decisions are made. Hence, Section 6.2.2 reviews the related literature whereas Section 6.2.3 and Section 6.2.4 describe the institutional labor market context as well as relevant parts of the education system.

6.2.2 Literature on paternal unemployment and child outcomes

A large number of studies investigate how unemployment and involuntary job loss affect an individual's well-being. Jacobson et al. (1993) and Ruhm (1991) find that permanent earnings of displaced workers are lowered by 25 and 10-13 percent respectively. It is also well-established that unemployment has quite dramatic effects on health, mental health and life satisfaction (Sullivan and Von Wachter, 2009; Eliason and Storrie, 2009; McKee-Ryan et al., 2005; Theodossiou, 1998; Kassenboehmer and Haisken-DeNew, 2009).[2] It also has negative effects on family well-being, marital disruption and family relocation (Charles and Stephens, 2004; Astone and McLanahan, 1994; Kind and Haisken-DeNew, 2012).

Despite the importance of the subject, few studies relate parental unemployment to child outcomes. Important exceptions are Oreopoulos et al. (2008), Bratberg et al. (2008) and Rege et al. (2011).[3] Rege et al. (2011) use Norwegian register data to estimate the causal effect of parental job loss due to plant closures during grade seven on the Grade Point Average (GPA) after grade ten. The analysis is based on a matching assumption and the authors control for a large number of region, industry and school fixed effects. They find that paternal job displacement reduces GPA by 6 percent of a standard deviation, while maternal job loss leads to a nonsignificant increase in GPA. Furthermore, the effect is

[2] McKee-Ryan et al. (2005), in a meta study summarizing 104 other studies, find large effects of unemployment but no significant effects of the current unemployment rate on mental health.

[3] Kalil and Ziol-Guest (2008) and Stevens and Schaller (2011) find that paternal job loss results in a higher probability of offspring grade retention in the Survey of Income and Program Participation (SIPP).Gregg et al. (2012) in a recent publication using the British Cohort Study Using the British Cohort Study find that children with displaced fathers obtain lower grades, lower wages and are at a higher risk of youth unemployment.

largest in municipalities with non-decreasing unemployment rates and below median pre-closure earnings. Focusing on paternal investment channels, the authors report that the effect of paternal job loss does not pass through subsequent earnings or time allocation of mothers, divorce or residential reallocation. Therefore, the authors conclude that parental mental health is the driving mechanism. A second paper by Bratberg et al. (2008) uses Norwegian data, too, and investigates the effect of paternal displacement when children are of age 12-16. The authors analyze matched employer-employee panel data but find no significant effects on earnings, non-employment or registered unemployment of the next generation. The third paper uses information about plant closures in Canadian administrative data (Oreopoulos et al., 2008). The authors show that sons in the age group 10-14 of displaced workers have adult annual earnings that are about 9 percent lower than similar children of fathers who did not experience an employment shock. The effect is largest for families in the lowest quartile of the income distribution. When it comes to mechanisms, they find displacement to slightly affect mobility but mobility not to affect child outcomes. Moreover, their results indicate that displacement does not affect marital status or spousal income. Therefore, Oreopoulos et al. (2008) conclude that income is the driving mechanism. The results of these studies are specific to individuals affected by plant closings. Also, fathers employed at closing plants are likely to be less skilled at foreseeing the future or less willing to change jobs beforehand (Pfann and Hamermesh, 2001).

From a methodological point of view, this chapter is close to other studies that use exogenous variation as a shifter for unemployment or family resources. Examples are Ham and Jacobs (2000) who use the unemployment rate in the household head's occupational category as an instrument for family resources or Fougère et al. (2009) and Gould et al. (2002) who use predicted changes in employment shares of different demographic groups in different regions as instruments for youth unemployment.

A large body of literature studies the effects of parental income and maternal employment patterns on child outcomes (Blau, 1999; Dahl and Lochner, 2012).[4] These studies find that the effect of income on child development is significant but modest, and less important than child characteristics and other family background variables.[5]

[4]Maternal employment tends to have small negative effects, especially while the children are small (Ruhm, 2004).

[5]Jenkins and Schluter (2002) also find small income effects for child achievement in Germany.

Most of this literature investigates the effect on small children. But while it is well-established that parental investments and parental income matter more for the younger the child, other mechanisms may be more important when children are older (Tominey, 2009; Cunha and Heckman, 2007).

6.2.3 Institutional changes and the German labor market during the great recession

The exogenous variation used in this chapter comes from temporary unemployment shocks, and the identification of the model is based on temporary employment fluctuations. Therefore, it is important to note that the period under study includes the 2008/2009 world recession. The crisis has hit Germany harder than most other OECD countries, mainly because the country heavily depends on export markets, which more or less collapsed during the crisis. However, despite a 5% drop in GDP, the crisis has affected the German labor market much less than in many other European countries. There were almost no mass layoffs and no general feeling of panic. Assisted by short-time work schemes, many firms buffered capacity (Möller, 2010).

The crisis has affected German industries very unequally. 37% of all firms reported to be affected by the crisis, but as many as 70% of metal producers (Möller, 2010). Because many of these producers are located in Western and Southern Germany, there were some regions that were quite heavily affected. These were mostly regions that had been economically strong before the crisis (Fuchs and Kempermann, 2011).

The period under investigation comprises the so-called Hartz reforms (I to IV) between 2003 and 2005, which was a substantial reform of the unemployment benefit system that lead to an increase in unemployment hardship. In 2005, unemployment and social assistance were merged into a single means-tested welfare payment. Since then, eligibility for unemployment benefits depends on being physically and mentally capable of working for at least 15 hours per week, active job search and the willingness to participate in welfare to work programs. Moreover, non-compliance to the unemployment benefit rules or the rejection of job offers can be sanctioned by means of temporary benefit cuts (Huber et al., 2011). The reforms have lead to a decline in the so called natural rate of unemployment by increasing the incentives for unemployed to search for and accept new jobs.

6.2.4 School choice in the German context

School choice at age 16 is an important stepping stone towards obtaining an upper secondary school degree (Abitur/Fachabitur). Taking 2-3 years to complete, it serves as a school graduation certificate and university entrance exam. Moreover, it grants access to colleges and universities and is a prerequisite for many apprenticeship and vocational training positions. Individuals who do not complete an upper secondary school degree mostly continue vocational training and eventually take up work in blue collar occupations.

Upper secondary school choice in Germany is influenced by a system of early tracking. At age 10-12, children are tracked into one of three separate hierarchical school strands. Although, formally, students from all three tracks can obtain an upper secondary school degree, the latter is more difficult if tracked into Haupt- or Realschule (the lower tracks) than Gymnasium (the highest track) (Jenkins and Schluter, 2002).

Results by Dustmann et al. (2012) imply that, at least for individuals at the margin, tracking is less decisive for obtaining a certain educational degree than upper secondary school choice itself. This is explainable by the substantial amount of student up- and downgrading between track types at the time of upper secondary school choice (Dustmann et al., 2012). At that time, individuals holding a German general secondary school degree (Hauptschulabschluss) or a German intermediate school degree (Realschulabschluss) can obtain an upper secondary school degree (Fachabitur or Abitur) if they change school after grade 10 or if they graduate from specialized vocational schools.

Rules and regulations concerning degrees and compulsory schooling ages vary greatly between the different federal states. While, in most states, schooling is compulsory until age 18 (secondary school + vocational school), some states only require 9 years of schooling.[6] Furthermore, secondary general schools finish after grade 9 in most federal states and after grade 10 in others. When investigating the impact of paternal unemployment at child age 16, it is thus very important to control for state of residence.

[6]These are Saarland, Thuringia and Hesse.

6.3 A Theoretical Framework for the Impact of Paternal Unemployment on Child Upper Secondary School Choice

Education choices are pivotal to the amount of human capital an individual accumulates in life. While skills and abilities are predominantly a function of parental investments during early childhood, human capital accumulation is essentially a matter of own choice at later ages. Section 6.3.1 considers a simple theoretical framework where each individual chooses between obtaining upper secondary education or not. Paternal unemployment influences this choice through its impact on the subjective probability of upper secondary school success as laid out in Section 6.3.2.

6.3.1 A human capital investment model

The effect of paternal unemployment can be incorporated into a Roy model of human capital investment decisions where individuals weigh the discounted flow of wages of the upper secondary school degree wage path with the subjective probability of being successful at obtaining the degree.

In a typical model of human capital investment, individuals make human capital investment decisions based on the present value of future wages (Becker, 1993). Individuals weigh the benefits of continued upper secondary schooling against the benefits of dropping out at age 16 when they decide whether to continue schooling.

Assume that there are two education levels, $S = \{0, 1\}$, where $S = 1$ denotes holding an upper secondary school degree and $S = 0$ denotes all lower education levels. Furthermore, suppose that there are two wage paths w^l and w^h over T periods of time, where $w^l(t)$ represents the wage in period t for a low educated individual and $w^h(t)$ represents the wage in period t of an individual who holds an upper secondary school degree. Assume that, while wage path l can be obtained with certainty, wage path h depends on successful school completion ($S = 1$) which is uncertain. That is, at age 16, an individual cannot be sure she will be successful at obtaining the corresponding degree whether after 2-3 years. Let $p \in [0, 1]$ denote the subjective probability of successful upper secondary school completion at age 16. Expected future wages conditional on choosing the high schooling

path then depend on p according to:

$$\mathrm{E}\big[w^{S=1}(t)\big] = pw^h(t) + (1-p)w^l(t) \tag{6.1}$$

If an individual chooses to drop out at age 16, this implies $p = 0$ and:

$$\mathrm{E}\big[w^{S=0}(t)\big] = w^l(t). \tag{6.2}$$

In that case the low wage path is a deterministic process from the point of view of the individual.

Agents maximize the expected net present value of education to make their decision. Let V^* denote this latent variable. Then an individual attends upper secondary schooling, $S = 1$, if:

$$V^* \quad \geq \quad 0,$$

and $S = 0$ otherwise. Using Equations 6.1 and 6.2, the net present value of upper secondary schooling, accounting for the discounted flow of ex post earnings is:

$$V^*(w^1, w^0, \delta, t_s, p) = \sum_{t=t_s}^{T} \delta^t \mathrm{E}\big[w^1(t)\big] - \sum_{t=0}^{T} \delta^t \mathrm{E}\big[w^0(t)\big], \tag{6.3}$$

where t_s represents the time required to achieve upper secondary schooling, T is the life horizon, and δ denotes the discount rate, which for is assumed to be constant over time for simplicity. If the decision process is also influenced by monetary costs of upper secondary school choice, such costs would be subtracted in Equation (6.3). However, German schools are almost exclusively public and do not charge fees. One can estimate the effect of an individual's characteristics on the probability of an individual to choose upper secondary schooling school with a reduced-form model using variables that influence earnings, determines the discount rate, δ, and determine the subjective probability of obtaining an upper secondary schooling certificate.

6.3.2 The role of paternal unemployment

Teenagers whose fathers become unemployed are likely to receive a temporary shock to their mental health, self-confidence and locus of control. Furthermore, these children may expect that school support and assistance of their parents will go down in the future. Paternal unemployment therefore reduces the subjective success probability p of obtaining

an upper secondary school degree:

$$p(D = 1) < p(D = 0) \tag{6.4}$$

where $D = \{0, 1\}$ denotes paternal unemployment. It is easy to see that the net present value of upper secondary schooling is increasing in p as long as wages in the high education sector are higher than in the low education sector:[7]

$$\frac{\partial V^*(w^1, w^0, \delta, t_s, p)}{\partial p} = \sum_{t=t_s}^{T} \delta^t \left[w^h(t) - w^l(t) \right] > 0 \tag{6.5}$$

It follows that paternal unemployment has a negative effect on the net present value of upper secondary schooling:

$$\frac{\Delta V^*}{\Delta D} < 0. \tag{6.6}$$

If p can be observed by the econometrician, this simple theoretical framework provides a testable mechanism for the effect of paternal unemployment on upper secondary school choice. Note that I make the assumption that paternal unemployment at age 16 does not have a direct impact on wages or the discount rate of individuals. This is a strong assumption, which fails to hold, for example, if paternal unemployment has a differential impact on an individual's work related skills needed in either of the two wage sectors. Note also that, in order to be able to use fluctuations in the local unemployment rate as an exclusion in the schooling equation, I need to make the assumption that future wage paths are not influenced by labor market fluctuations. E.g., my results would be biased upwards if a recession today permanently reduced wages more along the high education wage path than along the low education wage path. Research shows that temporary labor market downturns can indeed have lasting impacts but that the effect is larger for lower educated workers (Oreopoulos et al., 2012). This research shows that it is important to control for youth unemployment and the availability of vocational training positions.

6.4 Data

I match German representative household data with labor market information on 97 regional economic centers for the years 1998-2009. Section 6.4.1 describes the dataset and Section 6.4.2 explains the coding of the main variables. Section 6.4.3 lays out how

[7] Implicitly I assume that t_e is low enough, such that, if $p = 1$, education pays off in general.

I construct the cyclical component of adult male unemployment in the paternal labor market, and Section 6.4.4 describes the sample.

6.4.1 Dataset and sample construction

My sample is drawn from the German Socioeconomic Panel (GSOEP), a representative longitudinal household dataset that contains a wide range of socio-economic information on individuals in Germany comprising follow-ups for the years 1984-2010. Information was first collected in 1984 for about 12,200 randomly selected adult respondents in West Germany. After German reunification in 1990, the GSOEP was extended to around 4,500 persons from East Germany, and subsequently supplemented and expanded by additional samples.

This study draws on 3,138 individuals, born 1983-1993, from the GSOEP "youth survey", covering the children of all GSOEP panel members. A comprehensive set of background variables, schooling choices, preferences, opinions and traits of these individuals were collected over the years 2000-2010, when the subjects were 17 years of age.

The data provide four advantages. First, using a household's region of residence, the data can be matched with regional labor market information as well as a large battery of other regional measures, such as regional tax income or regional development indicators. Second, the data contain information on youth upper secondary school choice one year after the decision was made, such that revealed education preferences can be observed. Third, the youth data can be linked to detailed parental information including parental investments, skills, living patterns, labor force participation and unemployment histories. Fourth, the data contain rich information on child traits as well as the subjective probability of an individual to successfully complete her education.

I use the following sample selection criteria. First, I exclude all individuals who in elementary school received a track recommendation for the lowest track. The reason is that in most federal states the low-track schools only take 9 years to complete. Moreover, individuals in the lowest track may lack the cognitive ability or opportunity to obtain an upper secondary school degree. Second, I only include individuals who live in the same household with both of their parents. Third, I exclude individuals for which paternal unemployment in the previous year is missing. Fourth, some subjects were already 18 or 19 years of age when first completing the questionnaire in 2001. I exclude these individuals from the sample. Moreover, I also drop individuals with implausible values for paternal

age, missings for the subjective school success probability (very few) and missing regional indicators.[8] Last, I exclude all students with missing information in any of the covariates displayed in Table 6.3. Table 6.1 displays the final sample size (N=2,326), the fraction of individuals who choose upper secondary schooling and the fraction of individuals whose father is unemployed when they are 16 years of age. 52% of all individuals in the sample have chosen upper secondary education, compared to an average of 54% in Germany (in 2012) (OECD, 2012). The unemployment rate of 10% for the fathers in the sample is slightly higher than the average official unemployment rate of 9.6% over that period, which is due to the fact that I also include non-employed fathers who are currently not part of the labor force. The third reports the mean and standard deviation of the cyclical component of regional unemployment in the paternal labor market.

Table 6.1: Proportions of youths with higher education (*outcome*), paternal unemployment (*treatment*) and cyclical component of regional unemployment in paternal labor market (*exclusion*)

	Youths, age 17		
	Proportion	**SD**	**N**
Upper_secondary_education	.52	.5	1205
Father_unemployed	.1	.296	225
Regional_unemployment	.66	1.377	2326

Source: GSOEP youth sample 2000-2010.

6.4.2 Coding of main variables

Upper secondary school choice is the main outcome of interest. This study classifies all individuals as having chosen upper secondary education, if according to the international Comparative Analysis of Social Mobility in Industrial Nations (CASMIN), they have an education level that corresponds to CASMIN-categories (2c), (3a) or (3b).[9] Here, I use the latest available information. Furthermore, all youths who have not yet completed their education at the time of the last interview are classified as having chosen upper secondary education if they are still in school and are planning to take an upper secondary school exam that entitles them to enter a teaching college or university (German Abitur or Fachabitur) at the time of the interview.[10]

[8] I drop all individuals with fathers who are younger than 32.

[9] See Section 6.2.4 for a description of the German institutional context.

[10] A second measure of upper secondary schooling is whether an individual is still at school at age 17. This measure is used for robustness checks.

Paternal unemployment is the central explanatory variable. Fathers are classified as unemployed if they are not working at the time when the child is 16 years of age. This also discouraged workers who are not currently looking for a job count as unemployed. The assumption behind this is that voluntary unemployment among the fathers of school age children is rare. Business cycle fluctuations and sanctioning (see Section 6.2.4), on the other hand, may cause part of the unemployed workforce not to actively search for a job. Table 6.2 displays raw correlations between upper secondary school choice and a variable that indicates whether father or mother became unemployed in the past year. The table shows that (a) in terms of correlations, it hardly matters whether all unemployed or only those actively searching for a job are classified as unemployed;[11] (b) The association between paternal unemployment and school choice is three times as large as between maternal unemployment and school choice.[12]

Table 6.2: Raw correlation: paternal/maternal change to non(un-)employment and child upper secondary schooling

Newly unemployed	**Upper Secondary Education**
Father becomes unemployed	-0.3160***
	(0.053)
Father becomes involuntarily unemployed	-0.3472***
	(0.050)
Mother becomes unemployed	-0.1167**
	(0.057)
Mother becomes involuntarily unemployed	-0.1024
	(0.068)
Observations	2120 (father) 2092 (mother)
Covariates included	NO

Standard errors in parentheses
Source: GSOEP Youth Sample.
Note: Standard errors are robust.
Raw correlations displayed, no covariates included.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The subjective probability of school success is the main potential mechanism of interest. In the youth survey, individuals are asked "What is the probability that you will successfully complete your training or further studies?"[13] Individuals can indicate a probability in decimal steps. Note that the question is framed rather broadly and does

[11] All robustness checks show that the results with this coding of unemployment are conservative. If only those individuals are classified as unemployed who are not working and actively looking for a job, coefficients increase slightly and become somewhat more significant.

