

Anger, Silke; Schnitzlein, Daniel D.

**Working Paper**

## Cognitive skills, non-cognitive skills, and family background: Evidence from sibling correlations

Hannover Economic Papers (HEP), No. 561

**Provided in Cooperation with:**

School of Economics and Management, University of Hannover

*Suggested Citation:* Anger, Silke; Schnitzlein, Daniel D. (2015) : Cognitive skills, non-cognitive skills, and family background: Evidence from sibling correlations, Hannover Economic Papers (HEP), No. 561, Leibniz Universität Hannover, Wirtschaftswissenschaftliche Fakultät, Hannover

This Version is available at:

<https://hdl.handle.net/10419/121772>

**Standard-Nutzungsbedingungen:**

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

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

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

**Terms of use:**

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

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

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

# Cognitive Skills, Non-Cognitive Skills, and Family Background: Evidence from Sibling Correlations

Silke Anger

*IAB Nuremberg, University of Bamberg, IZA*

Daniel D. Schnitzlein\*

*Leibniz University Hannover, DIW Berlin*

This version: July 2015

## Abstract

This paper estimates sibling correlations in cognitive and non-cognitive skills to evaluate the importance of family background for skill formation. Based on a large representative German dataset including IQ test scores and measures of non-cognitive skills, a restricted maximum likelihood model indicates substantial influences of family background on skill formation. Sibling correlations in non-cognitive skills range from 0.223 to 0.464; therefore, at least one-fifth of the variance in these skills results from sibling-related factors. Sibling correlations in cognitive skills are higher than 0.50; therefore, more than half of the inequality in cognition can be explained by family background. Comparing these findings with those in the intergenerational skill transmission literature suggests that intergenerational correlations capture only part of the influence of family on children's cognitive and non-cognitive skills, as confirmed by decomposition analyses and in line with previous findings on educational and income mobility.

**JEL codes:** J24, J62

**Keywords:** Sibling correlations, family background, non-cognitive skills, cognitive skills, intergenerational mobility

---

\*Correspondence to: Daniel D. Schnitzlein; Leibniz University Hannover; Institute of Labour Economics; Königsworther Platz 1; 30167 Hannover; Germany; ✉: schnitzlein@aoek.uni-hannover.de; ☎: +49-(0)511-762-5298.

# 1 Introduction

Economic research emphasizes the importance of cognitive and non-cognitive skills for both individual labor market outcomes and social outcomes.<sup>1</sup> This finding has triggered a growing interest in the determinants of cognitive and non-cognitive skills. [Cunha and Heckman \(2007, 2008\)](#) present a model of skill formation that links the development of these skills to parental cognitive and non-cognitive skills as well as parental resources, among other factors. This link raises a question regarding equality of opportunity. According to [Roemer \(1998\)](#), equality of opportunity requires that an individual's economic success depends only on factors under the individual's control. Environmental factors, which are beyond an individual's control, should not influence future success or failure. The family into which a child is born is clearly beyond the child's control; therefore, the "accident of birth" ([Cunha and Heckman, 2007, p. 37](#)) should not influence individual outcomes. As cognitive and non-cognitive skills are important determinants of economic and social success, the normative goal of equality of opportunity is violated if the formation of these skills is influenced by family background.

A growing body of literature in the field of intergenerational mobility analyzes the transmission of both cognitive and non-cognitive skills from parents to children ([Black and Devereux, 2011](#)). Intergenerational transmission of cognitive skills has been analyzed in the contexts of Scandinavia ([Black et al., 2009](#); [Björklund et al., 2010](#); [Grönqvist et al., 2010](#)), the US ([Agee and Crocker, 2002](#)), the UK ([Brown et al., 2011](#)), and Germany ([Anger and Heineck, 2010](#); [Anger, 2012](#)). By contrast, the economic literature contains far less evidence on the intergenerational transmission of non-cognitive skills. The transmission of personality traits from parents to children has been examined in the contexts of the US ([Mayer et al., 2004](#); [Duncan et al., 2005](#)), Sweden ([Grönqvist et al., 2010](#)) and Germany ([Anger, 2012](#)).<sup>2</sup>

A number of authors emphasize that estimating intergenerational correlations or elasticities reveals only part of the impact of family background (see, e.g., [Björklund et al., 2010](#); [Björk-](#)

---

<sup>1</sup>See, for example, [Heckman et al. \(2006\)](#) and [Heineck and Anger \(2010\)](#). An extensive overview can be found in [Almlund et al. \(2011\)](#).

<sup>2</sup>Although economic research on non-cognitive skill formation is rather scarce, intergenerational correlations have been analyzed by psychologists for decades (e.g., [Loehlin, 2005](#)). However, the data used in most psychological studies are based on a small number of observations or lack representativeness.

lund and Jäntti, 2012).<sup>3</sup> Instead, researchers suggest estimating sibling correlations, especially for interpretation as an indicator of equality of opportunity. Compared with intergenerational correlations, sibling correlations are a much broader measure of the influence of family. An intergenerational correlation covers only a one-dimensional association between parental and offspring skill measures, a sibling correlation considers all factors that are shared by the siblings of one family.<sup>4</sup> In the context of skill formation, this capability is an important advantage of sibling correlations over intergenerational correlations, as Cunha and Heckman (2007, 2008) suggest that skill formation is dependent not only on parental skills but also on a variety of parental characteristics.

In the existing literature, sibling correlations are used to estimate the influence of family background on educational and labor market outcomes. The results show, for example, that intergenerational correlations explain less than half of the influence of family background on earnings (Mazumder, 2008). Moreover, research provides evidence of remarkable cross-country differences in sibling correlations in education and earnings (Björklund et al., 2002; Schnitzlein, 2014).<sup>5</sup> These cross-country differences might be attributed to different institutional settings in these countries, but the exact mechanisms remain unclear. To the best of our knowledge, existing studies of cognitive and non-cognitive skill correlations within families have covered only the US (Mazumder, 2008) and Sweden (Björklund et al., 2010; Björklund and Jäntti, 2012). Both analyses are based on few skill measures and on only a single skill measurement at one point in time.<sup>6</sup> Moreover, Swedish register data are restricted to males because these data are based on information from military enlistment tests (Björklund and Jäntti, 2012).

In this study, we contribute to the literature in the following ways. First, we estimate sibling correlations in a great variety of cognitive and non-cognitive skill test scores, providing

---

<sup>3</sup>Björklund and Jäntti (2012) call this partial effect the “tip of the iceberg.”

<sup>4</sup>This includes family background and community factors. Among others, Solon et al. (2000), Page and Solon (2003), Leckie et al. (2010), Nicoletti and Rabe (2013) and Lindahl (2011) show that an individual’s family is a more influential factor than his or her neighborhood. Bügelmayer and Schnitzlein (2014) present results on German adolescents suggesting that although the influence of the neighborhood is not negligible in Germany, family background is the predominant factor. Thus, in the following sections, when we speak of shared family background, this discussion includes shared community factors.

<sup>5</sup>For example, using brother correlations, Schnitzlein (2014) reports that approximately 45 percent of the variance in permanent earnings can be attributed to family or neighborhood factors in the US and Germany, whereas the corresponding estimate for Denmark is only 20 percent.

<sup>6</sup>Nicoletti and Rabe (2013) report sibling correlations on exam scores, which are similar in size to cognitive skills but refer to educational attainment.

measures of the importance of family factors to the formation of multiple individual skills. We thus provide evidence based on skill measures that are broader than those used in existing studies. Our data contain test scores from two ultra-short IQ tests that we use as our measure of cognitive skills. Furthermore, our study provides data on the locus of control, reciprocity, and the Big Five personality traits (openness, conscientiousness, extraversion, agreeableness, and neuroticism), which act as our measures of non-cognitive skills. The advantage of the present study is that our data are not restricted to males and that we rely on two repeated measurements of our non-cognitive skill measures.

Second, following the decomposition approach by [Mazumder \(2008\)](#), we investigate the factors that may drive the effect of family background on skill formation. Our rich data enable us to open the black box of family influence on skills by considering potential influence channels that include parental skills, family characteristics, and childhood environment.