[12] Table 6.15 repeats the same correlations with GPA instead of upper secondary school choice and shows that there is no significant negative association between parental unemployment and child GPA as defined by the average of the most recent grades obtained in German, math and first foreign language.

[13] "Wie wahrscheinlich ist es, dass Sie Ihre Ausbildung oder Ihr Studium erfolgreich abschließen?"

not specifically address upper secondary schooling. Yet, what this study is interested in is whether an individual loses confidence in her abilities to complete further education or training due to paternal unemployment, which this question should adequately capture. I use the subjective probability as described above as well as an indicator variable $\mathbb{1}[Percentage > 50]$ for whether an individual believes she is rather likely to complete her education.

To be able to control for childhood circumstances, I construct a large set of background variables comprising parental age, family size paternal education, parental investment variables, nationality and region. Moreover, in order to proxy cognitive skills and to account for the fact that schooling decisions may depend on prior track attendance (see Section 6.2.4), I include an individual's track recommendation after elementary school. In Germany, every student receives a track recommendation during 4th grade by her elementary school teacher. In most German states, track recommendations are non-mandatory. In some states they are compulsory. Last, I construct time and state fixed effects.

6.4.3 Regional labor market information

Using individual identifiers for 97 regional economic centers in Germany, I match the GSOEP data with external data on local economic and labor market variables for the year in which the child was 16 years old.[14] Macro variables are obtained from the German Federal Institute for Research on Building, Urban Affairs and Spatial Development (Bundesinstitut für Bau-, Stadt- und Raumforschung (BBSR), 2011). Specifically, I take the regional unemployment for male individuals in the age group of the father and deduct mean unemployment in that region over the entire observation period. Besides, I use year fixed effects in all specifications to account for institutional changes and country-wide shocks over time.[15] The exogenous labor market shifter is given by:

$$Z_{t,r,a} = A_{t,r,a} - \mu_{r,a},$$

[14]The 97 labor market regions are planning entities of the federal states in Germany and borders were drawn such that they reflect local labor markets and commuting areas. They are comparable to US commuting zones (see e.g. Autor and Dorn, 2009).

[15]See Section 6.2.3.

where $Z_{t,r,i}$ denotes the cyclical component of adult male unemployment in year t, region r and age group a. The average unemployment rate $\mu_{r,a}$ is given by:

$$\mu_{r,a} = \frac{1}{T}\sum_{t=1}^{T} A_{t,r,a} \qquad \text{for } t = 1998, ..., 2009$$

Region-age-gender specific unemployment rates imply, for example, that, if the father is 56 years of age, the unemployment rate A_t is given by the regional unemployment rate for males aged 55+.[16] Fluctuations in the regional unemployment rate for older men are included in the data, because they are likely correlated with youth unemployment, taxed-based school financing and the availability of apprenticeship training positions.

Table 6.3: Summary statistics, background variables

Variables	**Youths**		
	Father employed	**Father unemp**	**P-value**
Outcome			
Secondary schooling	0.55	0.25	0.00
Background variables			
Maternal age	44.54	44.56	0.95
Paternal age	47.04	49.14	0.00
One sibling	0.50	0.28	0.00
Two siblings	0.26	0.30	0.17
Three or more siblings	0.12	0.28	0.00
Father secondary intermediate	0.33	0.25	0.03
Father grammar school	0.31	0.16	0.00
Mother secondary intermediate	0.45	0.30	0.00
Mother grammar school	0.25	0.14	0.00
Childhood in large city	0.20	0.24	0.13
Childhood in medium city	0.19	0.20	0.55
Childhood in small city	0.27	0.23	0.21
Sex, male=1	0.51	0.48	0.46
Permanent component of unemployment	7.36	8.21	0.00
German nationality	0.95	0.84	0.00
Father has German nationality	0.92	0.80	0.00
Father cognitive skills	0.11	-0.30	0.00
Youth local labor market (mean deviation)			
Vocational training positions per 100 applicants	-0.78	-0.74	0.85
Youth unemployment	0.07	0.20	0.03
N	2326		

Source: SOEP youth data, waves 2000-2010. Own calculations.
Notes: p-values of a two-sided t-test for differences in means are reported.

6.4.4 Characteristics of the sample

Table 6.3 summarizes characteristics of the pooled sample of the 2,326 youths I analyze. The summary statistics clearly show that the children of unemployed fathers differ from

[16] An alternative way of defining the cyclical component of regional labor market fluctuations would be to subtract trend unemployment, as computed using the Hodrick-Prescott Filter. However, due to the extremely short time series available for each region, this approach would be dominated by endpoint problems. Robustness checks show that the results are robust to this alternative definition.

children of employed fathers in almost all characteristics. First, only 25% of the children of unemployed fathers choose upper secondary schooling, while these are 55% among families with employed fathers. Individuals with unemployed fathers tend to have more siblings, older fathers, less educated parents, fathers with lower cognitive skills, and are more likely to have a non-German background. Concerning child characteristics, Table 6.4 shows that children of jobless fathers are less likely to believe that they will successfully complete their education and display a significantly lower locus of control and a lower GPA.[17] Youth mental health is measured by the Mental Component Summary Scale (MCS), one of the two sub-dimensions of the SF-12 questionnaire and risk aversion is measured by an individual's willingness to take risks (for a description of the risk measure see Dohmen et al., 2011). Locus of control, youth mental health and risk aversion are standardized to have mean zero and standard deviation one. GPA is coded by averaging last school grades in math, German and their first foreign language. German grades have been reversed now ranging from 1 to 6 where more is better. Certainly, some of the

Table 6.4: Summary statistics, subjective school success and child traits

Variables	Youths		
	Father employed	**Father unemp**	**P-value**
Subjective school success			
Probability, successful school completion	78.59	74.67	0.00
Probability, successful school completion>50 %	0.88	0.81	0.01
Child trait channels			
Youth mental health	-0.01	0.12	0.07
Youth risk aversion	-0.01	0.04	0.52
Youth locus of control	0.02	-0.21	0.00
GPA			
Youth grade point average	0.02	-0.11	0.08
N	2326		

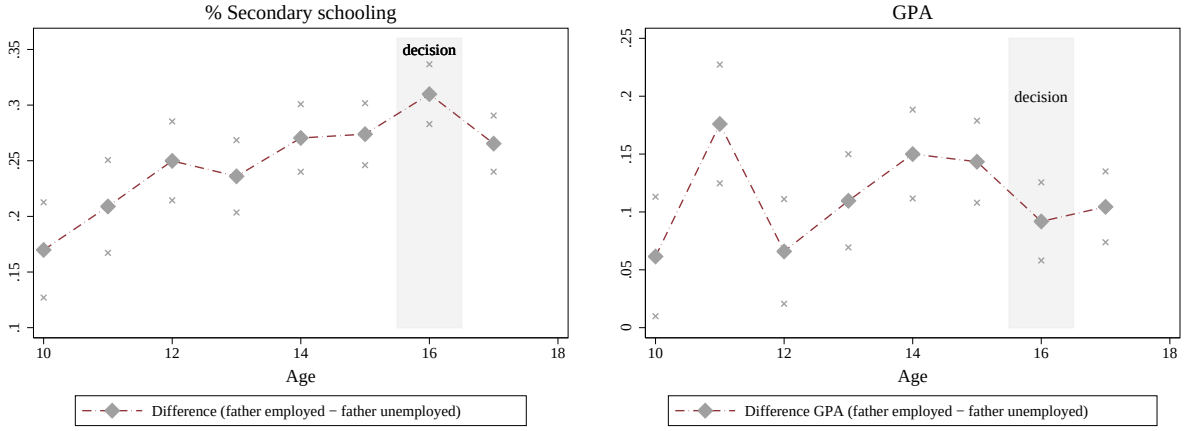
Source: SOEP youth data, waves 2000-2010. Own calculations.
Notes: p-values of a two-sided t-test for differences in means are reported.

observed difference in child upper secondary schooling cannot be ascribed to the causal effect of unemployment but is driven by observed and unobserved confounders. Panel (a) of Figure 6.1 provides a graphical assessment of the average difference in upper secondary schooling probability between youths whose fathers were employed or unemployed at different ages of the child. The graph shows that this difference in outcomes increases the closer in time paternal unemployment occurs to the education decision of the child. This

[17]For a description of the locus of control measure and underlying construct in the GSOEP see Piatek and Pinger (2010).

indicates a likely causal effect of unemployment. Panel (b) shows that no such relationship exists when looking at child GPA as an outcome.

Figure 6.1: % upper secondary schooling and GPA by paternal employment status and age.



(a) Upper secondary schooling **(b)** GPA

Notes: GSOEP youth sample 2000-2010. Sample contains all individuals whose parental unemployment history is available for at least the past 3 years.

6.5 Estimation Strategy

The theoretical framework of Section 6.3 has shown that paternal unemployment is likely to reduce the probability of upper secondary schooling and that the perceived probability of school success is a crucial parameter for the decision about upper secondary schooling. This section explains how I identify and estimate (a) the overall causal effect of paternal unemployment, (b) a latent factor model for paternal cognitive ability, (c) the direct, indirect and total effect of paternal characteristics, (d) the effect of paternal joblessness on the subjective school success probability, and (e) the impact of that probability on schooling decisions.

6.5.1 Simultaneous equation bivariate probit model

The utility of upper secondary schooling and the disutility of unemployment are unobserved latent variables, for which only final outcomes are observed. Because unobservables that drive paternal unemployment (D) and child secondary school choice (E) are likely to be correlated, it is important to jointly estimate the probability of

paternal unemployment and of child upper secondary schooling. The model is:

$$\begin{aligned} U^*_{S,i} &= \beta D_i + \alpha'_S X_i + \sum_{r=1}^{R} \gamma_S d_{r,i} + \sum_{t=1}^{T} \tau_S d_{t,i} + \lambda_S \theta_i + \epsilon_{S,i}, \qquad S_i = \mathbb{1}\left[U^*_{S,i} \geq 0\right] \\ U^*_{D,i} &= \alpha'_D X_i + \delta Z_i + \sum_{r=1}^{R} \gamma_D d_{r,i} + \sum_{t=1}^{T} \tau_D d_{t,i} + \lambda_D \theta_i + \epsilon_{D,i}, \quad D_i = \mathbb{1}\left[U^*_{D,i} \geq 0\right], \end{aligned} \tag{6.7}$$

where U^*_S and U^*_D denote latent (dis-)utility from education and unemployment respectively. X is a vector of background variables, Z denotes the cyclical component of adult male unemployment as defined in Section 6.4.3, and θ is latent paternal cognitive ability. $d_{r,i}$ and $d_{t,i}$ denote state (or region) and time dummies. A list of all included explanatory variables in each equation is given in Table 6.6. $(\epsilon_{S,i}, \epsilon_{D,i})$ are jointly distributed as standard bivariate normal with correlation ρ and independent of Z. I use standard maximum likelihood methods to estimate the parameters β, δ, α_D, α_S, λ_D, λ_S, γ_S, γ_D, τ_S, τ_D and ρ.[18] I compute standard errors that are robust and clustered at the level of the regional economic centers. For each model, I conduct a Likelihood Ratio (LR) test for the absence of correlation in the model under the null hypothesis that ρ equals zero. If the null hypothesis cannot be rejected, I report parameters of a restricted model with $\rho = 0$. There is a large number of different marginal effects that can be computed for this model. I compute average marginal effects for the unconditional probability that $S = 1$ (or $D = 1$), which is the effect of interest in this chapter. The marginal effect for continuous covariates is given by:

$$\frac{\partial \mathrm{E}[S|\mathbf{x}]}{\partial \mathbf{x}} = \Phi(\mathbf{x}'\zeta)\zeta, \tag{6.8}$$

where $\mathbf{x}$ denotes a combined vector of all explanatory variables and ζ a vector of all coefficients, some of which may be zero for variables that only appear in the other equation. For discrete variables, finite differences are computed. Standard errors of the marginal effects are bootstrapped using 200 bootstrap replications.

[18] The likelihood is given by:

$$\ln L = \sum_{i=1}^{N} \ln \Phi_2(a_{i,S}, a_{i,D}, \rho)$$

where Φ_2 denotes the bivariate normal cdf, $a_{i,S} = (2S-1)(\beta D_i + \alpha'_S X_i + \sum_{r=1}^{R} \gamma_S d_{r,i} + \sum_{t=1}^{T} \tau_S d_{t,i} + \lambda_S \theta_i)$ and $a_{i,D} = (2D-1)(\alpha'_D X_i + \delta Z_i + \sum_{r=1}^{R} \gamma_D d_{r,i} + \sum_{t=1}^{T} \tau_D d_{t,i} + \lambda_D \theta_i)$.

6.5.2 Latent factor model

In order to account for paternal cognitive ability, I use a factor model as an integral component of the simultaneous equation model described above. Paternal cognitive skills are assumed to depend on multiple measures M_k where $k \in \{1...K\}$ and K is the total number of measures available.

A factor model is necessary here to account for the fact that different measurements are going to be correlated to a different degree with the latent construct. In a factor model, different weights, called factor loadings, are estimated. By estimating a factor model, one can account for measurement error in proxies and avoid attenuation bias. The cognitive skill measurement system is:

$$M_{Ck,i} = \lambda_{Ck}\theta_{C,i} + \epsilon_{Ck,i} \qquad \text{for } k = 1, ..., K.$$

where λ_{Ck} are factor loadings associated with measurement k. Factor loadings are allowed to differ across equations giving measurements different weights. Since the scale of each factor is arbitrary, I restrict the variance of the factor to equal unity and require $K > 2$ for identification. In addition, $\mathrm{E}[\epsilon_{\theta C k}] = 0$ and $\mathrm{E}[\theta_{C,i}] = 0$. I use the Bartlett method to obtain unbiased factor scores (Bartlett, 1937).

Cognitive skills are measured using a test of symbol correspondence (administered in 2006) that was specifically developed for the GSOEP and corresponds to a sub-module of the Wechsler Adult Intelligence Scale (Lang et al., 2007). Missing cognitive ability measures were imputed on hands of information about an individual's education as well as family education measures using linear regression.

6.5.3 Linear IV

To investigate the effect of paternal unemployment on the subjective probability of school success, child mental health and locus of control, I use the Two-Stage-Least-Squares (2SLS) estimator with paternal unemployment as a treatment and the cyclical component of adult male unemployment in the region as an instrument. The estimator is:

$$\beta_{2SLS} = (\hat{D}'\hat{D})^{-1}\hat{D}'Y, \tag{6.9}$$

where Y denotes the outcome of interest and $\hat{D} = Z(Z'Z)^{-1}Z'D$. To use Z as an exclusion in Equation (6.7) or as an instrument in Equation (6.9), it is important that the reginal

Table 6.5: First stage regression of paternal unemployment on the age-specific regional unemployment rate

Paternal unemployment	**All**	
	OLS	**Probit**
Cyclical component of adult male unemployment	0.01476***	0.01294***
	(0.005)	(0.004)
Observations	2326	2326
Covariates included	YES	YES
F-stat (β (instrument)=0)	9.673	
R-squ adj./Ps R-squ.	0.086	0.139

Standard errors in parentheses
Source: GSOEP Youth Sample.
Note: Standard errors clustered by region. Coefficients of probit equations are average marginal effects. The analytical sample on which these estimates are based consists of all GSOEP youths that have no missings in any of the covariates. For covariates included see Table 6.6.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

labor market instrument is informative. Table 6.5 shows the results of a linear regression model and of a probit model, where the dependent variable is paternal unemployment. In both cases, the instrument is highly significant and the F-statistic of a test for $\delta = 0$ is close to 10.

6.5.4 Decomposition of the unemployment effect

In order to investigate how much of the overall gap in different upper secondary school choice probabilities is due to differences in the subjective probability of school success or due to other child characteristics, I use a nonlinear decomposition in the spirit of Fairlie (2005). It is the same decomposition used and described in Heckman et al. (2013b) except for nonlinear models. Assume that school choice is independent across children of employed and unemployed fathers conditional on a large vector of exogenous variables X that are not affected by paternal unemployment. The decomposition is based on the following non-linear probit model:

$$P(S_i^D = 1 | X_i^D, p_i^D) \quad = \quad \Phi(\alpha_S{}' X_i^D + \eta p_i^D), \quad with \quad D \in \{0, 1\} \tag{6.10}$$

where region and time dummies are also controlled for, but excluded from Equation 6.10 for notational simplicity. X_i^D is a vector of all background characteristics listed in Table 6.6 and p_i again denotes the subjective success probability of individual i. The above equation implies the assumption that the effect of background characteristics and

the subjective school probability is the same for individuals of employed and unemployed fathers $\alpha_S^0 = \alpha_S^1$ and $\eta^0 = \eta^1$. I test and do not reject this hypothesis.

The goal of the decomposition is to decompose the effect of paternal unemployment into components attributable to a change in the subjective school success probability. Following the notation used in Fairlie (1999) and using coefficients from a probit regression for a pooled sample, the contribution of p to the different in school choice probabilities $\Delta\bar{S}$ between children of employed ($D = 0$) and unemployed ($D = 1$) fathers can be written as:

$$\Delta\bar{S}_p = \sum_{i=1}^{N^1} \frac{1}{N^1} \Phi(\alpha_S{}' X_i^0 + \eta p_i^1) - \Phi(\alpha_S{}' X_i^0 + \eta p_i^0)^{19} \tag{6.11}$$

The contribution of each variable to the in upper secondary school choice probabilities is thus equal to the change in the average predicted probability from replacing the distribution of the subjective school success probability for children of unemployed fathers with the one of children from employed fathers. To this end both samples are matched on the basis of their rank in the distribution of school choice probabilities. Because there is a lower number of children from unemployed fathers a number of 100 repeated random subsamples of children from employed fathers are drawn and matched to compute Equation 6.11.

The percentage of the unemployment effect explained by p which cannot be accounted for by background variables can then be expressed as a percentage of the observed difference in school choice probabilities that cannot be accounted for by the exogenous variables. For most models this amounts roughly the percentage of the treatment effect explained

[19] In practice the value of $\Delta\bar{S}_p$ depends on whether background variables are fixed at $X_i^D = 0$ or X_i^1. To account for this a random ordering of variables is used.

by p. In addition I express p as a percentage of the treatment effect computed in Equation 6.7 (to be done!).

Table 6.6: Covariates in the different model equations

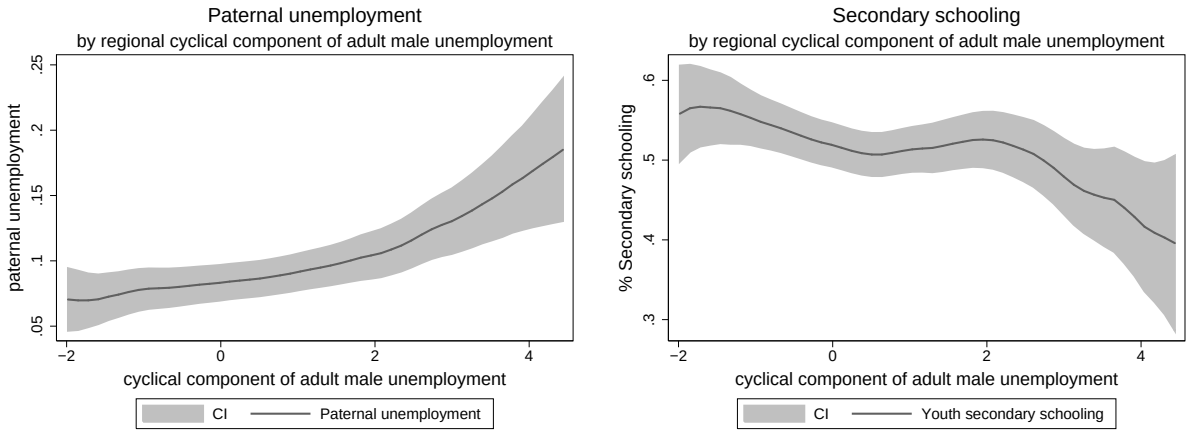
		Unemployment	Upper secondary schooling
Covariates	Constant	✓	✓
	Paternal age	✓	✓
	Maternal age		✓
	2 children	✓	
	3 children	✓	
	4 or more children	✓	
	1 sibling		✓
	2 children		✓
	3 or more siblings		✓
	Father secondary intermediate school	✓	✓
	Father grammar school	✓	✓
	Mother secondary intermediate school	✓	✓
	Mother grammar school	✓	✓
	Father German	✓	
	Large city	✓	✓
	Medium city	✓	✓
	Small city	✓	✓
	Permanent unemployment component		✓
	Father industry dummies	✓	✓
	Youth sex (male=1)		✓
	Youth German		✓
	Youth track recommendation		✓
	Father cognitive ability	(✓)	(✓)
	Vocational training positions per applicant in region		(✓)
	Youth unemployment in region		(✓)
	Subjective school success probability		(✓)
	Year FEs	✓	✓
	Federal state FEs	✓	✓
	Region FEs	(✓)	(✓)

Note: See table 5.3 for details.

6.6 Empirical Results

The results are presented and discussed in several stages. I first provide a description of the main findings in Section 6.6.1, including the reduced form effect, the causal effect of paternal unemployment and a comparison of the secondary school choice results with results for child GPA. Section 6.6.2 elicits further results displaying the heterogeneity of the paternal unemployment effect for different groups of children and families. Section 6.6.3 presents the impact of paternal unemployment on child traits and the subjective school success probability, as well as the effect of their effect on upper secondary school choice. Section 6.6.4 elaborates on some robustness checks.

Figure 6.2: Paternal unemployment and youth upper secondary schooling by regional cyclical component of adult male unemployment.



(a) Paternal unemployment **(b)** Youth upper secondary schooling

Notes: Graphs display kernel-weighted local polynomial regression outputs of paternal unemployment on the regional component of adult male unemployment. Smoothing is obtained from Epanechnikov Kernel weighted local polynomial estimates. Bandwidth selection follows Silverman's rule of thumb (Silverman, 1986). Shaded area displays 95% confidence bands.

6.6.1 The effect of paternal unemployment on education choices and GPA

If labor market fluctuations influence secondary school choice via paternal unemployment, one would expect to find an association between the cyclical component in regional adult male unemployment and child upper secondary school choice. Columns (1) and (2) of Table 6.7 show that, after controlling for background variables, the linear reduced form effect of an increase in the unemployment rate of the paternal labor market leads to a reduction in the probability that a child chooses upper secondary schooling by a little more than two percentage points.[20] This effect remains strong even after controlling for youth unemployment and the number of vocational training positions per applicant in the region.[21] The effect of unemployment fluctuations in the maternal labor market, on the other hand, displayed in columns (3) and (4), is close to zero and insignificant. The reason for why child school choice and unemployment in the maternal labor market are unrelated after controlling for background variables is probably that (a) female employment decisions are correlated less with local labor market developments, and (b)

[20]See also Figure 6.2 for a graph of the raw association between regional unemployment fluctuations and paternal unemployment and child schooling decisions, respectively.

[21]This strong reduced form effect is consistent with findings by Rampino and Taylor (2012) who report that changes in the unemployment rate have an effect on youth attitudes towards schooling in Britain.

the association between maternal unemployment and child schooling decisions is lower than for paternal unemployment.

Table 6.7: Reduced form regressions

reduced form	**Upper Secondary Education**			
Cyclical component of adult male unemployment	-0.0228**	-0.0237**		
	(0.011)	(0.011)		
Cyclical component of adult female unemployment			0.0005	0.0005
			(0.000)	(0.000)
Observations	2326	2326	2326	2326
Covariates included	YES	YES	YES	YES
Labor market controls included	NO	YES	NO	YES
R-squ adj./Ps R-squ.	0.202	0.202	0.201	0.201

Standard errors in parentheses
Source: GSOEP Youth Sample.
Note: Standard errors are robust. Covariates are sibling dummies, parental education, size of region, nationality, year FEs, region FEs.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Going beyond the reduced form estimation, I use a simulatenous equation bivariate probit model to assess the causal effects of paternal unemployment on child upper secondary school choice. The first row of Table 6.8 displays the average marginal effect of paternal unemployment on child upper secondary school choice for different model specifications. Paternal unemployment reduces the probability of upper secondary school choice by around 18 percentage points. This effect seems very robust even after including youth labor market and vocational training measures, paternal cognitive skills and several interaction effects. The latter comprise cross terms between paternal unemployment and paternal schooling, paternal cognitive skills and paternal age on the one hand, and between paternal labor market fluctuations and the three paternal traits, displayed in lines 2-4 on the other hand. The table also displays the estimated marginal effect of unemployment fluctuations in the second equation of model (5.1). It shows that a one percentage point increase in the cyclical component of adult male unemployment translates also into a one percentage points higher unemployment probability among the fathers in my sample. Note that the LR-test does not reject the null hypothesis of $\rho = 0$ for any of the models displayed in the table. Hence, marginal effects of a restricted model are presented.[22]

[22]Estimates that restrict $\rho = 0$ are much lower in size than for the unrestricted model. Compare Table 6.18.

Table 6.8: Bivariate probit results of the effect of paternal unemployment on youth higher education decision

Upper secondary education	**1**	**2**	**3**	**4**	**5**
Schooling equation					
Father unemployed	-0.1899***	-0.1896***	-0.1836***	-0.1938***	-0.1852***
	(0.036)	(0.037)	(0.038)	(0.040)	(0.037)
Father grammar school	0.3245***	0.3253***	0.3050***	0.3098***	0.3209***
	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)
Father cognitive skills			0.07373***	0.07440***	
			(0.018)	(0.018)	
Paternal age	0.004277*	0.004277*	0.003641	0.002527	0.003294
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Unemployment equation					
Cyclical component of adult male unemployment	0.01130**	0.01130**	0.01077**	0.01283**	0.009684*
	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)
Observations	2326	2326	2326	2326	2326
P-val LRtest of rho=0	0.48	0.52	0.94	0.01	0.50
Covariates included	YES	YES	YES	YES	YES
Labor market controls included	NO	YES	YES	YES	YES
Father cognitive Skills included	NO	NO	YES	YES	NO
Interaction effects	NO	NO	NO	YES	NO
Fixed Effects	state, time	state, time	state, time	state, time	region, time
Sample	All	All	All	All	All
log-lik	-1,900.67	-1,900.33	-1,870.71	-1,862.44	-1,785.39

Standard errors in parentheses
Source: GSOEP Youth Sample.
Note: Standard errors clustered by region and bootstrapped using 200 replications. For all covariates included see Table 6.6.
Interaction effects are interactions between paternal education/cognitive ability/age and unemployment/local labor market variation, respectively. Biprobit coefficients displayed are average marginal effects for the probability Pr(Upper secondary education = 1) and Pr(Paternal unemployment = 1) (for regional unemployment variation) are reported.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6.9 repeats the analysis of Table 6.8 for child GPA as a dependent variable, using the linear IV estimator and OLS.[23] The IV results suggest that paternal unemployment does not have an impact on child GPA. The 2SLS estimates are insignificant, and the coefficient of the OLS estimator even changes signs. At first sight these results deviate from the findings of Rege et al. (2011). Note, however, that, while I use grades in the same year as paternal unemployment, Rege et al. (2011) look at school grades around three years after paternal job loss.

Nevertheless, this result is striking because it indicates that paternal unemployment affects child upper secondary schooling decisions but not child school performance or abilities. In the context of the theoretical framework described in Section 6.3, this also indicates

[23] Again, a Durbin-Wu-Hausman test does not reject the H_0 of paternal unemployment being exogenous.

Table 6.9: The effect of paternal unemployment on youth GPA

GPA	2SLS		OLS
Treatment			
Father unemployed	4.1189	6.9192	-0.07813
	(3.287)	(5.995)	(0.073)
Vocational training positions per 100 applicants		0.05034	0.02031**
		(0.034)	(0.010)
Youth unemployment		-0.1298	-0.08605**
		(0.103)	(0.036)
Observations	2302	2302	2302
Covariates included	YES	YES	YES
Labor market controls included	NO	YES	YES
Father cognitive Skills			
P-val DWH-test (H0: unemp exogenous)	0.094	0.016	

Standard errors in parentheses
Source: GSOEP Youth Sample.
Note: For covariates included see Table 6.6.
GPA ranges from 6 (best grade) to 1 (worst grade).
German grades have been reversed.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

that paternal joblessness does not have an immediate effect on potential wages in either of the two education sectors, because wages are driven by child abilities and GPA is an indicator of ability and subsequent wages (Rose and Betts, 2004).

6.6.2 Heterogeneity of effects across different groups of individuals

For policy makers it is important to know which children are particularly vulnerable to paternal unemployment. This section first presents results generated by stratifying the sample on different parent and child characteristics. Then, I use my model results to show the degree of heterogeneity in the unemployment and labor market effects for fathers of different ages and cognitive abilities.

Regional labor market fluctuations only affect individuals who work in a respective region. Therefore, it is interesting to investigate this in order to see whether fathers who are willing to commute larger distances may be more likely to find a new job quickly and may be less affected by local labor market fluctuations. Column (1) of Table 6.10 indicates that the unemployment effect stays roughly constant after excluding those fathers from the sample who commuted more than 30km to work in the past year. This effect may be caused by similar unemployment fluctuations in neighboring regions. Section 6.2.4 explained that schools in the German school system usually belong to one of three different tracks, where only the highest track provides automatic access to an upper secondary school

degree. Hence, ex ante one may expect that paternal unemployment is more harmful for individuals with a lower track recommendation, who are more likely to attend the lower tracks where access to upper secondary schooling is not automatic. Columns (2) and (3) show that the marginal effect of paternal unemployment for individuals attending the highest school track is larger in absolute terms than for the middle track (recall that individuals of the lowest track are not part of the sample). However, note that the marginal effects are expressed in terms of percentage points. Regarding percentages the effect for individuals who had a low track recommendation (38 percent) is indeed significantly larger than for individuals with a high track recommendation (23 percent). Stratifying across gender in Columns (4) and (5) indicates that daughters are more prone to dropping out in response to paternal unemployment than boys. It may be that parents are more willing to cushion shocks towards their sons rather than daughters. What is also likely is that females lose confidence more easily in response to a family shock than boys. However, the difference in the effect between girls and boys is not significant. Due to institutional changes in the regulation and amount of unemployment benefits, unemployment hardship has increased in Germany from 2005 onwards. Hence, it is not surprising that the adverse effect of paternal unemployment is larger after the reform (Column (7)) than before (Column (6)). Again, the difference in the effects between Column (6) and Column (7) is not significant.

In late 2008 the most recent economic crisis started and if the exclusion restriction is violated this is most likely the case for the crisis years when newspapers were full of bad news about the economy. Hence, in Column (8) I reestimate the model also for all years except the crisis years. Again, the point estimate is slightly larger but not significantly different from the point estimate in Column (1).[24]

6.6.3 The role of the subjective school success probability

In Section 6.3 I argued on theoretical grounds that the subjective school success probability is an important channel through which paternal unemployment affects child upper secondary school choices. For this to be true, three circumstances need to hold: First, paternal unemployment has to have a causal effect on the success probability. Second,

[24]Heterogeneity results for fathers with high/low cognitive ability, fathers with upper secondary schooling or below and fathers of different ages are displayed in the appendix.

the subjective probability of school success has to be a predictor for upper secondary school choice. Third, the fraction of the paternal unemployment effect explained by this probability has to be sufficiently large. This section will explore all three of these conditions in turn.

Table 6.10: Bivariate probit results of the effect of paternal unemployment on youth higher education decision, different subsamples

Upper secondary education	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Schooling equation								
Father unemployed	-0.1878***	-0.1921***	-0.1141***	-0.1551***	-0.2181***	-0.1819***	-0.2166***	-0.1976***
	(0.037)	(0.052)	(0.036)	(0.054)	(0.048)	(0.037)	(0.066)	(0.035)
Observations	1969	990	1336	1171	1155	1407	919	1856
P-val LRtest of rho=0	0.563	0.000	0.800	0.625	0.150	0.871	0.089	0.603
Covariates included	YES	YES	YES	YES	YES	YES	YES	YES
Labor market controls included	YES	YES	YES	YES	YES	YES	YES	YES
Fixed Effects	state, time	state, time	state, time	state, time	state, time	state, time	state, time	state, time
Sample	No commuters	High Track	Low track	Males	Females	<2005	>=2005	<2008
log-lik	-1,646.846	-520.285	-1,106.998	-946.777	-922.239	-1,149.317	-712.632	-1,529.312

Standard errors in parentheses

Source: GSOEP Youth Sample.

Note: Standard errors clustered by region and bootstrapped using 200 replications.
For covariates included see Table 6.6.
Biprobit coefficients displayed are average marginal effects for the
probability Pr(Upper secondary education = 1). Non-commuters are fathers who in (t-2) commuted
less than 30km to work. High track denotes
individuals who are in the highest secondary school track (Gymnasium). Low track are individuals
in the low secondary school tracks, that is at schools which do not offer an upper secondary
school degree (Hauptschule and Realschule).
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

I again use the cyclical component in an adult male unemployment as an instrumental variable for paternal unemployment to investigate the effect on the probability of school success and on the binary indicator $\mathbb{1}[Percentage > 50]$ using 2SLS and probit models respectively. Table 6.11 shows that the effect of paternal unemployment on that probability is significant and large in absolute terms. The estimate generated by the linear IV estimator in Column (1) indicates that the causal effect of paternal unemployment on the subjective school probability is a non-significant reduction of 11 percentage points of this probability. With a binary probit, I find that the probability of finding it rather likely to graduate is significantly reduced by 6.3 percentage points.

Table 6.11: The effect of paternal unemployment on subjective school success

Probability of school success	**2SLS**	**Biprobit**
Treatment		
Father unemployed	-11.109	-0.06333**
	(26.025)	(0.032)
Observations	2326	2326
P-val LRtest of rho=0		0.342
Covariates included	YES	YES
P-val DWH-test (H0: unemp exogenous)	0.784	

Standard errors in parentheses
Source: GSOEP Youth Sample.
Note: For covariates included see Table 6.6.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Except for the subjective school success probability, other child traits and preferences are also likely to be affected by paternal unemployment. Table 6.12 displays the effect of paternal unemployment on child locus of control, child risk aversion and child mental health. The table shows that paternal unemployment reduces child mental health and child locus of control by a little over one standard deviation. The coefficient on risk aversion is positive but strongly insignificant.