Finally, by estimating sibling correlations in cognitive and non-cognitive skills based on representative German survey data, we add the German perspective to the existing literature. This contribution is important, given the cross-country differentials in sibling correlations in education and economic outcomes identified in previous studies. If the estimated sibling correlations in cognitive and non-cognitive skills follow the same cross-country patterns as the estimates for economic outcomes, this outcome would provide insight into the underlying mechanisms of these differentials. Our contribution is therefore to assess the extent to which differences in sibling correlations in skills between countries can explain cross-national differences in the influence of family background on education and labor market outcomes.

To summarize our main results, we show that family background is important for the cognitive and non-cognitive skills of both men and women. Sibling correlations of personality traits range from 0.223 to 0.464, indicating that even for the lowest estimate, one-fifth of the variance or inequality in personality can be attributed to factors shared by siblings. All of the calculated sibling correlations in cognitive skills are higher than 0.50, indicating that more than half of the inequality in cognitive abilities can be explained by family background. Comparing these findings to the results in the literature on intergenerational skill transmission suggests that sibling correlations are indeed able to provide a more complete picture of the influence of family on

children’s cognitive and non-cognitive skills.

Opening the black box of the influence of family background supports this result. Parental skills are important determinants, but including a rich set of family characteristics enhances the explanation of the observed influence of family background. Nevertheless, this rich set of characteristics is able to explain only up to 36 percent of the estimated sibling correlations.

Comparing our results to previous findings for the US and Sweden provides no evidence that the differential in sibling correlations in education and economic outcomes can be explained by differences in cognitive skill formation. The evidence from cross-country comparisons with respect to sibling correlations in non-cognitive skills is less clear.

The remainder of the paper is structured as follows. In the next section, we briefly discuss the existing theoretical model of skill formation. The third section presents our data. The fourth section contains our estimation strategy. Our main results are presented and discussed in section 5, followed by conclusions in the last section.

## 2 Theoretical background

The model of the family as formalized by [Becker and Tomes \(1979, 1986\)](#) underlies most empirical analyses of both intergenerational mobility and sibling correlations in economic outcomes. For the analysis of skill formation, this model has two weaknesses. First, the model includes only a single composite skill measure. Therefore, the complementarity and substitution of different skills cannot be analyzed. Second, in the original model, parental investment and complete skill formation occur in one single period (childhood). This limitation eliminates the possibility that investments in skill formation may be more important during certain periods of childhood than others and that skill production may depend on the existing stock of skills.

[Cunha and Heckman \(2007\)](#) suggest an extension of the model addressing these issues. In their model, an individual’s human capital stock contains *both* cognitive and non-cognitive skills. [Cunha and Heckman \(2007\)](#) present a production function for this aspect of accumulated human capital. According to their model, the vector of cognitive *and* non-cognitive skills ( $\theta$ ) of an individual in period  $(t + 1)$  is a function of the individual’s stock of both cognitive and non-cognitive skills in the previous period  $(t)$ , individual and parental investments in skill formation

in the previous period ( $I_t$ ), and parents' cognitive and non-cognitive skills, as well as other parental or environmental characteristics ( $h$ ):

$$\theta_{t+1} = f_t(\theta_t, I_t, h) \quad (1)$$

Cunha and Heckman (2007) propose that (i)  $\partial f_t(\theta_t, I_t, h)/\partial \theta_t > 0$  and (ii)  $\partial^2 f_t(\theta_t, I_t, h)/\partial \theta_t \partial I_t' > 0$ . Hence, the skill formation process is characterized by a multiplier effect through the (i) self-productivity and (ii) dynamic complementarity of skills. The former mechanism implies that stronger skills in one period create stronger skills in the subsequent period, which is also true across different skills through cross effects. Given the latter mechanism, the productivity of an investment in cognitive and non-cognitive skills is increasing for stronger existing skills. Cunha and Heckman (2008) present empirical evidence corroborating these assumptions; they identify early childhood as the most productive period for investing in cognitive and non-cognitive skills.

This paper focuses on the importance of family background to an individual's skill formation. Family background enters the above production function via two channels. First, the accumulation of cognitive and non-cognitive skills is directly determined by previous parental investments, and second, skill formation depends on the parental stock of cognitive and non-cognitive skills. As we cannot directly observe the arguments in the above function, we apply an indirect approach. If both of these channels – parental investments and the parental stock of cognitive and non-cognitive skills – are important, then siblings should have very similar outcomes because they share the same family background. We estimate sibling correlations in cognitive and non-cognitive skills to assess the similarity in skill levels between siblings. In the second step, we decompose the sibling correlations into different input factors related to individual skill formation. This step allows us to identify channels through which family background affects cognitive and non-cognitive skills.

## 3 Data

### 3.1 Estimation sample

We use data from the German Socio-Economic Panel Study (SOEP), which is a representative household panel survey that began in 1984 (Wagner et al., 2007).<sup>7</sup> The SOEP conducts annual personal interviews with all adult household members and provides rich information on socio-demographic characteristics, family background, and childhood environment on approximately 20,000 individuals in more than 11,000 families in the most recent wave (2012). Measures of cognitive and non-cognitive skills are included for the years 2005 (Big Five, locus of control, reciprocity), 2006 (two cognitive skill tests), 2009 (Big Five), 2010 (locus of control, reciprocity), and 2012 (three cognitive skill tests). Whereas the non-cognitive skill measures are surveyed using the main SOEP questionnaire with all respondents, the ultra-short IQ tests are performed only in computer-assisted personal interviews (CAPIs), which cover approximately one-third of all respondents.<sup>8</sup> This procedure results in a significantly lower number of observations compared with those available for non-cognitive skill measures. Unfortunately, for the repeated measurement in 2012, the sample is divided to conduct three instead of the original two ultra-short IQ tests. Therefore, we use only the 2006 measurement of cognitive skills in this study.

The information on family relations between household members and the follow-up concept of the SOEP allow us to observe children over time and to identify them as siblings even after they grow up and live in different households. In the survey, children must be observed in the same household as their parents only once to be assigned correctly to their mother and father. We consider two children to be siblings if they are assigned to the same father and mother.<sup>9</sup>

---

<sup>7</sup>We use SOEPv29 (DOI: 10.5684/soep.v29). For more information, see <http://www.diw.de/soep>.

<sup>8</sup>Although CAPIs are standard for newer SOEP subsamples, the initial subsamples are still interviewed using PAPI (paper and pencil interviewing).

<sup>9</sup>The SOEP data provide different parental identifiers. In this study, we use the identifiers provided in the SOEP file BIOPAREN. These parental identifiers are based on cohabitation at age 17 (or older if the respondent is older in the first interview). Although the SOEP also provides information on biological children for all women in the survey, information on the biological children of men has been recorded only since 2000. As our sample includes children primarily from the initial SOEP households, which were sampled before 2000, using the biological identifier for men would significantly reduce our sample size. However, we know that for approximately 92 percent of the mother-child pairs in our sample, the social mother is also the biological mother. Thus, nearly all of the siblings studied share at least a biological mother. If genetics are an important factor, then considering social instead of biological parents would result in underestimating the estimated sibling correlations. In this case, our estimates



We include all adult children of SOEP households with identified mothers and fathers who either participated in one of the cognitive tests in 2006 or successfully answered at least one of the question sets on non-cognitive skills in one of the respective waves. We restrict the sample to individuals aged 20 to 54 to avoid the risk of observing noisy skill measures at very young or old ages (Baltes et al., 1999; Cobb-Clark and Schurer, 2012, 2013). Our final sample consists of up to 4,380 individuals from 3,034 families in the non-cognitive skill analysis and 443 individuals from 364 families in cognitive skill analysis.<sup>10</sup>

### 3.2 Cognitive and non-cognitive skill measures

In 2006, information on cognitive skills was collected by measuring test scores from a word fluency test and a symbol correspondence test. Both of these ultra-short tests were developed especially for the SOEP, as full-length IQ tests cannot be incorporated into a large-scale panel survey (Lang et al., 2007). Because the symbol correspondence test is performed using a computer, these tests are conducted only in the CAPI-based subsamples of the SOEP. Both tests correspond to different modules of the Wechsler Adult Intelligence Scale (WAIS) and produce outcomes that are relatively well correlated with test scores from more comprehensive and well-established intelligence tests.<sup>11</sup>

The symbol correspondence test is conceptually related to the mechanics of cognition or fluid intelligence and encompasses general abilities. The test involves asking respondents to match as many numbers and symbols as possible within 90 seconds according to a given correspondence list that is permanently visible to the respondents on a screen.