After having shown that paternal unemployment has a large and significant negative effect on the subjective school success probability, I investigate whether that probability also has effects on child schooling decisions by including it as an additional covariate into the upper secondary school equation of Equation (6.7). The results are reported in Table 6.13. Column (1) includes the subjective school probability as a linear measure and Column (2) as a binary indicator for $\mathbb{1}[Percentage > 50]$. Only the marginal effect reported in Column (2) is significant and equals 0.07. Thus, if an individual thinks she is rather likely to succeed at school her probability to opt for upper secondary schooling increases by 7

Table 6.12: Child preferences and traits through which paternal unemployment can affect youth higher education decisions (IV-2SLS estimator)

Child traits	**Mental Health**	**Risk aversion**	**Locus of control**
Treatment			
Father unemployed	-1.0312*	0.2146	-1.1190*
	(0.565)	(0.759)	(0.612)
Observations	2115	1909	2085
Covariates included	YES	YES	YES

Standard errors in parentheses
Source: GSOEP Youth Sample.
Note: Standard errors clustered by region.
The analytical sample on which these estimates are based
consists of all GSOEP youths that have no missings in any of the covariates.
For covariates included see Table 6.6.
Trait measures are standardized with mean zero and standard deviation one.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

percentage points. The finding that only the the marginal effect in Column (2) is significant suggests that the effect of that probability is nonlinear and that individuals need to pass a certain threshold to choose higher schooling.[25]. This result is easily explained on hands of the model presented in Section 6.3. Departing from Equation (6.3), individuals will choose upper secondary education if it holds that

$$\sum_{t=t_e}^{T} \delta^t \mathrm{E}\big[w^1(t)\big] - \sum_{t=0}^{T} \delta^t \mathrm{E}\big[w^0(t)\big] > 0.$$

Solving this inequality for p gives:

$$p > \frac{\sum_{t=0}^{t_e} \delta^t w^l}{\sum_{t=t_e}^{T} \delta^t \left(w^h(t) - w^l(t)\right)} \tag{6.12}$$

$$p > \frac{\text{foregone earnings}}{\text{gain from education}}$$

Hence, the effect of the subjective probability in shifting an individual into upper secondary education is nonlinear. Individuals will choose upper secondary education only if that probability passes a certain threshold. Moreover, if individuals use the market interest rate to discount future earnings and have roughly equal wage streams in the two sectors, this threshold should be similar across individuals.

[25]Robustness checks show that this threshold is somewhere between 50% and 60%. Unfortunately, the coding of the variable in steps of 10 does not allow a more precise analysis.

Table 6.13: Bivariate probit results of the effect of subjective school success on youth higher education decision

Upper secondary education	**1**	**2**
Probability, successful school completion	0.0001189	0.07259***
	(0.000)	(0.025)
Schooling equation		
Father unemployed	-0.1908***	-0.1862***
	(0.038)	(0.038)
Unemployment equation		
Cyclical component of adult male unemployment	0.01130**	0.01130**
	(0.005)	(0.005)
Observations	2326	2326
P-val LRtest of rho=0	0.522	0.604
Covariates included	YES	YES
Labor market controls included	YES	YES
Measure of school success probability	linear	binary (>50%)
Fixed Effects	state, time	state, time
Sample	All	All
log-lik	-1,900.296	-1,896.733

Standard errors in parentheses
Source: GSOEP Youth Sample.
Note: Standard errors clustered by region and bootstrapped using 200 replications. For all covariates included see Table 6.6.
Biprobit coefficients displayed are marginal effects.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

After having shown that paternal unemployment has a large and significant effect on the subjective school success probability and that this probability also affects upper secondary school choice, it is interesting to see what part of the overall unemployment effect can be explained by that probability. I use the decomposition laid out in Section 6.5.4 and investigate how much of the difference in the unemployment effect can be ascribed to the probability that p is larger than 50 percent and to other child characteristics, after controlling for a large number of background variables.[26] Using the decomposition, paternal unemployment is estimated to reduce upper secondary schooling by 20 percentage points as can be seen from Table 6.14. Moreover, 2% of that reduction can be explained by the probability that p is larger than 50 percent. In Column (2) I add locus of control as an additional indicator of the school success probability. Hence, I make the assumption that locus of control is another measure of that same probability, because it evaluates whether an individual thinks that she can affect future wages by e.g. choosing education (Coleman and DeLeire, 2003). Doing so increases the explained part through upper secondary school success to 7.5 percent. Column (3) reveals that also adding mental

[26] All background variables of Column 2 in Table 6.6 are included.

health and risk aversion does further increase this explained part. Hence, both child mental health and risk aversion are not associated with a higher probability of upper secondary school choice.

Table 6.14: Decomposition of the probability of child upper secondary education into subjective success probability, traits and background variables: fractions of overall effect ascribed to investments and parental background

Decomposition	**(1)**	**(2)**	**(3)**
SuccessProb	0.004177**	0.01445***	0.01559***
	(0.002)	(0.003)	(0.004)
Observations	1728	1728	1728
N(treated)	1,568.000	1,568.000	1,568.000
N(control)	160.000	160.000	160.000
Unemployment effect	0.201	0.206	0.206
%of unemployment effect explained by p	**2.081**	**7.019**	**7.553**
Variables included	Success prob	adding Loc	adding risk, mhealth

Standard errors in parentheses
Source: GSOEP Youth Sample.
Note: Observations are weighted using coefficients of a pooled probit model.
Randomized ordering of variables used. Estimates based on 100 replications.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6.6.4 Additional sensitivity analyses

In addition to the sensitivity analyses discussed so far, I have performed a range of further estimations to assess the robustness of the results with respect to the choice of covariates, outcome definition, coding the treatment and estimation strategy. First, much of my analysis hinges on the assumption that fluctuations in the regional unemployment rate only influence child schooling decisions via their effect on paternal unemployment. Admittedly, local labor market conditions may impact educational choices in ways that may not be transmitted solely through the father's unemployment experience. Regional fluctuations in GDP may also affect the general wage development in that region or a decline in tax revenue may affect school operations. Therefore, I conducted robustness checks where, in the first equation of the bivariate probit model, I also control for mean deviations in regional tax income and for mean deviations in the general wage development. Both of these variables do not affect my main coefficient estimates and do not significantly affect upper secondary school choice.

Second, I use a different coding of the outcome variable. While in my main estimations I use final level of schooling if observed and planned level of schooling if the final level is not yet observed, Table 6.17 shows the main results of this analysis when using a binary

indicator for whether an individual is still in school at age 17. In this case, the results are still significant, but the marginal effect size reduces by about 5 percentage points. I focus on final schooling for two reasons. First, an indicator for whether a child is still in school at age 17 is likely to be flawed for individuals who started school at a different age or had to repeat a grade. Second, final schooling is a much better predictor of later outcomes.
Third, I estimate the effect of involuntary unemployment instead of non-employment on child education decisions. The results are displayed in Table 6.19, and the marginal effects obtained are extremely similar to the ones in Table 6.8.
Last, I use linear 2SLS-IV and OLS instead of nonlinear probit estimators. Using 2SLS entails considerable increase in effect sizes as can be seen in Table 6.16. Note that the IV-estimator provides a weighted local average treatment effect for children of fathers who are shifted into unemployment due to a regional labor market downturn, while the bivariate probit results are average marginal effects. Hence, a straightforward explanation for the larger effect size is treatment heterogeneity and that children of individuals affected by a regional downturn (*compliers*) are worse off when compared to children of individuals who do not experience a change in employment in times of recession (*never-takers* and *always-takers*).[27]

6.7 Conclusion

This chapter shows that paternal unemployment has a large and significant effect on child upper secondary school choice and that the subjective probability of successful school completion is a driving mechanism behind this effect. Moreover, the study identifies heterogeneity in the paternal unemployment effect for different groups of children and fathers. I estimate a simultaneous equation latent variable model for the joint probability of child upper secondary school choice and paternal unemployment using regional variation in the cyclical component of adult male unemployment in the labor market of the father as an exogenous shifter for paternal unemployment. To interpret my findings and to link them to the theory of human capital investment decisions, I present a simple theoretical framework that explains how paternal unemployment affects schooling decisions by means of the perceived school success probability. Within this framework, young individuals make

[27] 2SLS is also biased in small samples, such that part of the increase may be ascribed to that bias (Chiburis et al., 2012).

upper secondary schooling decisions by comparing discounted expected wage flows for each schooling choice.

Paternal unemployment reduces the probability of upper secondary school choice by roughly 18 percentage points or 34 percent. Paternal unemployment reduces the probability that an individual finds it rather likely that she will graduate successfully by 7 percentage points. It also reduces child locus of control and child mental health by roughly one standard deviation. The theoretical framework that motivates my analysis predicts that the subjective school success probability has a nonlinear effect on child upper secondary school choice, and the empirical analysis confirms this presumption. Overall, the subjective school success probability explains about 2-7.5 percent of the overall gap in different upper secondary school choice probabilities between employed and unemployed fathers.

Some of the results are specific to the German institutional system. First, in percentage terms, children who visit schools of the lower secondary school tracks are more affected from paternal unemployment than children who visit higher-track schools. Second, unemployment has a more detrimental effect after a substantive reform of the unemployment benefit system as of January 2005, the so-called Hartz IV reform, although the difference is not significant.

Given the finding that regional labor market downturns are an important driver of child education decisions via their effect on paternal unemployment, this chapter contributes to the discussion on second order effects of economic crises. Using the structure of my model to predict the effect of a recession on the marginal probability of upper secondary school choice, I find that a labor market downturn that is similar in size to that in the US during the Great Recession leads to a reduction in the upper secondary school choice probability by 2 percentage points.

My finding that paternal unemployment has adverse effects on child outcomes confirms earlier findings by Rege et al. (2011); Oreopoulos et al. (2008); Kalil and Ziol-Guest (2008) as well as Stevens and Schaller (2011). Yet, this chapter differs substantially from these other studies in the literature focusing primarily on the psychological and behavioral impacts of paternal unemployment on the child rather than on paternal investments. The children of unemployed fathers in this study are substantially older than in any of the above-named papers. At age 16, cognitive skill production is largely completed which

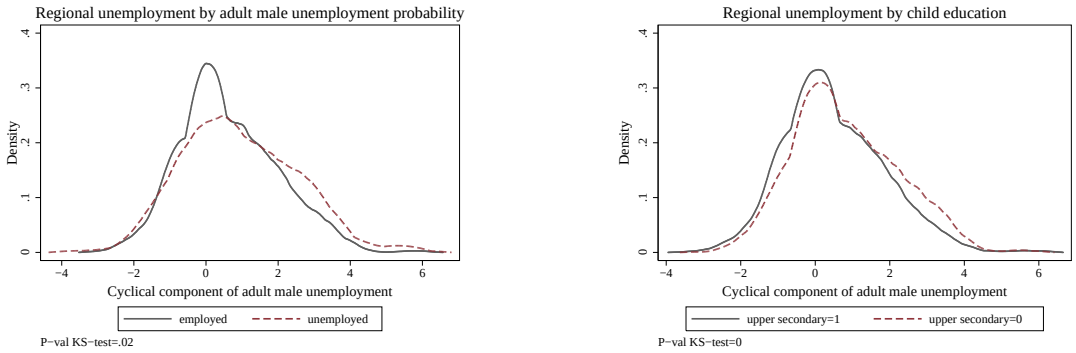
explains why in contrast to Rege et al. (2011) I find that child GPA is not affected by paternal unemployment.

This chapter shows that economic crises can have very important second order effects in terms of education outcomes of the next generation. From a policy perspective, this is relevant because it shows that part of the negative effect of paternal unemployment on education decisions can be mitigated by policy interventions focused on changing self-confidence and expectations about the future. Such policies are likely to be more cost-effective than seeking to directly improve cognitive abilities. Arguably, this chapter is limited in scope. First, it focuses on teenage children only, and second it only investigates the impact of paternal unemployment on child secondary school choice. The present chapter may, therefore, be seen as a motivation to construct models, which allow for different channels and mechanisms linking familial distress to human capital investment decisions.

6.A The Effect of Labor Market Fluctuations

I also consider the impact of regional labor market downturns. Figure 6.3 displays the association between regional unemployment deviations and unemployment or education by showing kernel densities of the cyclical component of adult male unemployment by education and employment, respectively. Unsurprisingly, densities for unemployed fathers and children without upper secondary schooling are shifted to the right. The overall effect of fluctuations in the cyclical unemployment component as predicted from the model can be seen in Figure 6.4. The gradient of the line shows the degree to which regional unemployment influences child upper secondary schooling decisions via the effect on paternal unemployment. Note that this gradient is surprisingly steep indicating that economic crises have considerable second order effects on next generation schooling choices.

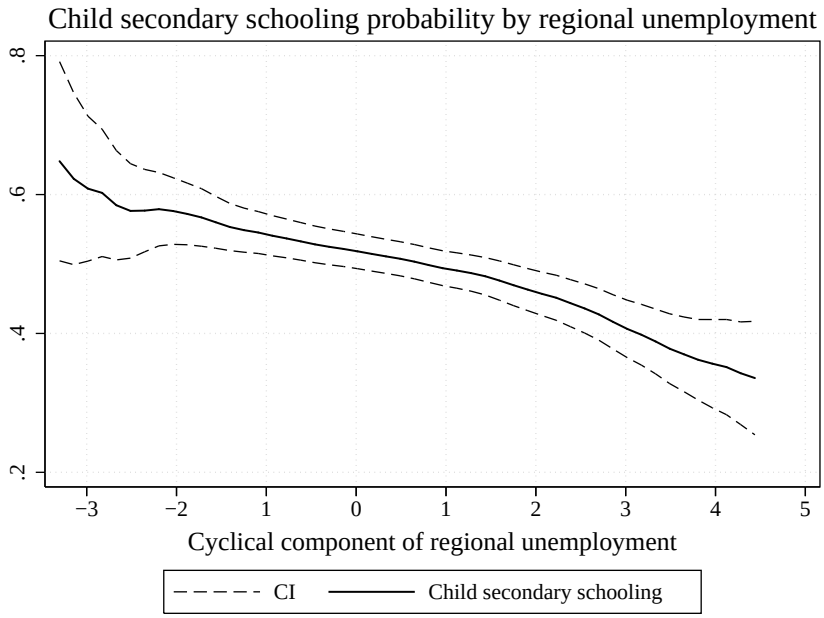
Figure 6.3: Cyclical component of adult male unemployment by paternal unemployment and youth upper secondary education.



(a) Paternal unemployment **(b)** Youth higher education

Notes: Model simulation results. Estimates of model (4) in Table 6.8. Kernel density estimation implemented using a Gaussian kernel with bandwidth selected using Silverman's rule of thumb (Silverman, 1986). Kolmogorov-Smirnov test: Two-sample KS-test with null hypothesis that the two distributions are the same. p-values reported underneath graphs.

Figure 6.4: Effect of cyclical unemployment fluctuations in paternal labor market on youth upper secondary schooling.

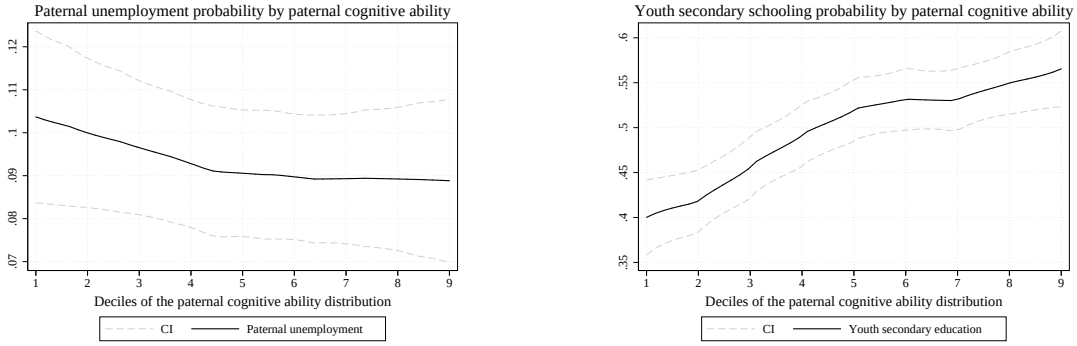


Notes: Model simulation results. Simulation based on estimates of model (4) in Table 6.8. 95% pointwise confidence interval between dashed lines.

6.B The Role of Paternal Characteristics

The analysis in the main part of the chapter investigates average marginal effects for different models and different strata of the sample. Model (4) contains interaction effects for different paternal characteristics, which are important when thinking about the intergenerational transmission of disadvantage in response to labor market shocks. When using interaction terms in nonlinear models, coefficients are hard to interpret. Therefore, I use the structure of the model and the interaction terms to predict probabilities for three different groups of fathers: Fathers with high/low cognitive ability, fathers with upper secondary schooling or below and fathers of different ages.[28] The main prediction results are based on coefficients of model (4) in Table 6.8. Figure 6.5 and Figure 6.6 show the predicted unemployment and education probabilities for fathers with different cognitive abilities and ages. Even after conditioning on schooling, higher paternal cognitive skills significantly reduce the unemployment probability and significantly increase the probability of child upper secondary schooling. If a father could be moved from the lowest

Figure 6.5: Effect of paternal cognitive ability (after conditioning on education) on paternal unemployment and youth upper secondary education.



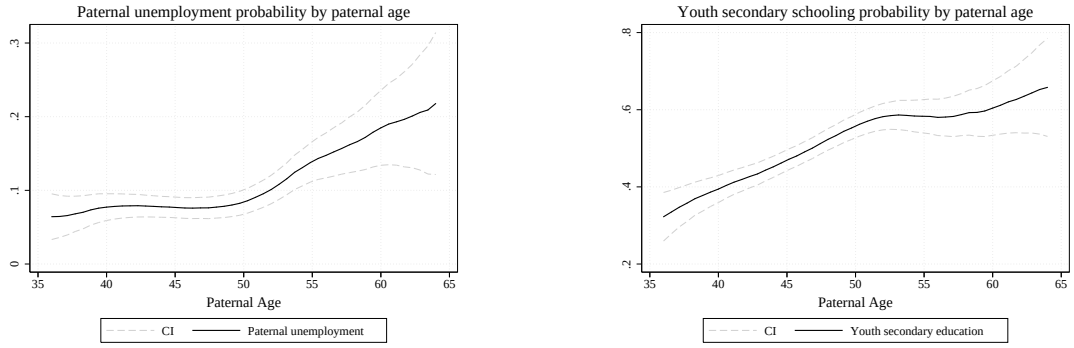
(a) Paternal unemployment **(b)** Youth higher education

Notes: Model simulation results. Simulations based on estimates of model (4) in Table 6.8. 95% pointwise confidence interval between dashed lines.

to the highest decile of the cognitive ability distribution, his child would be 15 percentage points more likely to choose upper secondary schooling. Paternal age also has a positive effect on child schooling decisions but a negative effect on the employment probability. Unsurprisingly, Figure 6.7 shows that individuals with upper secondary schooling are

[28] Goodness-of-fit statistics for using the model to predict outcomes are presented in Table 6.20.