The word fluency test is conceptually related to the pragmatics of cognition or crystallized intelligence. This test involves the fulfillment of specific tasks that improve in accordance with previously acquired knowledge and skills. The word fluency test implemented in the SOEP is based on the animal-naming task (Lindenberger and Baltes, 1995): respondents name as many different animals as possible within 90 seconds. Whereas verbal fluency is based on learning, speed of cognition is related to an individual's innate abilities (Cattell, 1987). In addition, we could be considered to be a lower bound.

<sup>10</sup>The share of women in our sample is 48 percent. Because there is no theoretical reason to expect differences between sons and daughters with respect to family background effects, we do not separate the analysis by gender.

<sup>11</sup>Lang et al. (2007) conduct reliability analyses and find test–retest coefficients of 0.7 for both the word fluency and symbol correspondence tests.

generate a measure of general intelligence by averaging the two types of ability test scores.<sup>12</sup>

Measures of non-cognitive skills are available for the 2005 survey (Dehne and Schupp, 2007; Richter et al., 2013), and these measures were repeated in 2009 and 2010. The personality measures in the 2005 survey include self-rated measures related to the Five-Factor Model (McCrae and Costa Jr., 2011) and comprise the five basic psychological dimensions (Big Five) – openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism (emotional instability) – each measured using 3 items. In addition, self-rated measures of Locus of Control (7 items) and reciprocity (6 items) are included in the 2005 survey.

Locus of control is the extent to which an individual believes that he or she has control over what happens in his or her life. Psychologists differentiate between external locus of control (i.e., individuals believing that events are largely the result of external effects) and internal locus of control (i.e., individuals believing that events are the results of their own actions). We follow the suggestions by Richter et al. (2013) and use a one-dimensional measure with higher scores representing a more internal locus of control and lower scores representing a more external locus of control.

Reciprocity measures the extent to which an individual is willing to respond to positive or negative behavior. One can distinguish positive reciprocity (i.e., the extent to which individuals respond positively to positive actions) from negative reciprocity (i.e., the extent to which individuals respond negatively to negative behavior).

All items related to non-cognitive skills are answered on 7-point Likert-type scales (1 – “disagree completely” to 7 – “agree completely”). The scores are summed along each dimension to create an index ranging from 1 to 7 and are standardized for each year. In 2009, respondents were repeatedly asked to rate their personality according to the dimensions of the Five-Factor Model. Self-ratings of locus of control and reciprocity were repeated in 2010.

### **3.3 Family background variables**

Our data not only enable the identification of parents and siblings but also provide information on parental characteristics and family background. To identify factors through which family

---

<sup>12</sup>Using average test scores is expected to reduce the error-in-variable bias by diminishing the random component of measured test scores. Furthermore, average test scores could be interpreted as an extract of a general ability type, which captures both coding speed and verbal fluency.

background may affect skills, we use data on parental socio-economic characteristics. In particular, we use information regarding both paternal and maternal years of education, individual labor earnings, and migration background; the mother's age at first birth; whether the family is originally from East Germany; and the total number of children reported by the mother.<sup>13</sup> As measures of parental non-cognitive skills, we include paternal and maternal personality measures from 2005, which are available for approximately half of our sample.<sup>14</sup> Given the small number of available observations with cognitive skill measures that would be further reduced when restricting the sample to observations with non-missing parental characteristics, we are unable to investigate family influence channels for cognitive skills. Hence, we perform the decomposition analysis only for the non-cognitive skill scores.

### 3.4 Descriptive statistics

The descriptive statistics of our main sample are shown in panel A of Table 1, which presents figures for the pooled subsamples for each skill. All skill measures are standardized within the entire population to have a mean of zero and a standard deviation of one for each year.<sup>15</sup>

In addition, the number of observations, the number of individuals, and the number of families are reported separately for each subsample. As we include only one observation for cognitive skills, the number of observations and number of individuals are identical for these outcomes. Dividing the pooled sample based on individual survey years for non-cognitive skills shows that the means are similar for each year (i.e., personality traits within the population change little over time; not displayed in the table).<sup>16</sup>

The descriptive statistics for the subsample with available parental information are presented in panel B of Table 1, which shows virtually the same mean non-cognitive skill test scores as

---

<sup>13</sup>As an earnings measure (in 2007 euros), to reduce measurement error resulting from transitory fluctuations, we use mothers' and fathers' average observed earnings between 25 and 60 years of age. We include years with zero earnings and use (earnings+1) in our calculations. On average, the earnings measure includes approximately 16 years of parental earnings information.

<sup>14</sup>Because of the low number of observations, we cannot include parental cognitive skill measures in the analysis. The effect of cognitive skills will to some extent be captured by parental education.

<sup>15</sup>The displayed means of the skills (particularly those for crystallized intelligence) deviate slightly from zero, as our sample consists of (adult) children who rated some of their personality traits differently and performed better in the cognitive tests than the relatively older generations in the SOEP. This result can be partially explained by age-related cognitive decline and by the so-called Flynn effect, which indicates a rise in average cognitive ability test scores for the last three generations (Flynn, 1994).

<sup>16</sup>Personality traits and locus of control are shown to be relatively stable across the adult lifespan (Cobb-Clark and Schurer, 2012, 2013).

in the main sample. An overview of parental characteristics is shown in Table 2. As shown in the table, mothers and fathers differ slightly in their personality traits. Whereas mothers appear to have a lower internal locus of control and negative reciprocity, they score higher on agreeableness while rating themselves as more neurotic. Both the educational attainment and earnings of mothers are lower than those of fathers. Note that the average number of children is relatively high (2.57), as all women in the sample are mothers (conditional average).

As discussed above, our sample includes only individuals whose parents we can identify. Naturally, as in all analyses of intergenerational mobility or family background, this sample characteristic reduces the number of individuals in the estimation sample. Figures A.1 and A.2 in the appendix show the distributions for our cognitive and non-cognitive skill measures for both the full SOEP sample and our full estimation sample. For all skill measures, the graphs show similar distributions in the two samples. Therefore, our results should not be contaminated by the restriction to individuals with identified parents. This finding is in line with the results obtained by Richter et al. (2014), who find only minor differences in personality traits between SOEP respondents who stay in the survey and those who drop out of the sample.<sup>17</sup>

## 4 Estimation strategy

Let  $y_{ij}$  be a cognitive or non-cognitive test score for child  $j$  of family  $i$ . We assume that this score can be decomposed into two orthogonal components (Solon et al., 1991; Solon, 1999).

$$y_{ij} = \alpha_i + \mu_{ij} \tag{2}$$

where  $\alpha_i$  covers the combined effect of all factors that are shared by siblings from family  $i$  and  $\mu_{ij}$  covers all factors that are purely idiosyncratic to sibling  $j$ . Orthogonality arises because we observe each child in only one family. Therefore, the variance of the observed test score  $\sigma_y^2$  can be expressed as the sum of the variances of the two components:

$$\sigma_y^2 = \sigma_\alpha^2 + \sigma_\mu^2 \tag{3}$$

---

<sup>17</sup>Moreover, because family background is identified based on siblings in our analysis, the question arises as to whether children with siblings and singletons have different cognitive and non-cognitive skills. However, apart from emotional stability and fluid intelligence, which seem slightly lower for children without (identified) siblings in our dataset, both personality traits and cognitive abilities appear to be fairly equal for all family types.

The correlation coefficient  $\rho$  of the skill measure of two siblings  $j$  and  $j'$  then equals the ratio of the variance of the family component  $\sigma_\alpha^2$  to the total variance of the measure  $\sigma_\alpha^2 + \sigma_\mu^2$ :

$$\rho = \text{corr}(y_{ij}, y_{ij'}) = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\mu^2} \quad \text{with } j \neq j' \quad (4)$$

The interpretation of  $\rho$  is that the correlation in skills between two siblings (i.e., the sibling correlation) equals the proportion of the variance (or inequality) in the skills that can be attributed to factors shared by siblings, such as family factors or neighborhood factors.  $\sigma_\alpha^2$  and  $\sigma_\mu^2$  cannot be negative; thus,  $\rho$  can take values between 0 and 1. A correlation of 0 indicates no influence of family and community factors, and 1 indicates no individual influence. The first case would describe a fully mobile society and the latter a fully deterministic one.

[Solon \(1999\)](#) shows that the relationship of the sibling correlation defined above and the often-estimated intergenerational correlation is as follows:

$$\rho_{\text{skill}} = IGC_{\text{skill}}^2 + \text{other shared factors uncorrelated with the parental skill measure} \quad (5)$$

The sibling correlation in a specific cognitive or non-cognitive skill equals the square of the intergenerational correlation in this skill plus the influence of all shared factors that are uncorrelated with the corresponding parental skill measure. Although sibling correlation is a much broader measure of family background than intergenerational correlation is, sibling correlation is still a lower bound of the true influence of family background, as some family-related factors are not shared by siblings (see the discussion in [Björklund and Jäntti, 2012](#)).

Following [Mazumder \(2008\)](#), we estimate the sibling correlation in our skill measures as the within-cluster correlation in the following linear multilevel model:

$$y_{ijt} = \beta X_{ijt} + \alpha_i + \mu_{ij} + \nu_{ijt} \quad (6)$$

with  $y_{ijt}$  being an annual ( $t$ ) observation of a specific outcome,  $X_{ijt}$  being a matrix of fixed year, age and gender effects (including year dummies, age, age<sup>2</sup>, and a gender dummy as well as interaction terms of the gender dummy and the age variables), the family component

( $\alpha_i$ ), the individual component ( $\mu_{ij}$ ), and a transitory component ( $\nu_{ijt}$ ). The sum of the family and individual components represents the permanent part of the observed outcome. We apply restricted maximum likelihood (REML) to estimate this model and to estimate the variances of  $\alpha_i$  and  $\mu_{ij}$ . The standard error for the sibling correlation is calculated using the delta method. For specifications with only one observation in time (cognitive skill test scores), the model is estimated with only two levels.

To identify the relative importance of different inputs in the skill formation process, we follow the decomposition approach suggested by [Mazumder \(2008\)](#). We add family background characteristics as explanatory variables to equation (6). If these characteristics are important determinants of the formation of the respective skill, this should decrease the variance of the family-specific component and its relative importance and therefore reduce the sibling correlation. This reduction can be considered an upper bound estimate of the importance of the additional family background characteristics.

## 5 Results

### 5.1 Sibling correlations in cognitive and non-cognitive skills

We begin the discussion of our results with the measure of cognitive skills. [Figure 1](#) shows the estimated sibling correlations and the corresponding standard errors, and [Table A.1](#) shows the underlying basic estimates for this figure, including the variance of the family and individual components. We find a strong influence of family background on all three dimensions of cognitive abilities. The strongest sibling correlation can be found for crystallized intelligence, with a coefficient of 0.607, whereas the sibling correlation in fluid intelligence is slightly lower, at 0.547. The sibling correlation in general intelligence lies between these figures, at 0.578. Hence, family and community background explains more than 50 percent of the variation in cognitive test scores between individuals, and this result applies to both types of cognitive skills: those related to innate abilities and those based on learning. Even compared to a value of 0.45 for sibling correlations in earnings in Germany ([Schnitzlein, 2014](#)), these coefficients are considerably large.

[Figure 2](#) and [Table A.2](#) show the results for non-cognitive skills. The highest sibling corre-

lation is estimated for locus of control, which shows a coefficient of 0.464. This result indicates another strong effect of family background on skills, as forty-six percent of the variation in locus of control can be attributed to factors shared by siblings. The corresponding estimates for positive and negative reciprocity are 0.434 and 0.383, respectively, which still indicate substantial influences of family background on personality traits. The estimates for Big Five personality traits show greater variation. Whereas shared background factors appear to be important for conscientiousness (0.412), the estimated sibling correlation in extraversion is only 0.223. Agreeableness (0.349), openness (0.293) and neuroticism (0.308) fall between those figures. Hence, even if the difference between sibling correlations in cognitive abilities and locus of control is rather small, family and community background appear to explain more of the variation in cognitive skills than that in non-cognitive skills. One possible explanation is that measurement error is higher when measuring non-cognitive skills than when measuring cognitive skills (Grönqvist et al., 2010), which would imply that our sibling correlations in non-cognitive skills are a lower bound of the true influence of family background.

As shown in equation (5), sibling correlations cover a much greater influence of family background than intergenerational correlations do. As argued in the introduction, this greater influence is one explanation for why sibling correlations are a preferable measure to assess equality of opportunity. In Figure 3, we draw on the intergenerational skill correlations reported by Anger (2012), who uses the same dataset and outcomes that we use.<sup>18</sup> For all analyzed outcomes, the estimated sibling correlations are considerably higher than the corresponding (squared) intergenerational correlations. This finding suggests that intergenerational correlations are actually able to capture only some of the influence of family on children's cognitive and non-cognitive skills. This result is in line with findings in the literature on educational and income mobility.

In summary, we showed that family background has a significant and usually substantial influence on an individual's cognitive and non-cognitive skills. As these skills are important determinants of economic success, this finding indicates that the normative goal of equality of opportunity is violated.

---

<sup>18</sup>Note that Anger (2012) does not report results for reciprocity.

## 5.2 Decomposition of the influence of family background

As [Cunha and Heckman \(2007\)](#) show, the formation of skills is affected by different input factors. In this section, we provide insight into the question regarding which channels are most important in determining the influence of family on non-cognitive skill formation. As noted previously, we must restrict our decomposition analysis to the formation of non-cognitive skills because of the limited number of observations for cognitive test scores.

In the first step, we estimate sibling correlations for different subgroups of our estimation sample to investigate whether the family effect differs by the socio-economic status of the family. [Table 3](#) shows the results divided by family income and mother's education as well as the results for the full sample and for the subsample of individuals with non-missing parental characteristics for comparison. Siblings with high-income parents<sup>19</sup> show higher sibling correlations with respect to locus of control and in four of the five Big Five personality traits, indicating a stronger family influence for these siblings than for those from low-income families. Although the estimated sibling correlations for neuroticism are virtually the same for high- and low-income families, the influence of family background on both measures of reciprocity is greater for low-income families than for high-income families. Furthermore, children of highly educated mothers<sup>20</sup> show higher sibling correlations in most outcomes, thus indicating a greater influence of family background on skill formation. One notable exception is the estimated sibling correlations for neuroticism. Here, the sibling correlation for families with a less educated mother is more than twice as large as the estimate for families with a highly educated mother. Overall, the results in [Table 3](#) indicate that the influence of family on non-cognitive skills differs for various family types, with most outcomes showing a stronger influence of families with higher socio-economic status. This result may indicate that the skill formation of children from low-SES families is more idiosyncratic than those from higher-SES families.

Next, we provide insight into the question regarding which parental characteristics best explain the influence of family background on skill formation. [Table 4](#) shows the results of the

---

<sup>19</sup>We use the sum of the mother's and father's average individual labor earnings as defined above. Families above the median are labeled as high-income families.

<sup>20</sup>Mothers with at least 12 years of education are defined as highly educated, including all mothers who have at least an intermediate secondary degree plus a vocational school degree.



decomposition approach described in section 4. The first column shows the estimated sibling correlations in non-cognitive skills for the full estimation sample, and the second column shows the estimated sibling correlations for the subsample with non-missing parental characteristics. Overall, the sibling correlations are similar in both samples.