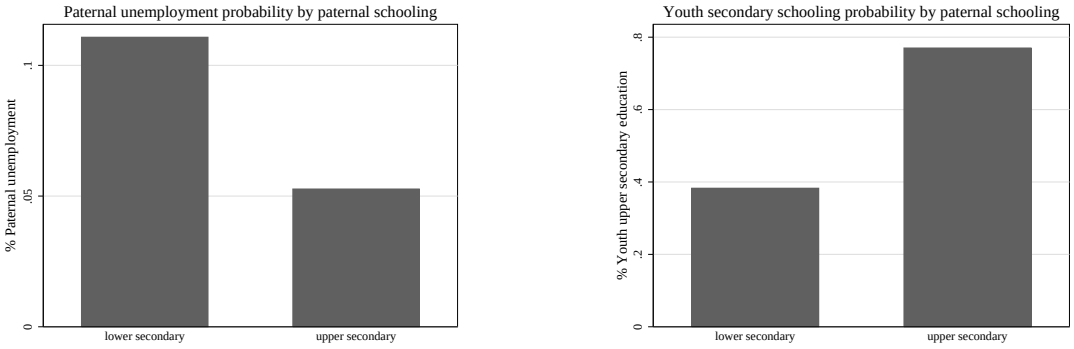
Figure 6.6: Effect of paternal age on paternal unemployment and youth upper secondary education.



(a) Paternal unemployment **(b)** Youth higher education

Notes: Model simulation results. Simulations based on estimates of model (4) in Table 6.8. 95% pointwise confidence interval between dashed lines.

Figure 6.7: Effect of paternal upper secondary education on paternal unemployment and youth upper secondary education.



(a) Paternal unemployment **(b)** Youth higher education

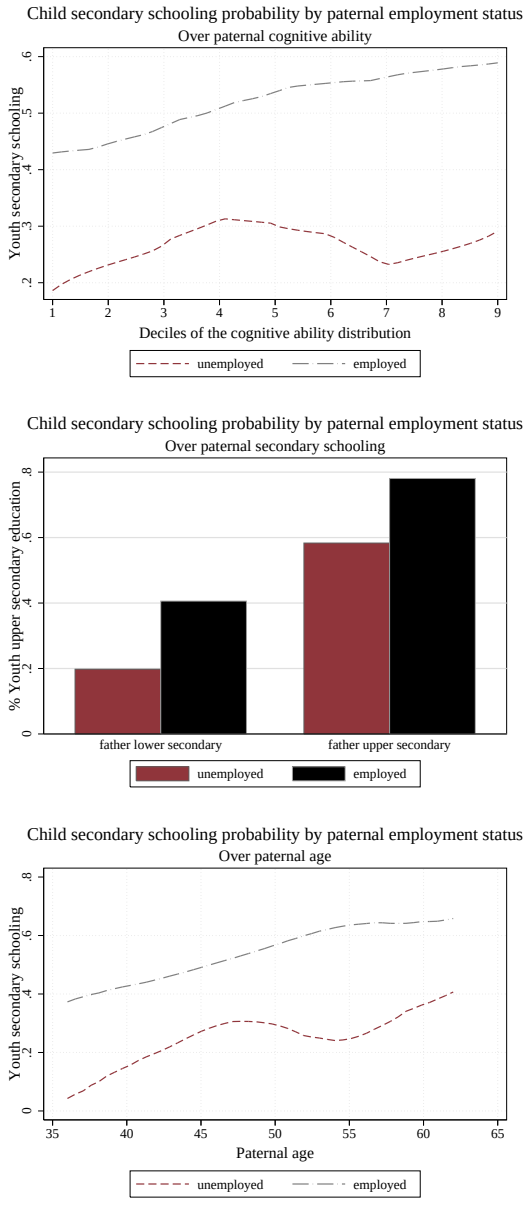
Notes: Model simulation results. Simulations based on estimates of model (4) in Table 6.8.

less likely to become unemployed while their children are more likely to choose upper secondary education.

Figure 6.8 investigates the overall effect of paternal unemployment for fathers with high/low cognitive abilities, with high and low education levels and with different ages. For each of these graphs, a move along the x-axis leads to a widening in the difference of upper secondary school probabilities between employed and unemployed fathers. Hence, the effect of unemployment tends to be somewhat more detrimental for fathers with high cognitive abilities, high education and higher age. This seems surprising at first, given that families with higher ability endowments should be able to cushion shocks more easily.

However, at the same time, the graphs display the average treatment effect of individuals with different endowments. Hence, given that the event of unemployment is very unlikely for fathers with high cognitive abilities, it may simultaneously be psychologically more detrimental to these individuals.

Figure 6.8: Heterogenous effect of paternal unemployment on child upper secondary schooling by paternal cognitive ability, education and age

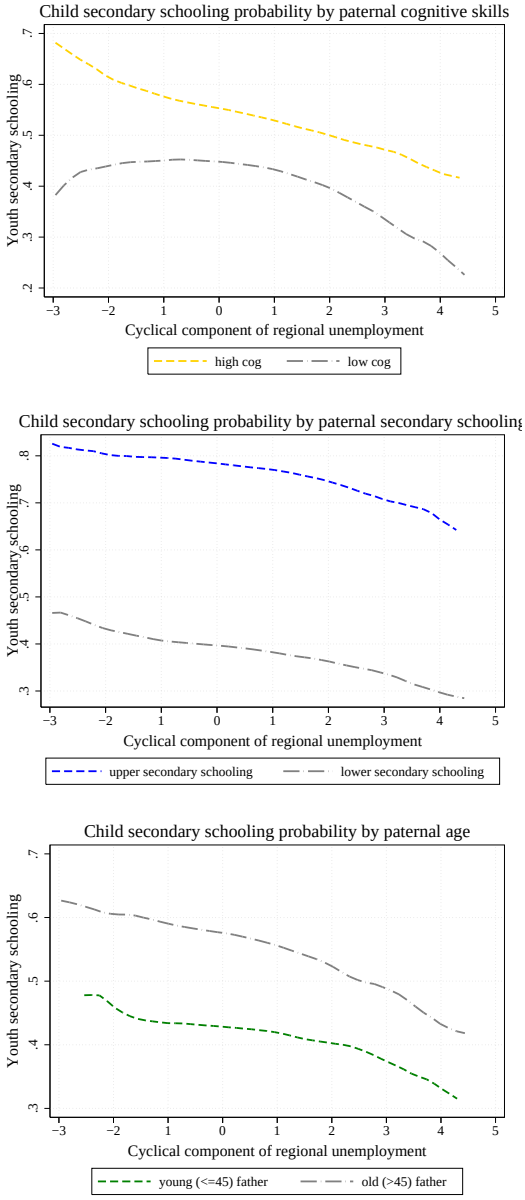


Notes: Model simulation results. Simulation based on estimates of model (4) in Table 6.8.

I also use the model to predict the effects of a recession where unemployment increases in all regions by 4 percentage points. This is about the effect that the most recent crisis had on US unemployment rates. When predicting probability of upper secondary school success for different levels in regional unemployment rates, I find that overall there is

not much heterogeneity in the response of child upper secondary schooling to regional labor market downturns for fathers with different characteristics. Panel 1 of Figure 6.9 shows that children of fathers with an upper secondary school degree suffer slightly more from a regional labor market downturn than children from fathers with lower education. Moreover, children of older fathers are slightly less affected by labor market downturns (Panel 3 of Figure 6.9).

Figure 6.9: Heterogenous effect of regional cyclical unemployment fluctuations on child upper secondary schooling by paternal cognitive ability, education and age



Notes: Model simulation results. Simulation based on estimates of model (4) in Table 6.8.

6.B.1 Direct, indirect and total effects of paternal characteristics

From a policy perspective, it is interesting to see how the effect of regional labor market downturns and paternal unemployment varies for fathers with different observable characteristics. Different paternal characteristics may affect the probability to choose upper secondary schooling either directly or because they affect the probability of paternal unemployment. The total effect of a paternal characteristic c^j can be decomposed into a direct and an indirect effect according to:

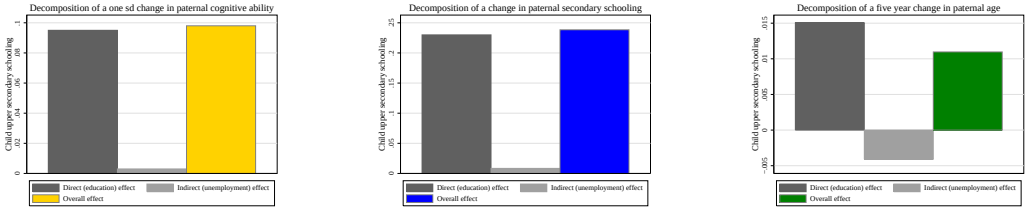
$$\overbrace{\frac{dP\left(S=1|X=x\right)}{dc^j}}^{\text{Total effect}}$$

$$= \overbrace{\sum_{D=0}^{1} P(D=d|X=x,C^j=c^j)\frac{\partial P(S=s|X=x,C^j=c^j,D=d)}{\partial c^j}}^{\text{direct effect}}$$

$$+ \overbrace{\sum_{D=0}^{1} \frac{\partial P(D=d|X=x,C^j=c^j)}{\partial c^j} P(S=s|X=x,D=d,C^j=c^j)}^{\text{indirect effect}}$$

The indirect effect represents a reduced probability to opt for upper secondary schooling induced by a change in the probability of paternal unemployment, which is induced by a change in the respective paternal characteristic. The direct effect is the part of the effect of a characteristic that is unrelated to unemployment and directly influences the education probability. The results of the decomposition of Section 6.B.1 are displayed in Figure 6.10. The graphs show that paternal cognitive abilities and paternal education have large positive direct effects on child upper secondary schooling. The effect of paternal age, on the other hand, is very small. Moreover, the indirect effect of paternal age is negative because higher paternal age leads to a reduction in the employment probability.

Figure 6.10: Decomposing the effect of a change in paternal cognitive ability, schooling and age on child upper secondary schooling decisions.



Notes: Model simulation results. Simulation based on estimates of model (4) in Table 6.8.

6.C Additional Tables

Table 6.15: Correlation: paternal/maternal change to non(un-)employment and gpa

Newly unemployed	**GPA**
Father becomes unemployed	-0.06062
	(0.113)
Father becomes involuntarily unemployed	-0.08875
	(0.126)
Mother becomes unemployed	0.07079
	(0.127)
Mother becomes involuntarily unemployed	-0.06773
	(0.141)
Observations	2098 (father) 2071 (mother)
Covariates included	NO
R-squ adj.	-0.00

Standard errors in parentheses
Source: GSOEP Youth Sample.
Note: Standard errors are robust.
Raw correlations displayed, no covariates included.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6.16: 2SLS/OLS-results of the effect of paternal unemployment on youth higher education decision

Upper secondary education	**2SLS**	**2SLS**	**2SLS**	**OLS**
Treatment				
Father unemployed	-0.5479**	-0.6279**	-0.5224**	-0.1951***
	(0.263)	(0.245)	(0.264)	(0.032)
Vocational training positions per 100 applicants			-0.003490	-0.004321
			(0.004)	(0.003)
Youth unemployment			-0.01381	-0.02057*
			(0.011)	(0.012)
Father cognitive skills		0.08385***	0.08522***	0.09054***
		(0.017)	(0.017)	(0.010)
Observations	2326	2326	2326	2326
Covariates included	YES	YES	YES	YES
Labor market controls included	NO	NO	YES	YES
Father cognitive Skills	NO	YES	YES	YES
R-squ adj.	0.160	0.165	0.191	0.226
P-val DWH-test (H0: unemp exogenous)	0.198	0.074	0.197	

Standard errors in parentheses
Source: GSOEP Youth Sample.
Note: For covariates included see Table 6.6.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6.17: Using higher secondary school at age 17 as outcome variable instead of final schooling

Upper secondary education (at 17)	**1**	**2**	**3**	**4**	**5**	**6**
Treatment						
Father unemployed	-0.1340***	-0.1051***	-0.09708*	-0.03441	-0.08829	-0.1413***
	(0.036)	(0.039)	(0.057)	(0.030)	(0.062)	(0.051)
Vocational training positions per 100 applicants	0.001466	0.0007131	0.1585	-0.002730	0.004609	-0.001726
	(0.004)	(0.005)	(0.099)	(0.004)	(0.007)	(0.006)
Youth unemployment	-0.002948	-0.01312	0.003744	-0.01320	0.002544	0.02961
	(0.015)	(0.016)	(0.006)	(0.017)	(0.025)	(0.034)
Father cognitive skills	0.04713***	0.04777***	-0.007381	0.03067*	0.06477***	0.03160*
	(0.015)	(0.016)	(0.024)	(0.016)	(0.018)	(0.019)
Observations	2326	1969	990	1336	1171	1155
P-val LRtest of rho=0	0.419	0.743	0.715	0.257	0.272	0.578
Covariates included	YES	YES	YES	YES	YES	YES
Labor market controls included	YES	YES	YES	YES	YES	YES
Father cognitive Skills	YES	YES	YES	YES	YES	YES
Sample	All	No commuters	High Track	Low track	Males	Females
log-lik	-1,881.471	-1,625.069	-601.984	-926.557	-894.890	-947.155

Standard errors in parentheses
Source: GSOEP Youth Sample.
Note: Standard errors clustered by region and bootstrapped using 500 replications.
For covariates included see Table 6.6.
Biprobit coefficients displayed are average marginal effects for the conditional probability Pr(Upper secondary education = 1).
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6.18: Binary probit results, $\rho > 0$

Upper secondary education	**1**	**2**
Treatment		
Father unemployed	-0.2965**	-0.3011*
	(0.136)	(0.158)
Observations	2326	2326
P-val LRtest of rho=0	0.475	0.499
Covariates included	YES	YES
Labor market controls included	NO	YES
Sample	All	All
log-lik	-1,785.556	-1,785.058

Standard errors in parentheses
Source: GSOEP Youth Sample.
Note: Standard errors clustered by region.
For covariates included see Table 6.6.
Biprobit coefficients displayed are average marginal effects for the conditional probability Pr(Upper secondary education = 1).
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6.19: Binary probit results using involuntary unemployment instead of unemployment as treatment variable

Upper secondary education	**1**	**2**
Schooling equation		
Father unemployed, actively looking	-0.1899***	-0.1896***
	(0.036)	(0.037)
Observations	2326	2326
P-val LRtest of rho=0	0.160	0.619
Covariates included	YES	YES
Labor market controls included	NO	YES
Sample	All	All
log-lik	-1,717.647	-1,717.303

Standard errors in parentheses
Source: GSOEP Youth Sample.
Note: Standard errors clustered by region and bootstrapped using 200 replications. For all covariates included see Table 6.6.
Interaction effects are interactions between paternal education/cognitive ability/age and unemployment/local labor market variation, respectively. Biprobit coefficients displayed are average marginal effects for the probability Pr(Upper secondary education = 1) and Pr(Paternal unemployment = 1) (for regional unemployment variation) are reported.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6.20: Model simulation, goodness-of-fit statistics

			goodness-of-fit		
	Predicted	**Actual**	**Difference**	**P-val chi2**	**Pct corr pred**
Unemployment	.0937	.0967	-.003	0	.8491
Child upper secondary education	.4966	.5181	-.0215	0	.6028

Note: Statistics for simulation of model (4) in Table 6.8 shown.

Table 6.21: Probit model results for the probability of child upper secondary education on child traits

Probit base for composition	**(1)**	**(2)**	**(3)**	**(4)**
Success Probability				
Probability, successful school completion > 50%	0.1036***	0.08325**	0.06852**	0.08247**
	(0.036)	(0.032)	(0.032)	(0.033)
Traits				
Youth mental health			-0.004553	0.003031
			(0.011)	(0.011)
Youth risk aversion			0.002757	-0.0003157
			(0.011)	(0.011)
Youth locus of control			0.04711***	
			(0.011)	
Observations	1728	1728	1728	1728
R-squared adj.	0.003	0.200	0.207	0.200
Variables included	all	all	all	excl Locus
Background variables included	YES	YES	YES	YES

Standard errors in parentheses
Source: GSOEP Youth Sample.
Note: Model serves as a basis for the decompositions displayed in Table 6.14.
Average marginal effects reported.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Bibliography

Abrevaya, J. (2006). Estimating the effect of smoking on birth outcomes using a matched panel data approach. *Journal of Applied Econometrics 21*(4), 489–519.

Achenbach, T., S. McConaughy, and C. Howell (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin 101*(2), 213.

Ahmed, F. (2010). Epigenetics: Tales of adversity. *Nature 468*(7327), S20–S20.

Aigner, D., C. Hsiao, A. Kapteyn, and T. Wansbeek (1984). Latent variable models in econometrics. *Handbook of Econometrics 2*, 1321–1393.

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

Alexander, G., J. Himes, R. Kaufman, J. Mor, and M. Kogan (1996). A united states national reference for fetal growth. *Obstetrics & Gynecology 87*(2), 163–168.

Almond, D., K. Chay, and D. Lee (2005). The Costs of Low Birth Weight. *The Quarterly Journal of Economics 120*(3), 1031–1083.

Almond, D. and B. Mazumder (2005). The 1918 influenza pandemic and subsequent health outcomes: an analysis of SIPP data. *American Economic Review 95*(2), 258–262.