The middle part of Table 4 presents the results of our decomposition. In the third column, we add the respective parental (father's and mother's) non-cognitive skills as explanatory variables in equation (6).<sup>21</sup> The resulting decline in the estimated sibling correlation indicates the importance of the respective parental skill in the influence of family on the skill formation process. In the fourth column, instead of parental skills, we add parental education by including both the father's and mother's education in the model. Parental education serves as both an indicator of parental resources and an indicator of parental cognitive skills. Although the inclusion of parental education has little effect on the size of most sibling correlations in non-cognitive skills, adding the respective parental skill clearly reduces the family influence that can be attributed to the remaining factors shared by siblings. Finally, in the fifth column, we add the full set of parental characteristics (as presented in Table 2) to our model; for most outcomes, the inclusion of these characteristics leads to further decreases in the remaining sibling correlations.

For ease of interpretation, the right-hand side of Table 4 shows the respective percentage reduction in the estimated sibling correlation for each of these decompositions. The results yield two important insights. First, for all outcomes, the corresponding parental skill is the most important of all observed family characteristics. Moreover, including the full set of parental characteristics still contributes to explaining the observed sibling correlations for most outcomes. Second, even our rich set of parental characteristics is able to capture only up to 36 percent of the influence of family background as measured by the estimated sibling correlations. Although we would like to further investigate possible channels by including more family background and childhood environment variables, we cannot do so because interpreting the decomposition requires relying on factors that are truly shared by siblings and thus are not sibling specific.<sup>22</sup>

---

<sup>21</sup>See Table A.3 for the decomposition results when only the father's or the mother's characteristics are included. The separate decompositions yield similar results for the inclusion of the father's and mother's characteristics. However, including both parents' characteristics clearly best explains the influence of family background on skill formation.

<sup>22</sup>For example, we have information on whether an individual's parents divorced during her childhood or whether she spent her childhood in a rural or an urban area. However, these factors may differ between – and

Because sibling-specific family factors are most likely to be important determinants in the skill formation process, our sibling correlations provide a lower bound for the true influence of family background on skills.

In summary, we show that the influence of family background on an individual's skill formation process differs based on parental socio-economic status. Our decomposition approach reveals that parental skills are important factors in determining the influence of family background, but as suggested above, only considering the skills of the parents' generation does not capture the full picture. In addition, our results show that even a rich set of parental characteristics captures no more than 36 percent of the influence of family on the skill formation process. This result points to the importance of sibling-specific factors of family and neighborhood (i.e., factors that are not shared by siblings).

### 5.3 Cross-national comparisons

Next, we discuss our findings relative to the existing evidence in the literature for the US and Sweden. Any differences in the influence of family background on cognitive and non-cognitive skills may help to explain the observed cross-country differences in the importance of family background for economic outcomes.

Björklund et al. (2002) and Schnitzlein (2014) report significant cross-national differences in sibling correlations in earnings. In particular, those researchers find that family background is more important in the US and Germany than in the Scandinavian countries. In the US and Germany, approximately 45 percent of the variance in earnings can be attributed to family factors, whereas this share is only 20 percent in Denmark based on brother correlations (Schnitzlein, 2014).<sup>23</sup>

Sibling correlations in cognitive skills are reported by Mazumder (2008), who finds coefficients of approximately 0.6 for the US. Hence, compared with the estimates presented in Table A.1, the influence of family background on the formation of cognitive skills in the US context

---

hence would not be shared by – siblings of different ages.

<sup>23</sup>Cross-country differences in the importance of family background are also found for educational attainment. In Nordic countries, approximately 45 percent of the variance in education can be attributed to family and neighborhood (Raaum et al., 2006; Lindahl, 2011), whereas this share is more than 60 percent in Germany (Schnitzlein, 2014) and up to 70 percent in the US (Mazumder, 2011).

is only slightly different from the German context.<sup>24</sup> Björklund and Jäntti (2012) find brother correlations of approximately 0.5 for cognitive skills in Swedish data. Again, these estimates differ only slightly from those presented in Table A.1. Based on these findings, we find no evidence that differences in the influence of family background on cognitive skill formation can explain the cross-country differentials in sibling correlations in earnings.

With respect to non-cognitive skills, Mazumder (2008) uses the Rotter scale for locus of control and finds sibling correlations of 0.11 for brothers and 0.07 for sisters. These estimates are much lower than those presented in Table A.2 for Germany. However, the Rotter questionnaire in the NLSY used by Mazumder (2008) is only a four-item version of the original 60-item scale (instead of 7 items in the SOEP), and each item is scored on a scale of only 1 to 4 (instead of 1 to 7 in the SOEP), which may yield both larger measurement error and attenuation bias. In addition, Mazumder (2008) has only one skill observation available in the data and therefore cannot control for transitory fluctuations.<sup>25</sup> Solon et al. (1991) show that using multiple measurements reduces transitory fluctuations and measurement error that lead to the underestimation of the sibling correlation.<sup>26</sup> Table A.4 in the appendix shows this effect for our non-cognitive skill measures. The estimates in columns (1) - (3) are based on a balanced panel of individuals with two valid observations in both waves in which non-cognitive skills were included in the SOEP (2005, 2009/2010). Columns (1) and (2) show the estimated sibling correlations using only single-year measures for either 2005 (column 1) or 2009/2010 (column 2) without controlling for transitory fluctuations.<sup>27</sup> Column (3) contains estimates based on two observations.<sup>28</sup> All estimated sibling correlations based on single-year measures are clearly lower than those presented in column (3). However, even our single-year estimates for locus of control are higher than those reported in Mazumder (2008).

Björklund and Jäntti (2012) present the second available estimate in the literature for sibling correlations in non-cognitive skills. They use an aggregate measure of leadership skills derived

---

<sup>24</sup>However, it may be problematic to directly compare these results to those presented by Mazumder (2008), as he uses a different measure of cognitive skills (AFQT test scores).

<sup>25</sup>As shown in the last row of Table A.2, the variance of the transitory component is of substantial size in all estimations.

<sup>26</sup>This result is similar to the findings of Solon (1989, 1992) and Zimmerman (1992) for intergenerational mobility estimates.

<sup>27</sup>In these cases, the model in equation (6) is estimated without the transitory component.

<sup>28</sup>Column (4) contains estimates based on the full sample.

from interviews during a military enlistment test. These researchers report a brother correlation of 0.3, which falls within the range of sibling correlation for personality traits revealed by our estimates for Germany. Again, however, this measure is not directly comparable to our measures of non-cognitive skills.

To summarize these cross-national comparisons, our results on cognitive skills are remarkably similar to existing results for the US and Sweden. We find no evidence that differences in the influence of family background on cognitive skills explain differences in the importance of family background for economic success. The picture for non-cognitive skills is less clear because the different measures used are not fully comparable.

## 6 Conclusion

In this study, we investigate the importance of family background for cognitive and non-cognitive skills based on sibling correlations in order to provide a measure of the role of family in the skill formation process that is broader than the previously used intergenerational correlations. Our estimates are based on data from SOEP, which is a large representative household survey that provides measures of cognitive skills from two ultra-short IQ tests, as well as self-rated measures of locus of control, reciprocity, and the Big Five personality traits. Previous analyses for Sweden and the US are restricted because they are based only on males (Björklund et al., 2010; Björklund and Jäntti, 2012) and/or use few non-cognitive skill measures (i.e., locus of control (Mazumder, 2008) and leadership skills (Björklund and Jäntti, 2012)) and are based on only a single measurement at one point in time. Hence, our study contributes to the literature by providing evidence on sibling correlations using broader measures and repeated measurements of skills and by including both men and women.

We show that family background is important for cognitive and non-cognitive skill formation. Sibling correlations of personality traits range from 0.223 to 0.464, indicating that even for the lowest estimate, more than one-fifth of the variance or inequality in non-cognitive skills can be attributed to factors shared by siblings. All calculated sibling correlations for cognitive skills are higher than 0.50, indicating that more than half of the inequality can be explained by shared family background. Comparing these findings to the results in the intergenerational skill

transmission literature suggests that sibling correlations are indeed able to provide a more complete picture of the influence of family on children's cognitive and non-cognitive skills. This result is in line with findings in the literature on educational and income mobility.

Opening the black box of the influence of family background supports this result. Our decomposition analyses show that parental skills are the most important influencing factors, but including a rich set of family characteristics enhances the explanation of the observed influence of family background for most outcomes.