Almond, D. and B. Mazumder (2011). Health capital and the prenatal environment: the effect of ramadan observance during pregnancy. *American Economic Journal-Applied Economics 3*(4), 56.

Andersen, H., A. Mühlbacher, M. Nübling, J. Schupp, and G. Wagner (2007). Computation of standard values for physical and mental health scale scores using the SOEP version of SF-12v2. *Schmollers Jahrbuch 127*(1), 171–182.

Andrisani, P. J. (1977). Internal-External Attitudes, Personal Initiative, and the Labor Market Experience of Black and White Men. *Journal of Human Resources 12*(3), 308–328.

Andrisani, P. J. (1981). Internal-External Attitudes, Sense of Efficacy, and Labor Market Experience: A Reply to Duncan and Morgan. *Journal of Human Resources 16*(4), 658–66.

Angrist, J. D., G. W. Imbens, and D. Rubin (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association 91*(434), 444–455.

Anway, M. D. (2005, June). Epigenetic Transgenerational Actions of Endocrine Disruptors and Male Fertility. *Science 308*(5727), 1466–1469.

Arellano, M. and C. Meghir (1992). Female labour supply and on-the-job search: an empirical model estimated using complementary data sets. *The Review of Economic Studies 59*(3), 537.

Astone, N. and S. McLanahan (1994). Family structure, residential mobility, and school dropout: A research note. *Demography 31*(4), 575–584.

Autor, D. and D. Dorn (2009). This job is getting old: Measuring changes in job opportunities using occupational age structure. *The American Economic Review 99*(2), 45–51.

Babson, S. and G. Benda (1976). Growth graphs for the clinical assesment of infants of varying gestational age. *The Journal of pediatrics 89*(5), 814–820.

Banning, C. (1946). Food shortage and public health, first half of 1945. *Annals of the American Academy of Political and Social Science 245*, 93–110.

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

Barker, D. (1995). Fetal origins of coronary heart disease. *British Medical Journal 311*, 171–174.

Barker, D. (2007). The origins of the developmental origins theory. *Journal of internal medicine 261*(5), 412–417.

Bartlett, M. S. (1937, July). The statistical conception of mental factors. *British Journal of Psychology 28*(1), 97–104.

Bateson, P. (2001). Fetal experience and good adult design. *International Journal of Epidemiology 30*(5), 928–934.

Becker, G. and N. Tomes (1976). Child endowments and the quantity and quality of children. *Journal of Political Economy 84*(4), 143–162.

Becker, G. S. (1964). *Human Capital: A Theoretical and Empirical Analysis, With Special Reference to Education*. New York: National Bureau of Economic Research.

Becker, G. S. (1993). *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education* (3rd ed.). University of Chicago Press.

Behrman, J. R. (1997). Intrahousehold distribution and the family. In M. R. Rosenzweig and O. Stark (Eds.), *Handbook of Population and Family Economics*, Volume 1A, pp. 125–187. New York: Elsevier Science.

Behrman, J. R., R. A. Pollak, and P. Taubman (1982). Parental preferences and provision for progeny. *Journal of Political Economy 90*(1), 52–73.

Behrman, J. R. and M. R. Rosenzweig (2004, May). Returns to birthweight. *Review of Economics and Statistics 86*(2), 586–601.

Bhargava, A. and M. Fox-Kean (2003). The effects of maternal education versus cognitive test scores on child nutrition in Kenya. *Economics & Human Biology 1*(3), 309–319.

Blau, D. (1999). The effect of income on child development. *Review of Economics and Statistics 81*(2), 261–276.

Blum, M. (2011). Government decisions before and during the first world war and the living standards in Germany during a drastic natural experiment. *Explorations in Economic History 48*(4), 556–567.

Bonellie, S., J. Chalmers, R. Gray, I. Greer, S. Jarvis, and C. Williams (2008). Centile charts for birthweight for gestational age for Scottish singleton births. *BMC pregnancy and childbirth 8*(1), 5.

Borghans, L., A. L. Duckworth, J. J. Heckman, and B. ter Weel (2008). The economics and psychology of personality traits. *Journal of Human Resources 43*(4), 972–1059.

Börsch-Supan, A., A. Brugiavini, H. Jürges, A. Kapteyn, J. Mackenbach, J. Siegrist, and G. Weber (2008). *Health, ageing and retirement in Europe (2004-2007). Starting the longitudinal dimension*. Mannheim Research Institute for the Economics of Aging (MEA).

Bowles, S. and H. Gintis (1976). *Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life*. Routledge.

Bowles, S. and H. Gintis (2002). The Inheritance of Inequality. *Journal of Economic Perspectives 16*(3), 3–30.

Bowles, S., H. Gintis, and M. A. Osborne (2001). The Determinants of Earnings: A Behavioral Approach. *Journal of Economic Literature 39*(4), 1137–1176.

Bratberg, E., Ø. Nilsen, and K. Vaage (2008). Job losses and child outcomes. *Labour Economics 15*(4), 591–603.

Bundesinstitut für Bau-, Stadt- und Raumforschung (BBSR) (2011). Inkar. indikatoren und karten zur raum- und stadtentwicklung [elektronische ressource]. CD-ROM.

Bygren, L., G. Kaati, and S. Edvinsson (2001). Longevity determined by paternal ancestors nutrition during their slow growth period. *Acta Biotheoretica 49*(1), 53–59.

Caliendo, M., D. A. Cobb-Clark, and A. Uhlendorff (2010). Locus of control and job search strategies. Institute for the Study of Labor (IZA) Discussion Paper No. 4750.

Cameron, N. and E. W. Demerath (2002). Critical periods in human growth and their relationship to diseases of aging. *Yearbook of Physical Anthropology 45*, 159–184.

Campbell, S., S. Warsof, D. Little, D. Cooper, et al. (1985). Routine ultrasound screening for the prediction of gestational age. *Obstetrics and Gynecology 65*(5), 613.

Card, D. (1999). The causal effect of education on earnings. *Handbook of labor economics 3*, 1801–1863.

Carneiro, P., C. Crawford, and A. Goodman (2007). The impact of early cognitive and non-cognitive skills on later outcomes. CEE Discussion Papers 0092, Centre for the Economics of Education, LSE.

Carneiro, P., K. Hansen, and J. J. Heckman (2003). Estimating distributions of treatment effects with an application to the returns to schooling and measurement of the effects of uncertainty on college choice. *International Economic Review 44*(2), 361–422.

Carneiro, P., J. Heckman, and E. Vytlacil (2010). Evaluating marginal policy changes and the average effect of treatment for individuals at the margin. *Econometrica 78*(1), 377–394.

Carneiro, P., J. Heckman, and E. Vytlacil (2011). Shorter papers-estimating marginal returns to education. *American Economic Review 101*(6), 2754.

Carneiro, P., C. Meghir, and M. Parey (2007). Maternal Education, Home Environments and the Development of Children and Adolescents. IFS Working Papers.

Case, A. and C. Paxson (2005). Sex differences in morbidity and mortality. *Demography 42*(2), 189–214.

Case, A. and C. Paxson (2010). Causes and consequences of early-life health. *Demography 47*, S65–S85.

Cebi, M. (2007). Locus of control and human capital investment revisited. *Journal of Human Resources 42*(4), 919–932.

Charles, K. and M. Stephens (2004). Disability, job displacement and divorce. *Journal of Labor Economics 22*(2), 489–522.

Chervenak, F., D. Skupski, R. Romero, M. Myers, M. Smith-Levitin, Z. Rosenwaks, and H. Thaler (1998). How accurate is fetal biometry in the assessment of fetal age? *American journal of obstetrics and gynecology 178*(4), 678–687.

Chiburis, R., J. Das, and M. Lokshin (2012). A practical comparison of the bivariate probit and linear iv estimators. *Economics Letters 117*(3), 762–766. Forthcoming.

Chou, S., J. Liu, M. Grossman, and T. Joyce (2010). Parental education and child health: evidence from a natural experiment in Taiwan. *American Economic Journal: Applied Economics 2*(1), 33–61.

Cole, S., L. Hawkley, J. Arevalo, and J. Cacioppo (2011). Transcript origin analysis identifies antigen-presenting cells as primary targets of socially regulated gene expression in leukocytes. *Proceedings of the National Academy of Sciences 108*(7), 3080–3085.

Coleman, J. S. (1968). Equality of Educational Opportunity. *Equity & Excellence in Education 6*(5), 19–28.

Coleman, M. and T. DeLeire (2003). An economic model of locus of control and the human capital investment decision. *Journal of Human Resources 38*(3), 701–721.

Conti, G., S. Frühwirth-Schnatter, J. J. Heckman, and R. Piatek (2012). A Framework For Understanding the Social and Economic Determinants of Adult Health. Unpublished manuscript. University of Chicago, Department of Economics.

Conti, G., J. J. Heckman, and S. Urzua (2009). Early endowments, education, and health. Unpublished manuscript, University of Chicago, Department of Economics.

Conti, G., J. J. Heckman, and S. Urzua (2010). The education-health gradient. *American Economic Review Papers and Proceedings 100*, 234–238.

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

Costa, D. L. (1993). Height, wealth, and disease among the native-born in the rural, antebellum north. *Social Science History 17*(3), 355–383.

Cowles, M. K. (1996). Accelerating Monte Carlo Markov chain convergence for cumulative-link generalized linear models. *Statistics and Computing 6*(2), 101–111.

Croon, M. (2002). Latent variable and latent structure models, chapter 10. In *Using Predicted Latent Scores in General Latent Structure Models*, pp. 195. Lawrence Erlbaum.

Cunha, F. and J. J. Heckman (2007, May). The technology of skill formation. *American Economic Review 97*(2), 31–47.

Cunha, F. and J. J. Heckman (2008). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources 43*(4), 738–782.

Cunha, F., J. J. Heckman, and S. Navarro (2005). Separating uncertainty from heterogeneity in life cycle earnings. *Oxford Economic Papers 57*(2), 191–261. The 2004 Hicks Lecture.

Cunha, F., J. J. Heckman, and S. M. Schennach (2010, May). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica 78*(3), 883–931.

Currie, J. (2009). Healthy, Wealthy, and Wise: Socioeconomic status, poor health in childhood, and human capital development. *Journal of Economic Literature 47*(1), 87–122.

Currie, J. and R. Hyson (1999). Is the impact of health shocks cushioned by socioeconomic status? The case of low birthweight. *The American Economic Review 89*(2), 245–250.

Currie, J. and E. Moretti (2003). Mother's Education and the Intergenerational Transmission of Human Capital: Evidence From College Openings. *Quarterly Journal of Economics 118*(4), 1495–1532.

Currie, J. and E. Moretti (2007). Biology as destiny? Short-and long-run determinants of intergenerational transmission of birth weight. *Journal of Labor Economics 25*(2), 231–264.

Dagsvik, J., T. Hægeland, and A. Raknerud (2011). Estimating the returns to schooling: a likelihood approach based on normal mixtures. *Journal of Applied Econometrics 26*(4), 613–640.

Dahl, G. and L. Lochner (2012). The impact of family income on child achievement: Evidence from the earned income tax credit. *The American Economic Review 102*(5), 1927–1956.

Datar, A., M. Kilburn, and D. Loughran (2010). Endowments and parental investments in infancy and early childhood. *Demography 47*(1), 145–162.

De Rooij, S., H. Wouters, J. Yonker, R. Painter, and T. Roseboom (2010). Prenatal undernutrition and cognitive function in late adulthood. *Proceedings of the National Academy of Sciences 107*(39), 16881–16886.

Dean, R. F. A. (1951). The size of the baby at birth and the yield of breast milk. In *Medical Research Council Special Report Series No. 275: Studies of Undernutrition, Wuppertal 1946-9, by Members of the Department of Experimental Medicine, Cambridge, and Associated Workers*, Chapter XXVIII, pp. 346–378. His Majesty's Stationary Office.

Del Bono, E., J. Ermisch, and M. Francesconi (2008). Intrafamily resource allocations: a dynamic model of birth weight. Institute for the Study of Labor (IZA) Discussion Paper No. 3704.

Doblhammer, G., G. J. Van den Berg, and L. Lumey (2013). Long-term Effects of Famine on Life Expectancy: A Re-analysis of the Great Finnish Famine of 1866-1868. *Population Studies 5534*, forthcoming.

Dohmen, T., A. Falk, D. Huffman, U. Sunde, J. Schupp, and G. Wagner (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association 9*(3), 522–550.

Dohmen, T., U. Sunde, A. Falk, and D. Huffman (2010). Are risk aversion and impatience related to cognitive ability? *American Economic Review 100*(3), 1238–1260.

Dols, M. J. L. and D. J. A. M. van Arcken (1946). Food supply and nutrition in the netherlands during and immediately after world war II. *The Milbank Memorial Fund Quarterly 24*(4), 319–358.

Duncan, G. and J. Morgan (1981). Sense of Efficacy and Subsequent Change in Earnings–A Replication. *Journal of Human Resources 16*(4), 649–657.

Duncan, G. J. and R. Dunifon (1998). Soft-Skills and Long-Run Labor Market Success. *Research in Labor Economics 17*, 123–149.

Dustmann, C., P. A. Puhani, and U. Schönberg (2012). The long-term effects of school quality on labor market outcomes and educational attainment. Technical report, Centre for Research and Analysis of Migration (CReAM), Department of Economics, University College London.

Echternkamp, J. (2003). *Nach dem Krieg – Alltagsnot, Neuorientierung und die Last der Vergangenheit 1945–1949*. Pendo Verlag Zürich. in German only.

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

Eliason, M. and D. Storrie (2009). Job loss is bad for your health–swedish evidence on cause-specific hospitalization following involuntary job loss. *Social Science & Medicine 68*(8), 1396–1406.

Elo, I. and S. Preston (1992). Effects of early-life conditions on adult mortality: a review. *Population index 58*(2), 186–212.

Eriksson, J. G., E. Kajantie, C. Osmond, K. Thornburg, and D. J. Barker (2010). Boys live dangerously in the womb. *American Journal of Human Biology 22*(3), 330–335.

Escobar, M. D. and M. West (1995). Bayesian Density Estimation and Inference Using Mixtures. *Journal of the American Statistical Association 90*, 577–588.

Evans, W. and J. Ringel (1999). Can higher Cigarette Taxes improve Birth Outcomes? *Journal of Public Economics 72*(1), 135–154.

Fairlie, R. (1999). The absence of the african-american owned business: An analysis of the dynamics of self-employment. *Journal of Labor Economics 17*(1), 80–108.

Fairlie, R. (2005). An extension of the blinder-oaxaca decomposition technique to logit and probit models. *Journal of Economic and Social Measurement 30*(4), 305–316.

Falk, A., D. Huffman, and U. Sunde (2006). Self-confidence and search. Institute for the Study of Labor (IZA) Discussion Paper No. 2525.

Farquharson, J. (1985). *The Western Allies and the Politics of Food – Agrarian Management in Postwar Germany*. Berg Publishers Ltd.

Fenton, T. (2003). A new Growth Chart for Preterm Babies: Babson and Benda's Chart updated with Recent Data and a New Format. *BMC Pediatrics 3*(1), 13.

Ferguson, T. S. (1983a). Bayesian density estimation by mixtures of normal distributions. In H. Chernoff, M. Rizvi, J. Rustagi, and D. Siegmund (Eds.), *Recent Advances in Statistics: Papers in Honor of Herman Chernoff on his Sixtieth Birthday*, pp. 287–302. New York: Academic Press.

Ferguson, T. S. (1983b). Bayesian Density Estimation by Mixtures of Normal Distributions. *Recent Advances in Statistics 24*, 287–303.

Fertig, A. (2010). Selection and the effect of prenatal smoking. *Health Economics 19*(2), 209–226.

Finch, B. (2003). Early Origins of the Gradient: The Relationship between Socioeconomic Status and Infant Mortality in the United States. *Demography 40*(4), 675–699.

Flossmann, A. and W. Pohlmeier (2006). Causal returns to education: A survey on empirical evidence for Germany. *Journal of Economics and Statistics 226*(1), 6–23.

Flossmann, A. L., R. Piatek, and L. Wichert (2007). Going beyond returns to education: The effect of noncognitive skills on wages in Germany. Paper presented at the European Meeting of the Econometric Society, Budapest.

Fougère, D., F. Kramarz, and J. Pouget (2009). Youth unemployment and crime in france. *Journal of the European Economic Association 7*(5), 909–938.

Fox, M., R. Sear, J. Beise, G. Ragsdale, E. Voland, and L. a. Knapp (2010). Grandma plays favourites: X-chromosome relatedness and sex-specific childhood mortality. *Proceedings. Biological sciences / The Royal Society 277*(1681), 567–73.

Fraga, M. F., E. Ballestar, M. F. Paz, S. Ropero, F. Setien, M. L. Ballestar, D. Heine-Suner, J. C. Cigudosa, M. Urioste, J. Benitez, M. Boix-Chornet, A. Sanchez-Aguilera, C. Ling, E. Carlsson, P. Poulsen, A. Vaag, Z. Stephan, T. D. Spector, Y.-Z. Wu, C. Plass, and M. Esteller (2005, July). Epigenetic Differences Arise During the Lifetime of Monozygotic Twins. *Proceedings of the National Academy of Sciences of the United States of America 102*(30), 10604–10609.

Frey, B. S. and S. Meier (2004). Pro-social behavior, reciprocity or both? *Journal of Economic Behavior and Organization 54*, 65–68.

Frisancho, A., S. Fields, and S. Smith (1994). Small for gestational age associated with short stature during adolescence. *American journal of human biology 6*(3), 305–309.

Fuchs, M. and H. Kempermann (2011). Regional effects of the crisis in german engine building industries. In *ERSA conference papers*. European Regional Science Association.

Gallo, W., J. Endrass, E. Bradley, D. Hell, and S. Kasl (2003). The Influence of Internal Control on the Employment Status of German Workers. *Schmollers Jahrbuch 123*(1), 71–81.

Gardosi, J. (2006). New definition of small for gestational age based on fetal growth potential. *Hormone Research in Paediatrics 65*(3), 15–18.

Gasser, T. (1996). Development of fat tissue and body mass index from infancy to adulthood. *Pediatric Nephrology 10*(3), 340–342.

Gasser, T., P. Ziegler, R. H. Largo, L. Molinari, and a. Prader (1994). A longitudinal study of lean and fat areas at the arm. *Annals of human biology 21*(4), 303–14.

Gebel, M. and F. Pfeiffer (2010). Educational expansion and its heterogeneous returns for wage workers. *Schmollers Jahrbuch 130*(1), 19–42.

Ghosal, S. and A. Van Der Vaart (2001). Entropies and rates of convergence for maximum likelihood and bayes estimation for mixtures of normal densities. *The Annals of Statistics 29*(5), 1233–1263.

Ginja, R. (2010). Income shocks and investments in human capital. Job Market Paper.

Gluckman, P. and M. Hanson (2004a). Living with the past: evolution, development, and patterns of disease. *Science 305*(5691), 1733–1736.

Gluckman, P. and M. Hanson (2004b). Maternal constraint of fetal growth and its consequences. In *Seminars in fetal and neonatal medicine*, Number 5, pp. 419–425. Elsevier.

Gluckman, P., M. Hanson, and C. Pinal (2005). The developmental origins of adult disease. *Maternal & child nutrition 1*(3), 130–141.

Goldberger, A. (1972). Structural equation methods in the social sciences. *Econometrica: Journal of the Econometric Society 40*(6), 979–1001.

Gomella, T., M. Cunningham, F. Eyal, and K. Zenk (1999). *Neonatology: management, procedures, on-call problems, diseases, and drugs*. Appleton & Lange.

Gottschalk, P. (2005). Can work alter welfare recipients' beliefs? *Journal of Policy Analysis and Management 24*(3), 485–498.

Gould, E., B. Weinberg, and D. Mustard (2002). Crime rates and local labor market opportunities in the united states: 1979-1997. *Review of Economics and Statistics 84*(1), 45–61.

Gráda, C. (2009). *Famine: a short history*. Princeton University Press.

Gräff, J. and I. M. Mansuy (2008, September). Epigenetic codes in cognition and behaviour. *Behavioural brain research 192*(1), 70–87.

Gregg, P., L. Macmillan, and B. Nasim (2012). The impact of fathers' job loss during the recession of the 1980s on their children's educational attainment and labour market outcomes*. *Fiscal Studies 33*(2), 237–264.

Griliches, Z. (1979, October). Sibling models and data in economics: Beginnings of a survey. *Journal of Political Economy 87*(5), 37–64.

Hadley, C. and C. Patil (2006). Food insecurity in rural tanzania is associated with maternal anxiety and depression. *American Journal of Human Biology 18*(3), 359–368.

Hales, C. and D. Barker (1992). Type 2 (non-insulin-dependent) diabetes mellitus: the thrifty phenotype hypothesis. *Diabetologia 35*(7), 595–601.

Ham, J. and K. Jacobs (2000). Testing for full insurance using exogenous information. *Journal of Business & Economic Statistics 18*(4), 387–397.

Hamelin, A., J. Habicht, and M. Beaudry (1999). Food insecurity: consequences for the household and broader social implications. *The Journal of nutrition 129*(2), 525–528.

Hansen, K. T., J. J. Heckman, and K. J. Mullen (2004). The effect of schooling and ability on achievement test scores. *Journal of Econometrics 121*(1–2), 39–98.

Harper, L. V. (2005). Epigenetic inheritance and the intergenerational transfer of experience. *Psychological bulletin 131*(3), 340–60.

Harville, E., R. Boynton-Jarrett, C. Power, and E. Hypponen (2010). Childhood hardship, maternal smoking, and birth outcomes: a prospective cohort study. *Archives of Pediatrics and Adolescent Medicine 164*(6), 533.

Häusser, A. and G. Maugg (1985). *Hungerwinter. Deutschlands humanitäre Katastrophe 1946/47*. Berlin: Prophyläen.

Heckman, J., J. Humphries, S. Urzua, and G. Veramendi (2011). The effects of educational choices on labor market, health, and social outcomes. Working Paper, Human Capital and Economic Opportunity Working Group.

Heckman, J., H. Ichimura, and P. Todd (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The Review of Economic Studies 64*(4), 605.

Heckman, J., L. Lochner, and P. Todd (2006). Earnings functions, rates of return and treatment effects: The Mincer equation and beyond. *Handbook of the Economics of Education 1*, 307–458.

Heckman, J., R. Pinto, and P. Savelyev (2013a). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review*, forthcoming.

Heckman, J., R. Pinto, and P. Savelyev (2013b). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review*, Forthcoming.

Heckman, J., J. Stixrud, and S. Urzua (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics 24*(3), 411–482.

Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. *Econometrica 47*, 153–161.

Heijmans, B. T., E. W. Tobi, A. D. Stein, H. Putter, G. J. Blauw, E. S. Susser, P. E. Slagboom, and L. H. Lumey (2008). Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proceedings of the National Academy of Sciences 105*(44), 17046–17049.

Heineck, G. and S. Anger (2010). The returns to cognitive abilities and personality traits in Germany. *Labour Economics 17*(3), 535–546.

Hennessy, E. and E. Alberman (1998). Intergenerational influences affecting birth outcome. Birthweight for gestational age in the children of the 1958 british birth cohort. *Paediatric and perinatal epidemiology 12*(1), 45–60.

Hermanussen, M., H. Danker-Hopfe, and G. Weber (2001). Body weight and the shape of the natural distribution of weight, in very large samples of german, austrian and norwegian conscripts. *International journal of obesity 25*(10), 1550–1553.

Hill, M. S. (1979). The Wage Effects of Marital Status and Children. *Journal of Human Resources 14*, 579–594.

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

Hochberg, Z., R. Feil, M. Constancia, M. Fraga, C. Junien, J.-C. Carel, P. Boileau, Y. Le Bouc, C. L. Deal, K. Lillycrop, R. Scharfmann, a. Sheppard, M. Skinner, M. Szyf, R. a. Waterland, D. J. Waxman, E. Whitelaw, K. Ong, and K. Albertsson-Wikland (2011). Child health, developmental plasticity, and epigenetic programming. *Endocrine reviews 32*(2), 159–224.

Horny, G., R. Mendes, and G. Berg (2009). Job Durations with Worker and Firm Specific Effects: MCMC Estimation with Longitudinal Employer-Employee Data. Institute for the Study of Labor (IZA) Discussion Paper No. 3992.

Huber, D. and E. Fogel (1920). Food conditions and agricultural production. *Annals of the American Academy of political and social science, Social and Industrial Conditions in the Germany of Today 92*, 131–136.

Huber, M., M. Lechner, C. Wunsch, and T. Walter (2011). Do german welfare-to-work programmes reduce welfare dependency and increase employment? *German Economic Review 12*(2), 182–204.

Hutcheon, J. and R. Platt (2008). The missing data problem in birth weight percentiles and thresholds for small-for-gestational-age. *American journal of epidemiology 167*(7), 786–792.

Hutchinson, A., R. A. McCance, and E. M. Widdowson (1951). The size of the baby at birth and the yield of breast milk. In *Medical Research Council Special Report Series No. 275: Studies of Undernutrition, Wuppertal 1946-9, by Members of the Department of Experimental Medicine, Cambridge, and Associated Workers*, Chapter XXVIII, pp. 346–378. His Majesty's Stationary Office.

Huxley, R., C. Owen, P. Whincup, D. Cook, J. Rich-Edwards, G. Smith, and R. Collins (2007). Is birth weight a risk factor for ischemic heart disease in later life? *The American journal of clinical nutrition 85*(5), 1244.

Ido, P., R. Gale, A. Laor, Y. Danon, D. Stevenson, and D. Seidman (1995). The cognitive outcome of full-term small for gestational age infants at late adolescence. *Obstetrics & Gynecology 85*(3), 452–456.

Imbens, G. and J. Angrist (1994). Identification and estimation of local average treatment effects. *Econometrica: Journal of the Econometric Society 62*(2), 467–475.

Inoue, A. and G. Solon (2010). Two-sample instrumental variables estimators. *The Review of Economics and Statistics 92*(3), 557–561.

Iwata, S. (1992). Errors-in-variables regression using estimated latent variables. *Econometric Reviews 11*(2), 195–200.

Jacobson, L., R. LaLonde, and D. Sullivan (1993). Earnings losses of displaced workers. *The American Economic Review 83*(4), 685–709.

Jefferis, B., C. Power, and C. Hertzman (2002). Birth weight, childhood socioeconomic environment, and cognitive development in the 1958 british birth cohort study. *Bmj 325*(7359), 305.

Jenkins, S. and C. Schluter (2002). The Effect of Family Income During Childhood on Later-Life Attainment: Evidence from Germany. Institute for the Study of Labor (IZA) Discussion Paper No. 604.

Joreskog, K. and A. Goldberger (1975). Estimation of a model with multiple indicators and multiple causes of a single latent variable. *Journal of the American Statistical Association 70*(351), 631–639.

Jöreskog, K. G. and I. Moustaki (2001). Factor analysis of ordinal variables: A comparison of three approaches. *Multivariate Behavioral Research 36*(3), 347–387.

Judge, T. A. and J. E. Bono (2001). Relationship of core self-evaluations traits-self-esteem, generalized self-efficacy, locus of control, and emotional stability-with job satisfaction and job performance: A meta-analysis. *Journal of Applied Psychology 86*(1), 80–92.

Jürges, H. (2012). Collateral damage: Educational attainment and labor market outcomes among German war and post-war cohorts. Discussion paper, University of Wuppertal, Wuppertal, Germany.

Jyoti, D., E. Frongillo, and S. Jones (2005). Food insecurity affects school children's academic performance, weight gain, and social skills. *The Journal of nutrition 135*(12), 2831.

Kaati, G., L. O. Bygren, and S. Edvinsson (2002, December). Cardiovascular and diabetes mortality determined by nutrition during parents' and grandparents' slow growth period. *European journal of human genetics : EJHG 10*(11), 682–8.

Kaati, G., L. O. Bygren, M. Pembrey, and M. Sjöström (2007). Transgenerational response to nutrition, early life circumstances and longevity. *European journal of human genetics 15*(7), 784–90.

Kalil, A. and K. Ziol-Guest (2008). Parental employment circumstances and children's academic progress. *Social Science Research 37*(2), 500–515.

Karlberg, J., K. Albertsson-Wikland, et al. (1995). Growth in full-term small-for-gestational-age infants: from birth to final height. *Pediatric Research 38*(5), 733.

Kassenboehmer, S. and J. Haisken-DeNew (2009). You're fired! the causal negative effect of entry unemployment on life satisfaction. *The Economic Journal 119*(536), 448–462.

Kind, M. and J. Haisken-DeNew (2012). Sons' unexpected long term scarring due to fathers' unemployment. Melbourne Institute Working Paper.

Klatt, W. (1950). Food and farming in Germany: I. food and nutrition. *International Affairs 26*(1), 45–58.

Klein, F. (1968). *Deutschland im ersten Weltkrieg*. Leipziger Universitätsverlag.

Koop, G. and J. Tobias (2004). Learning about heterogeneity in returns to schooling. *Journal of Applied Econometrics 19*(7), 827–849.

Kotlarski, I. (1967). On characterizing the gamma and the normal distribution. *Pacific Journal of Mathematics 20*(1), 69–76.

Kramer, M. (1987). Determinants of Low Birth Weight: Methodological Assessment and Meta-Analysis. *Bulletin of the World Health Organization 65*(5), 663.

Kramer, M., L. Goulet, J. Lydon, L. Séguin, H. McNamara, C. Dassa, R. Platt, M. Fong Chen, H. Gauthier, J. Genest, et al. (2001). Socio-economic Disparities in Preterm Birth: Causal Pathways and Mechanisms. *Paediatric and Perinatal Epidemiology 15*, 104–123.

Kramer, M., F. McLean, M. Boyd, and R. Usher (1988). The validity of gestational age estimation by menstrual dating in term, preterm, and postterm gestations. *Journal of the American Medical Association 260*(22), 3306.

Kuzawa, C. (2005). Fetal origins of developmental plasticity: are fetal cues reliable predictors of future nutritional environments? *American Journal of Human Biology 17*(1), 5–21.

Kuzawa, C. (2008). The developmental origins of adult health: intergenerational inertia in adaptation and disease. *Evolutionary Medicine and Health: New Perspective. Oxford University Press, New York 102*(2), 23–58.

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

Lang, F., D. Weiss, A. Stocker, and B. Von Rosenbladt (2007). Assessing cognitive capacities in computer-assisted survey research: Two ultra-short tests of intellectual ability in the german socio-economic panel (soep). *Schmollers Jahrbuch 127*(1), 183–192.

Lauer, C. and V. Steiner (2000). Private Erträge von Bildungsinvestitionen in Deutschland. *Applied Economics Quarterly 51*, 71–101.

Lawlor, D. (2008). The developmental origins of health and disease: where do we go from here? *Epidemiology 19*(2), 206.

Lawlor, D. A., Y. Ben-Shlomo, and D. A. Leon (2004). Pre-adult influences on cardiovascular. In *A life course approach to chronic disease epidemiology*. Oxford University Press.

Lee, P., S. Chernausek, A. Hokken-Koelega, P. Czernichow, et al. (2003). International small for gestational age advisory board consensus development conference statement: management of short children born small for gestational age, april 24–october 1, 2001. *Pediatrics 111*(6), 1253.

Leigh, J. P. (1983). Direct and indirect effects of education on health. *Social Science and Medicine 17*(4), 227–234.

Leon, D. A., H. O. Lithell, D. Vågerö, I. Koupilová, R. Mohsen, L. Berglund, U.-B. Lithell, and P. M. McKeigue (1998). Reduced fetal growth rate and increased risk of death from ischaemic heart disease: cohort study of 15 000 swedish men and women born 1915-29. *BMJ: British Medical Journal 317*(7153), 241.

Li, K. and D. Poirier (2003a). An econometric model of birth inputs and outputs for Native Americans. *Journal of Econometrics 113*, 337–361.

Li, K. and D. Poirier (2003b). Bayesian analysis of an econometric model of birth inputs and outputs. *Journal of Population Economics 16*(3), 597–625.

Li, K., D. Poirier, and C. Drive (2003). Relationship between maternal behavior during pregnancy, birth outcome, and early childhood development: An exploratory study. CESifo Working Paper.

Li, M. (2006). High school completion and future youth unemployment: new evidence from high school and beyond. *Journal of Applied Econometrics 21*(1), 23–53.

Lien, D. and W. Evans (2005). Estimating the impact of large cigarette tax hikes. *Journal of Human Resources 40*(2), 373–392.

Lindeboom, M., F. Portrait, and G. Van den Berg (2010). Long-run effects on longevity of a nutritional shock early in life: The dutch potato famine of 1846-1847. *Journal of health economics 29*(5), 617–629.

Liu, J. S. and C. Sabatti (2000). Generalised Gibbs sampler and multigrid Monte Carlo for Bayesian computation. *Biometrika 87*(2), 353–369.

Loewenberg, P. (1971). The psychohistorical origins of the Nazi youth cohort. *The American Historical Review 76*(5), 1457–1502.

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

Lu, I. and D. R. Thomas (2008). Avoiding and Correcting Bias in Score-Based Latent Variable Regression With Discrete Manifest Items. *Structural Equation Modeling: A Multidisciplinary Journal 15*(3), 462–490.

Lumey, L. and A. Stein (1997). Offspring birth weights after maternal intrauterine undernutrition: a comparison within sibships. *American journal of epidemiology 146*(10), 810.

Lumey, L., A. Stein, H. Kahn, K. van der Pal-de Bruin, et al. (2007). Cohort profile: the dutch hunger winter families study. *International journal of epidemiology 36*(6), 1196.

Lumey, L., A. Stein, and E. Susser (2011). Prenatal famine and adult health. *Annual Review of Public Health 32*, 237–262.

Lundborg, P. and A. Stenberg (2010). Nature, nurture and socioeconomic policy–what can we learn from molecular genetics? *Economics & Human Biology 8*(3), 320–330.

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

Magnusson, P., F. Rasmussen, and U. Gyllensten (2006). Height at age 18 years is a strong predictor of attained education later in life: cohort study of over 950 000 Swedish men. *International Journal of Epidemiology 35*(3), 658–663.

Mamelund, S.-E. (2003). Effects of the Spanish Influenza Pandemic of 1918-19 on Later Life Mortality of Norwegian Cohorts Born About 1900. Working Paper, Oslo University, Department of Economics.

Marshall, W. and J. Tanner (1968). Growth and physiological development during adolescence. *Annual Review of Medicine 19*(1), 283–300.

Masterpasqua, F. (2009). Psychology and epigenetics. *Review of General Psychology 13*(3), 194–201.

Mayer, W., a. Niveleau, J. Walter, R. Fundele, and T. Haaf (2000, February). Demethylation of the zygotic paternal genome. *Nature 403*(6769), 501–2.

McCloskey, D. N. and S. T. Ziliak (1996). The standard error of regressions. *Journal of Economic Literature 34*(1), 97–114.

McCrary, J. and H. Royer (2011). The Effect of Female Education on Fertility and Infant Health: Evidence from School Entry Policies Using Exact Date of Birth. *The American Economic Review 101*(1), 158–195.

McDermott, P. and M. Watkins (1981). Dimensions of maladaptive behavior among kindergarten level children. *Behavioral Disorders 7*(1), 11–17.

McGowan, P. O., A. Sasaki, T. C. T. Huang, A. Unterberger, M. Suderman, C. Ernst, M. J. Meaney, G. Turecki, and M. Szyf (2008). Promoter-wide hypermethylation of the ribosomal RNA gene promoter in the suicide brain. *PloS one 3*(5), 1–10.