Comparing our results to previous findings for the US and Sweden provides no evidence that the differential in sibling correlations in economic outcomes can be explained by differences in the formation of cognitive skills. The evidence from cross-country comparisons with respect to sibling correlations in non-cognitive skills is less clear.

## **Acknowledgments**

We would like to thank Anders Björklund, Markus Jäntti, Matthew Lindquist, Shelly Lundberg, Bhashkar Mazumder, and Catherine Weinberger; seminar participants of SOFI in Stockholm, UC Santa Barbara, ISER at the University of Essex, RWI Essen, the University of Hamburg, the University of Bath, the University of Bristol and The Danish National Centre for Social Research; and conference participants at the Annual Conference of the Scottish Economic Society 2013, SOLE 2013, ESPE 2013, IWAE 2013, SMYE 2013, the 2013 Annual conference of the German Economic Association, and EALE 2013 for their useful comments and discussions.

## References

- Agee, M. D., Crocker, T. D., 2002. Parents' discount rate and the intergenerational transmission of cognitive skills. *Economica* 69, 143–154.
- Almlund, M., Duckworth, A. L., Heckman, J., Kautz, T., 2011. Personality psychology and economics. In: Hanushek, E. A., Machin, S., Woessmann, L. (Eds.), *Handbook of the Economics of Education*. Vol. 4. Elsevier, Ch. 1, pp. 1–181.
- Anger, S., 2012. Intergenerational transmission of cognitive and noncognitive skills. In: Ermisch, J., Jäntti, M., Smeeding, T. (Eds.), *From Parents to Children: The Intergenerational Transmission of Advantage*. Russell Sage Foundation, New York, pp. 393–421.
- Anger, S., Heineck, G., 2010. Do smart parents raise smart children? the intergenerational transmission of cognitive abilities. *Journal of Population Economics* 23 (3), 1105–1132.
- Baltes, P. B., Staudinger, U. M., Lindenberger, U., 1999. Lifespan psychology: Theory and application to intellectual functioning. *Annual Review of Psychology* 50, 471–507.
- Becker, G. S., Tomes, N., 1979. An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of Political Economy* 87 (6), 1153–1189.
- Becker, G. S., Tomes, N., 1986. Human capital and the rise and fall of families. *Journal of Labor Economics* 4 (3), 1–39.
- Björklund, A., Eriksson, T., Jäntti, M., Raaum, O., Österbacka, E., 2002. Brother correlations in earnings in Denmark, Finland, Norway and Sweden compared to the United States. *Journal of Population Economics* 15 (4), 757–772.
- Björklund, A., Hederos Eriksson, K., Jäntti, M., 2010. IQ and family background: Are associations strong or weak? *The B.E. Journal of Economic Analysis & Policy* 10 (1), Article 2.
- Björklund, A., Jäntti, M., 2012. How important is family background for labor-economic outcomes? *Labour Economics* 19 (4), 465–474.
- Black, S. E., Devereux, P., 2011. Recent developments in intergenerational mobility. In: Ashenfelter, O., Card, D. (Eds.), *Handbook of Labor Economics*. Vol. 4B. Elsevier, pp. 1487–1542.
- Black, S. E., Devereux, P. J., Salvanes, K. G., 2009. Like father, like son? A note on the intergenerational transmission of IQ scores. *Economics Letters* 105, 138–140.
- Brown, S., McIntosh, S., Taylor, K., 2011. Following in your parents' footsteps? empirical analysis of matched parent-offspring test scores. *Oxford Bulletin of Economics and Statistics* 73 (1), 40–58.
- Bügelmayer, E., Schnitzlein, D. D., 2014. Is it the family or the neighborhood? evidence from sibling and neighbor correlations in youth education and health. Report, SOEP Papers No. 716, DIW Berlin.
- Cattell, R. B., 1987. *Intelligence: Its structure, growth, and action*. Elsevier Science, New York.
- Cobb-Clark, D. A., Schurer, S., 2012. The stability of big-five personality traits. *Economics Letters* 115 (1), 11–15.

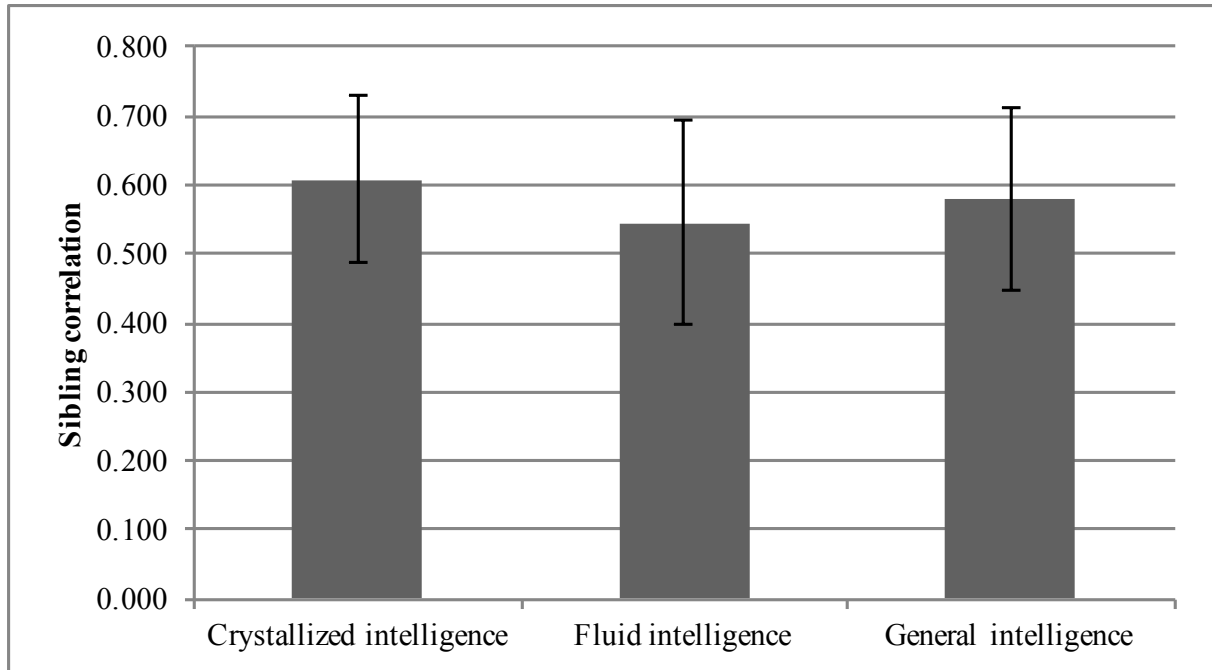
- Cobb-Clark, D. A., Schurer, S., 2013. Two economists' musings on the stability of locus of control. *The Economic Journal* 123 (570), F358–F400.
- Cunha, F., Heckman, J. J., 2007. The technology of skill formation. *American Economic Review* 97 (2), 31–47.
- Cunha, F., Heckman, J. J., 2008. A new framework for the analysis of inequality. *Macroeconomic Dynamics* 12 (S2), 315–354.
- Dehne, M., Schupp, J., 2007. Persönlichkeitsmerkmale im Sozio-oekonomischen panel (SOEP) - Konzept, Umsetzung und empirische Eigenschaften.
- Duncan, G., Kalil, A., Mayer, S. E., Tepper, R., Payne, M. R., 2005. The apple does not fall far from the tree. In: Bowles, S., Gintis, H., Osborne Groves, M. (Eds.), *Unequal Chances: Family Background and Economic Success*. Russell Sage Foundation, Princeton University Press.
- Flynn, J. R., 1994. IQ gains over time. In: Sternberg, R. J. (Ed.), *Encyclopedia of human intelligence*. Macmillan, New York, pp. 617–623.
- Grönqvist, E., Öckert, B., Vlachos, J., 2010. The intergenerational transmission of cognitive and non-cognitive abilities.
- Heckman, J. J., Stixrud, J., Urzua, S., 2006. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics* 24 (3), 411–482.
- Heineck, G., Anger, S., 2010. The returns to cognitive abilities and personality traits in germany. *Labour Economics* 17 (3), 535–546.
- Lang, F. R., Weiss, D., Stocker, A., von Rosenblatt, B., 2007. Assessing cognitive capacities in computer-assisted survey research: Two ultra-short tests of intellectual ability in the German Socio-Economic Panel. *Schmollers Jahrbuch* 127, 183–191.
- Leckie, G., Pillinger, R., Jenkins, J., Rasbash, J., 2010. School, family, neighbourhood: which is most important to a child's education? *Significance* 7 (2), 67–70.
- Lindahl, L., 2011. A comparison of family and neighborhood effects on grades, test scores, educational attainment and income—evidence from Sweden. *The Journal of Economic Inequality* 9 (2), 207–226.
- Lindenberger, U., Baltes, B., 1995. Kognitive Leistungsfähigkeit im Alter: Erste Ergebnisse aus der Berliner Altersstudie. *Zeitschrift für Psychologie* 203 (4), 283–317.
- Loehlin, J. C., 2005. *Resemblance in Personality and Attitudes between Parents and Their Children: Genetic and Environmental Contributions*. Princeton University Press, Princeton.
- Mayer, S. E., Duncan, G., Kalil, A., 2004. Like mother, like daughter? SES and the intergenerational correlation of traits, behaviors and attitudes.
- Mazumder, B., 2008. Sibling similarities and economic inequality in the US. *Journal of Population Economics* 21 (3), 685–701.

- Mazumder, B., 2011. Family and community influences on health and socioeconomic status: Sibling correlations over the life course. *B.E. Journal of Economic Analysis & Policy (Contributions)* 11 (3), Article 1.
- McCrae, R. R., Costa Jr., P. T., 2011. *The Five-Factor Theory of Personality*, 3rd Edition. Guilford, New York.
- Nicoletti, C., Rabe, B., 2013. Inequality in pupils' test scores: How much do family, sibling type and neighbourhood matter? *Economica* 80 (318), 197–218.
- Page, M., Solon, G., 2003. Correlations between brothers and neighboring boys in their adult earnings: The importance of being urban. *Journal of Labor Economics* 21 (4), 831–855.
- Raaum, O., Salvanes, K. G., Sorensen, E. O., 2006. The neighbourhood is not what it used to be. *The Economic Journal* 116 (508), 200–222.
- Richter, D., Körtner, J. L., Saßenroth, D., 2014. Personality has minor effects on panel attrition. *Journal of Research in Personality* 53, 31–35.
- Richter, D., Metzging, M., Weinhardt, M., Schupp, J., 2013. *Soep scales manual*. Report, SOEP Survey Papers Series C - Data Documentations No 138, DIW Berlin.
- Roemer, J. E., 1998. *Equality of opportunity*. Harvard University Press.
- Schnitzlein, D. D., 2014. How important is the family? Evidence from sibling correlations in permanent earnings in the US, Germany and Denmark. *Journal of Population Economics* 27 (1), 69–89.
- Solon, G., 1989. Biases in the estimation of intergenerational earnings correlations. *The Review of Economics and Statistics* 71 (1), 172–174.
- Solon, G., 1992. Intergenerational income mobility in the united states. *The American Economic Review* 82 (3), 393–408.
- Solon, G., 1999. Intergenerational mobility in the labor market. In: Ashenfelter, O., Card, D. (Eds.), *Handbook of Labor Economics*. Vol. 3A. Elsevier, pp. 1761–1800.
- Solon, G., Corcoran, M., Gordon, R., Laren, D., 1991. A longitudinal analysis of sibling correlations in economic status. *The Journal of Human Resources* 26 (3), 509–534.
- Solon, G., Page, M. E., Duncan, G. J., 2000. Correlations between neighboring children in their subsequent educational attainment. *Review of Economics and Statistics* 82 (3), 383–392.
- Wagner, G. G., Frick, J. R., Schupp, J., 2007. The German socio-economic panel study (SOEP): Scope, evolution and enhancements. *Schmollers Jahrbuch* 127 (1), 139–169.
- Zimmerman, D. J., 1992. Regression toward mediocrity in economic stature. *American Economic Review* 82 (3), 409–429.



## Figures and tables

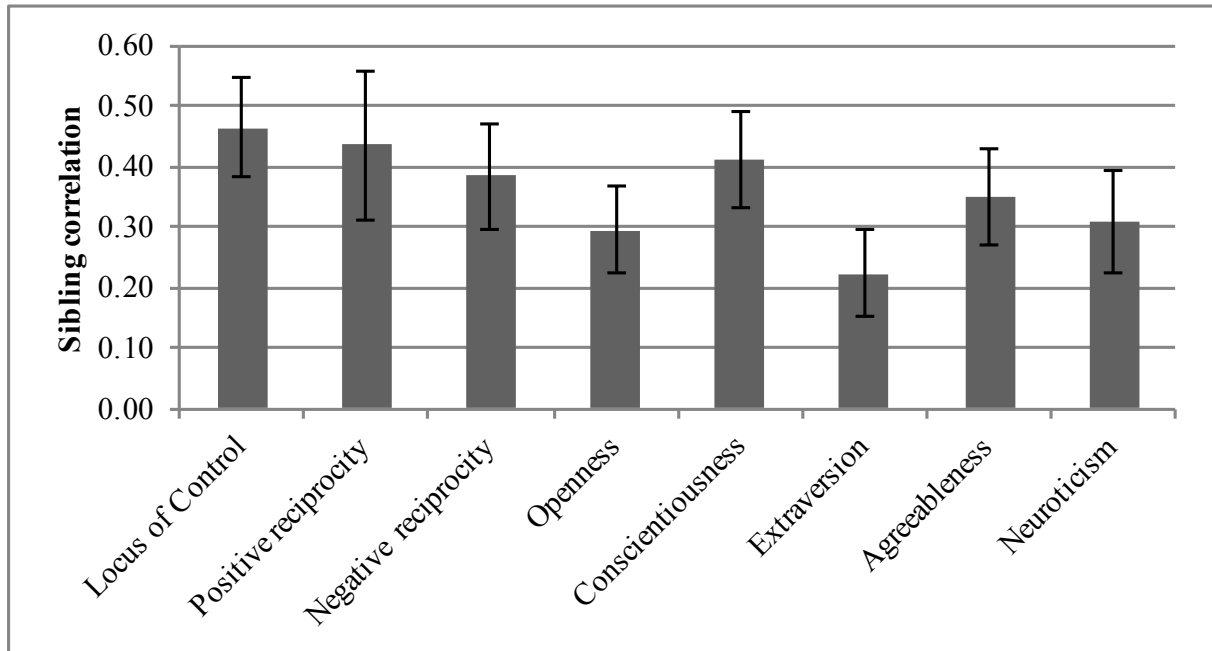
Figure 1: Sibling correlations in cognitive skills



**Note:** Sibling correlations for cognitive skill measures and standard errors are presented. The models are estimated via REML. Standard errors of the sibling correlations are calculated via the delta method. All estimations control for fixed aged profiles (age and age squared), a gender dummy and interactions of the gender dummy as well as polynomials of age.