McKee-Ryan, F., Z. Song, C. Wanberg, and A. Kinicki (2005). Psychological and physical well-being during unemployment: a meta-analytic study. *Journal of Applied Psychology 90*(1), 53.

Meyer, S., A. Raisig, L. Gortner, M. Ong, M. Bücheler, and E. Tutdibi (2009). In utero tobacco exposure: The effects of heavy and very heavy smoking on the rate of sga infants in the federal state of saarland, Germany. *European Journal of Obstetrics & Gynecology and Reproductive Biology 146*(1), 37–40.

Miller, G. and E. Chen (2010). Harsh family climate in early life presages the emergence of a proinflammatory phenotype in adolescence. *Psychological science 21*(6), 848.

Miller, G., E. Chen, A. Fok, H. Walker, A. Lim, E. Nicholls, S. Cole, and M. Kobor (2009). Low early-life social class leaves a biological residue manifested by decreased glucocorticoid and increased proinflammatory signaling. *Proceedings of the National Academy of Sciences 106*(34), 14716–14721.

Mincer, J. (1958). Investment in human capital and personal income distribution. *Journal of Political Economy 66*(4), 281–302.

Möller, J. (2010). The german labor market response in the world recession–de-mystifying a miracle. *Journal for Labour Market Research 42*(4), 325–336.

Morozink, J., E. Friedman, C. Coe, and C. Ryff (2010). Socioeconomic and psychosocial predictors of interleukin-6 in the midus national sample. *Health Psychology 29*(6), 626.

Mustafa, G. and R. David (2001). Comparative accuracy of clinical estimate versus menstrual gestational age in computerized birth certificates. *Public Health Reports 116*(1), 15.

Nandram, B. and M. H. Chen (1996). Reparameterizing the generalized linear model to accelerate Gibbs sampler convergence. *Journal of Statistical Computation and Simulation 54*(1), 129–144.

Nathanielsz, P. W. (2003). Fetal programming: from gene to functional systems - an overview. *The Journal of Physiology 547*(1), 3–4.

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

Neugebauer, R., H. W. Hoek, and E. Susser (1999). Prenatal exposure to wartime famine and development of antisocial personality disorder in early adulthood. *JAMA : the journal of the American Medical Association 282*(5), 455–62.

OECD (2012). *OECD economic surveys: Germany*. Organisation for Economic Co-operation and Development.

Onland-Moret, N., P. Peeters, C. Van Gils, F. Clavel-Chapelon, T. Key, A. Tjønneland, A. Trichopoulou, R. Kaaks, J. Manjer, S. Panico, et al. (2005). Age at menarche in relation to adult height. *American Journal of Epidemiology 162*(7), 623–632.

Oppers, V. (1963). *Analyse van de acceleratie van de menselijke lengtegroei door bepaling van het tijdstip van de groeifasen.* Ph. D. thesis, University of Amsterdam.

Oreopoulos, P., M. Page, and A. Stevens (2008). The intergenerational effects of worker displacement. *Journal of Labor Economics 26*(3), 455–000.

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

Oreopoulos, P., T. von Wachter, and A. Heisz (2012). The short-and long-term career effects of graduating in a recession. *American Economic Journal: Applied Economics 4*(1), 1–29.

Osborne, M. A. (2000). *The Power of Personality: Labor Market Rewards and the Transmission of Earnings*. Ph. D. thesis, University of Massachusetts.

Painter, R., T. Roseboom, and O. Bleker (2005). Prenatal exposure to the Dutch famine and disease in later life: an overview. *Reproductive Toxicology 20*(3), 345–352.

Painter, R. C., C. Osmond, P. Gluckman, M. Hanson, D. I. W. Phillips, and T. J. Roseboom (2008). Transgenerational effects of prenatal exposure to the Dutch famine on neonatal adiposity and health in later life. *BJOG : An International Journal of Obstetrics and Gynaecology 115*(10), 1243–9.

Pembrey, M. E. (2002). Time to take epigenetic inheritance seriously. *European Journal of Human Genetics 10*(11), 669–71.

Pembrey, M. E., L. O. Bygren, G. Kaati, S. Edvinsson, K. Northstone, M. Sjöström, and J. Golding (2006). Sex-specific, male-line transgenerational responses in humans. *European Journal of Human Genetics 14*(2), 159–66.

Pfann, G. and D. Hamermesh (2001). Two-sided learning, labor turnover and displacement. NBER working paper no. 8273.

Piatek, R. (2010). Bayesian inference for factor structure models via gibbs sampling. Unpublished technical report, University of Konstanz.

Piatek, R. and P. Pinger (2010). Maintaining (locus of) control? assessing the impact of locus of control on education decisions and wages. Institute for the Study of Labor (IZA) Discussion Paper No. 3646.

Pischke, J. and T. Von Wachter (2008). Zero returns to compulsory schooling in Germany: Evidence and interpretation. *The Review of Economics and Statistics 90*(3), 592–598.

Pitkaïnen, K. J. (1993). *Mortality during the Great Finnish Famine of the 1860s*. Helsinki: Finnish Demographic Society.

Pollitt, R., K. Rose, and J. Kaufman (2005). Evaluating the evidence for models of life course socioeconomic factors and cardiovascular outcomes: a systematic review. *BMC public health 5*(1), 7.

Poulsen, G., J. Kurinczuk, D. Wolke, E. Boyle, D. Field, Z. Alfirevic, and M. Quigley (2011). Accurate reporting of expected delivery date by mothers 9 months after birth. *Journal of Clinical Epidemiology 64*(12), 1444–1450.

Poulter, N., C. Chang, A. MacGregor, H. Snieder, and T. Spector (1999). Association between birth weight and adult blood pressure in twins: historical cohort study. *Bmj 319*(7221), 1330.

Psacharopoulos, G. (1981). Returns to education: an updated international comparison. *Comparative Education 17*(3), 321–341.

Quandt, R. (1972). A new approach to estimating switching regressions. *Journal of the American Statistical Association 67*(338), 306–310.

Radtke, K. M., M. Ruf, H. M. Gunter, K. Dohrmann, M. Schauer, a. Meyer, and T. Elbert (2011). Transgenerational impact of intimate partner violence on methylation in the promoter of the glucocorticoid receptor. *Translational Psychiatry 1*(7), e21.

Raico, R. (1989, December). The politics of hunger: A review. *The Review of Austrian Economics 3*(1), 253–259.

Rakyan, V. K., S. Chong, M. E. Champ, P. C. Cuthbert, H. D. Morgan, K. V. K. Luu, and E. Whitelaw (2003). Transgenerational inheritance of epigenetic states at the murine Axin(Fu) allele occurs after maternal and paternal transmission. *Proceedings of the National Academy of Sciences of the United States of America 100*(5), 2538–43.

Rampino, T. and M. Taylor (2012). Educational aspirations and attitudes over the business cycle. Working Paper, Institute for Social and Economic Research.

Rasmussen, K. (2001). The fetal origins hypothesis: challenges and opportunities for maternal and child nutrition. *Annual Review of Nutrition 21*(1), 73–95.

Rege, M., K. Telle, and M. Votruba (2011). Parental job loss and children's school performance. *The Review of Economic Studies 78*(4), 1462–1489.

Reichardt, S. and M. Zierenberg (2008). *Damals nach dem Krieg: eine Geschichte Deutschlands 1945 bis 1949*. München: Goldmann.

Roesle, E. (1928). Die Geburts- und Sterblichkeitsverhältnisse. In F. Blum (Ed.), *Deutschlands Gesundheitsverhältnisse unter dem Einfluss des Weltkrieges*, Deutsche Serie. Yale University Press, New Haven: Deutsche Verlags-Anstalt.

Rose, H. and J. Betts (2004). The effect of high school courses on earnings. *Review of Economics and Statistics 86*(2), 497–513.

Rosenzweig, M. and W. K. (1988). Heterogeneity, intrafamily distribution and child health. *Journal of Human Resources 23*(4), 437.

Rosenzweig, M. and T. Schultz (1983). Estimating a Household Production Function: Heterogeneity, the Demand for Health Inputs, and Their Effects on Birth Weight. *The Journal of Political Economy 91*(5), 723.

Rosenzweig, M. and K. Wolpin (1991). Inequality at Birth: The Scope for Policy Intervention. *Journal of Econometrics 50*(1-2), 205–228.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General & Applied 80*(1), 1–28.

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

Rubalcava, L. and G. Teruel (2004). The role of maternal cognitive ability on child health. *Economics & Human Biology 2*(3), 439–455.

Ruhm, C. (1991). Are workers permanently scarred by job displacements? *The American Economic Review 81*(1), 319–324.

Ruhm, C. (2004). Parental employment and child cognitive development. *Journal of Human Resources 39*(1), 155–192.

Saris, W., M. de Pijper, and J. Mulder (1978). Optimal procedures for estimation of factor scores. *Sociological Methods & Research 7*(1), 85.

Scholte, R. S., G. J. Van den Berg, and M. Lindeboom (2012). Long-run effects of gestation during the dutch hunger winter famine on labor market and hospitalization outcomes. Institute for the Study of Labor (IZA) Discussion Paper No. 6307.

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

Schulz, L. (2010). The dutch hunger winter and the developmental origins of health and disease. *Proceedings of the National Academy of Sciences 107*(39), 16757–16758.

Scott, D. W. (1992). *Multivariate Density Estimation: Theory, Practice, and Visualization.* Wiley Interscience.

Shah, P. (2010). Parity and low birth weight and preterm birth: a systematic review and meta-analyses. *Acta obstetricia et gynecologica Scandinavica 89*(7), 862–875.

Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis.* CRC Press.

Skrondal, A. and P. Laake (2001). Regression among factor scores. *Psychometrika 66*(4), 563–575.

Sparén, P., D. Vågerö, D. Shestov, S. Plavinskaja, N. Parfenova, V. Hoptiar, D. Paturot, and M. Galanti (2004). Long term mortality after severe starvation during the siege of leningrad: prospective cohort study. *British Medical Journal 328*(7430), 11.

Spence, D., M. Stewart, F. Alderdice, C. Patterson, and H. Halliday (2012). Intra-uterine growth restriction and increased risk of hypertension in adult life: A follow-up study of 50-year-olds. *Public Health 126*(7), 561–565.

St Clair, D., M. Xu, P. Wang, Y. Yu, Y. Fang, F. Zhang, X. Zheng, N. Gu, G. Feng, P. Sham, and Others (2005). Rates of adult schizophrenia following prenatal exposure to the Chinese famine of 1959-1961. *Journal of the American Medical Association 294*(5), 557.

Starling, E., A. McDougall, and C. Guillebaud (1919). *Report on food conditions in Germany*. HM Stationery Office.

Steckel, R. H. (2008). Biological measures of the standard of living. *The Journal of Economic Perspectives 22*(1), 129–152.

Stevens, A. and J. Schaller (2011). Short-run effects of parental job loss on children's academic achievement. *Economics of Education Review 30*(2), 289–299.

Strauss, R. (2000). Adult Functional Outcome of Those Born Small for Gestational Age Twenty-six-Year Follow-up of the 1970 British Birth Cohort. *JAMA 283*(5), 625–632.

Subramanian, S., L. Ackerson, G. Davey Smith, and N. John (2009). Association of maternal height with child mortality, anthropometric failure, and anemia in india. *Journal of the American Medical Association 301*(16), 1691.

Sullivan, D. and T. Von Wachter (2009). Job displacement and mortality: An analysis using administrative data. *The Quarterly Journal of Economics 124*(3), 1265–1306.

Susser, E. and S. Lin (1992). Schizophrenia after prenatal exposure to the dutch hunger winter of 1944-1945. *Archives of general psychiatry 49*(12), 983.

Susser, M. and Z. Stein (1994). Timing in prenatal nutrition: a reprise of the dutch famine study. *Nutrition reviews 52*(3), 84–94.

Tanner, M. A. and W. H. Wong (1987). The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association 82*(398), 528–540.

Theodossiou, I. (1998). The effects of low-pay and unemployment on psychological well-being: A logistic regression approach. *Journal of Health Economics 17*(1), 85–104.

Thomson, A., W. Billewicz, and F. Hytten (1968). The assessment of fetal growth. *BJOG: An International Journal of Obstetrics & Gynaecology 75*(9), 903–916.

Tobi, E. W., L. H. Lumey, R. P. Talens, D. Kremer, H. Putter, A. D. Stein, P. E. Slagboom, and B. T. Heijmans (2009). DNA methylation differences after exposure to prenatal famine are common and timing- and sex-specific. *Human molecular genetics 18*(21), 4046–53.

Tominey, E. (2007). Maternal Smoking During Pregnancy and Early Child Outcomes. CEP Discussion Papers, Centre for Economic Performance, LSE.

Tominey, E. (2009). The timing of parental income and child human capital: Estimating the role of permanent and transitory income shocks. Mimeo.

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

Trzcinski, E. and E. Holst (2010). Interrelationship among locus of control and years in management and unemployment: Differences by gender. Discussion Papers of DIW Berlin 974, DIW Berlin, German Institute for Economic Research.

Uhlendorff, A. (2004). Der Einfluss von Persönlichkeitseigenschaften und sozialen Ressourcen auf die Arbeitslosigkeitsdauer. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie 56*(2), 279–303.

Valaoras, V. (1946a). Some effects of famine on the population of greece. *The Milbank Memorial Fund Quarterly 24*(3), 215–234.

Valaoras, V. G. (1946b). Effects of famine on the population of greece. *The Milbank Memorial Fund Quarterly 24*, 215–235.

van den Berg, G., G. Doblhammer-Reiter, and K. Christensen (2011). Being born under adverse economic conditions leads to a higher cardiovascular mortality rate later in life: Evidence based on individuals born at different stages of the business cycle. *Demography 48*(2), 507–530.

Van den Berg, G. J., G. Doblhammer, K. Christensen, et al. (2009). Exogenous determinants of early-life conditions, and mortality later in life. *Social Science & Medicine 68*(9), 1591–1598.

Van den Berg, G. J. and S. Gupta (2007). Early Life Conditions, Marital Status, and Mortality. Institute for the Study of Labor (IZA) Discussion Paper No. 3350.

Van den Berg, G. J., M. Lindeboom, and F. Portrait (2006, March). Economic conditions early in life and individual mortality. *American Economic Review 96*(1), 290–302.

Van den Berg, G. J. and P. Pinger (2011). A validation study of intergenerational effects of early-life conditions on offspring's economic and health outcomes potentially driven by epigenetic imprinting: Evidence from the German famine of 1916-1918. Unpublished manuscript, University of Mannheim, Department of Economics.

Van den Berg, G.J. and Lindeboom, M. (2012). Aging: conditions early in life and health and mortality at advanced ages. In e. A. J. Culyer (Ed.), *Encyclopedia of Health Economics*. Elsevier, Amsterdam.

Van den Berg, G.J. and Lundborg, P. and Nystedt, P. and Rooth, D.O. (2010). Critical Periods During Childhood and Adolescence: A Study of Adult Height Among Immigrant Siblings. Institute for the Study of Labor (IZA) Discussion Paper No. 4140.

van Dyk, D. A. and X.-L. Meng (2001). The art of data augmentation. *Journal of Computational & Graphical Statistics 10*(1), 1–50.

Vincent, C. P. (1985). *The Politics of Hunger: The Allied Blockade of Germany, 1915-1919*. Ohio University Press.

Waaler, H. (1984). Height. Weight and Mortality The Norwegian Experience. *Acta Medica Scandinavica 215*(S679), 1–56.

Wagner, G. G., J. R. Frick, and J. Schupp (2007). The German socio-economic panel study (SOEP)–scope, evolution and enhancements. *Schmollers Jahrbuch 127*(1), 139–169.

Walker, M., E. Tekin, and S. Wallace (2009). Teen Smoking and Birth Outcomes. *Southern Economic Journal 75*(3), 892–907.

Whitaker, R., S. Phillips, and S. Orzol (2006). Food insecurity and the risks of depression and anxiety in mothers and behavior problems in their preschool-aged children. *Pediatrics 118*(3), 859.

Whitelaw, E. (2006). Epigenetics: sins of the fathers, and their fathers. *European journal of human genetics : EJHG 14*(2), 131–2.

Willenborg, K.-H. (1979). Überleben nach dem Zusammenbruch – die Not der Besiegten. In J. Weber (Ed.), *Geschichte der Bundesrepublik Deutschland, Band1: Auf dem Wege zur Republik 1945–1947*, Chapter 6, pp. 187–210. Schöningh. in German only.

Wingate, M., G. Alexander, P. Buekens, and A. Vahratian (2007). Comparison of gestational age classifications: date of last menstrual period vs. clinical estimate. *Annals of epidemiology 17*(6), 425–430.

Witte, W. (2008). *Tollkirschen und Quarantäne*. Wagenbach.

Zemel, B. (2002). Body composition during growth and development. In N. Cameron (Ed.), *Human growth and development*, pp. 271–291. London, UK: Academic Press.

Zhang, T., R. Bagot, C. Parent, C. Nesbitt, T. Bredy, C. Caldji, E. Fish, H. Anisman, M. Szyf, and M. Meaney (2006). Maternal programming of defensive responses through sustained effects on gene expression. *Biological psychology 73*(1), 72–89.

Ziliak, S. T. and D. N. McCloskey (2004). Size matters: the standard error of regressions in the american economic review. *Journal of Socio-Economics 33*(5), 527–546.

Curriculum Vitae

2013 onwards	Assistant professor University of Bonn, Germany
2007-2013	PhD University of Mannheim, Germany
2008 and 2011	Visiting scholar University of Chicago, USA
2006-2007	Advanced Studies Program Kiel Institute for the World Economy, Germany
2001-2006	Master in International Economic Studies Maastricht University, The Netherlands
2004-2005	Erasmus exchange (Master) Institut d'Etudes Politiques (Sciences Po), Paris, France