**Source:** SOEPv29.

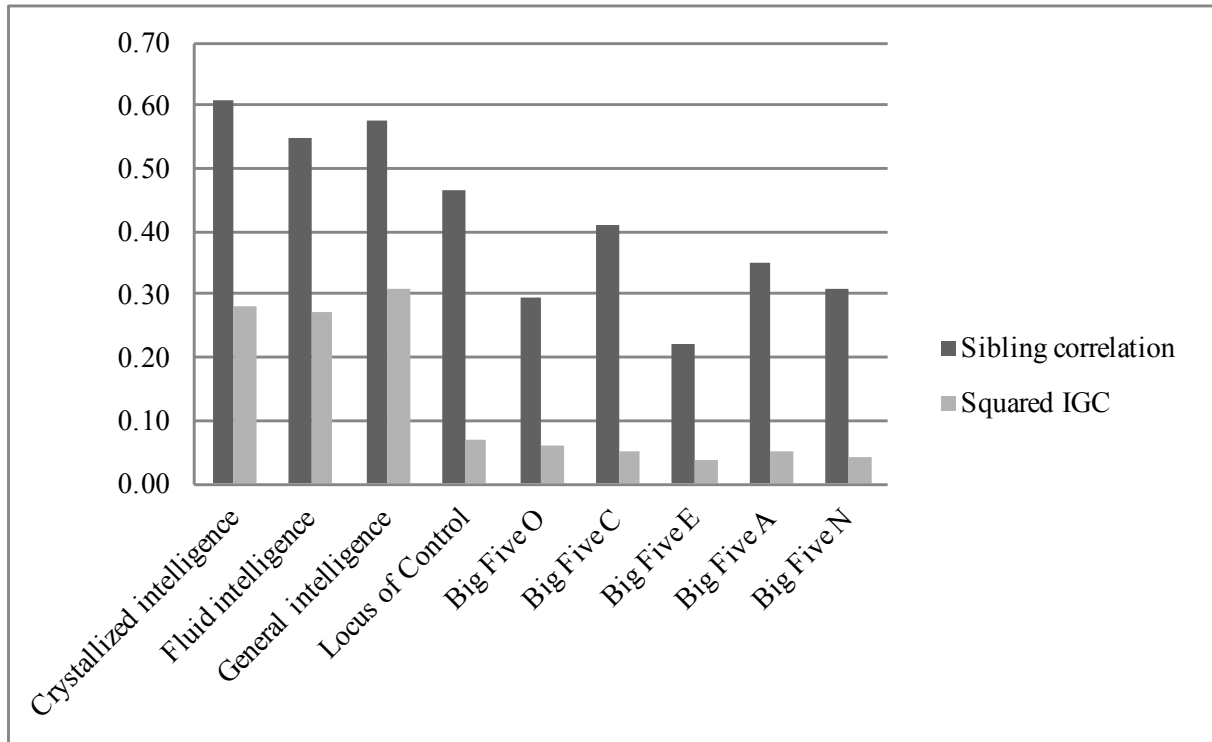
Figure 2: Sibling correlations in non-cognitive skills



**Note:** Sibling correlations for non-cognitive skill measures are presented. The models are estimated via REML. Standard errors of the sibling correlations are calculated via the delta method. All estimations control for fixed aged profiles (age and age squared), a survey year dummy, and a gender dummy as well as interactions of the gender dummy and polynomials of age.

**Source:** SOEPv29.

Figure 3: Comparison of sibling and intergenerational correlations



**Note:** Sibling correlations and squared intergenerational correlations for cognitive and non-cognitive skills are presented.

**Source:** SOEPv29. Intergenerational correlations are obtained from [Anger \(2012\)](#).

Table 1: Descriptive statistics - Main sample and sample with parental characteristics

<b>Outcome</b>	<b>Mean</b>	<b>Min</b>	<b>Max</b>	<b>N of Obs.</b>	<b>N of Ind.</b>	<b>N of Fam.</b>
<i>A: Main sample</i>						
<i>Cognitive skills</i>						
Crystallized intelligence	0.558	-1.812	2.962	443	443	364
Fluid intelligence	0.159	-2.135	3.118	443	443	364
General intelligence	0.342	-1.890	2.572	443	443	364
<i>Non-cognitive skills</i>						
Locus of Control	0.047	-3.660	2.407	6,290	4,352	3,014
Positive reciprocity	-0.067	-5.332	1.276	6,346	4,380	3,034
Negative reciprocity	0.151	-1.442	2.780	6,346	4,380	3,034
Openness	0.073	-2.895	2.131	6,415	4,237	2,942
Conscientiousness	-0.215	-4.866	1.241	6,415	4,237	2,942
Extraversion	0.085	-3.373	1.947	6,415	4,237	2,942
Agreeableness	-0.100	-4.554	1.690	6,415	4,237	2,942
Neuroticism	-0.070	-2.425	2.604	6,415	4,237	2,942
<i>B: Sample with parental characteristics</i>						
<i>Non-cognitive skills</i>						
Locus of Control	0.070	-3.660	2.407	4,126	2,866	1,891
Positive reciprocity	-0.078	-5.332	1.276	4,149	2,877	1,898
Negative reciprocity	0.162	-1.442	2.780	4,149	2,877	1,898
Openness	0.096	-2.895	2.131	4,181	2,778	1,837
Conscientiousness	-0.259	-4.866	1.241	4,181	2,778	1,837
Extraversion	0.118	-3.373	1.947	4,181	2,778	1,837
Agreeableness	-0.101	-4.554	1.690	4,181	2,778	1,837
Neuroticism	-0.099	-2.425	2.604	4,181	2,778	1,837

**Note:** The table shows descriptive statistics for our main sample and for the subsample with non-missing parental characteristics. The sample with cognitive skills is obtained from the 2006 wave, and the sample with non-cognitive skills is obtained from the 2005 and 2009 (Big Five) and 2005 and 2010 (locus of control, reciprocity) waves.

**Source:** SOEPv29.

Table 2: Descriptive statistics for parental characteristics

	<b>A: Mothers' characteristics</b>			<b>B: Fathers' characteristics</b>		
	<b>Locus of Control</b>	<b>Reciprocity</b>	<b>Big Five</b>	<b>Locus of Control</b>	<b>Reciprocity</b>	<b>Big Five</b>
Locus of Control	-0.156	-0.154	-0.160	-0.002	0.000	-0.001
Positive reciprocity	0.039	0.041	0.028	0.036	0.035	0.035
Negative reciprocity	-0.131	-0.133	-0.134	0.111	0.114	0.108
Openness	-0.054	-0.052	-0.058	-0.144	-0.147	-0.151
Conscientiousness	0.153	0.154	0.154	0.073	0.071	0.071
Extraversion	-0.016	-0.013	-0.024	-0.166	-0.167	-0.172
Agreeableness	0.183	0.182	0.188	-0.239	-0.239	-0.239
Neuroticism	0.205	0.206	0.202	-0.048	-0.050	-0.041
Years of education	11.79	11.79	11.77	12.41	12.41	12.40
(log) Earnings	7.812	7.810	7.785	9.940	9.939	9.919
East German	0.284	0.284	0.284	0.281	0.282	0.281
Migration background	0.186	0.186	0.190	0.187	0.186	0.190
Number of kids	2.570	2.570	2.576	-	-	-
Age at first birth	23.68	23.67	23.63	-	-	-
N of Obs.	4,126	4,149	4,181	4,126	4,149	4,181
N of Ind.	2,866	2,877	2,778	2,866	2,877	2,778
N of Fams.	1,891	1,898	1,837	1,891	1,898	1,837

**Note:** The table shows descriptive statistics for parental characteristics. Each column represents the subsample corresponding to one of the children's non-cognitive skill measures (locus of control, reciprocity, and Big Five personality traits).

**Source:** SOEPv29.

Table 3: Sibling correlations by parental background

	Locus of Control	Positive reciprocity	Negative reciprocity	Big Five O	Big Five C	Big Five E	Big Five A	Big Five N
High income family	0.512 (0.075)	0.445 (0.123)	0.295 (0.084)	0.302 (0.064)	0.507 (0.064)	0.250 (0.062)	0.352 (0.072)	0.327 (0.073)
Low income family	0.473 (0.067)	0.525 (0.109)	0.525 (0.064)	0.276 (0.062)	0.368 (0.071)	0.220 (0.062)	0.337 (0.064)	0.328 (0.072)
High educated mother	0.537 (0.089)	0.515 (0.129)	0.461 (0.093)	0.365 (0.071)	0.422 (0.075)	0.223 (0.074)	0.374 (0.080)	0.187 (0.081)
Low educated mother	0.462 (0.060)	0.458 (0.104)	0.404 (0.062)	0.236 (0.056)	0.456 (0.061)	0.241 (0.054)	0.324 (0.059)	0.420 (0.064)
All ind. with nonmis, par. characteristics	0.495 (0.049)	0.477 (0.081)	0.434 (0.051)	0.291 (0.044)	0.447 (0.047)	0.232 (0.044)	0.337 (0.048)	0.329 (0.051)
Full sample	0.464 (0.042)	0.434 (0.063)	0.383 (0.044)	0.293 (0.037)	0.412 (0.041)	0.223 (0.037)	0.349 (0.041)	0.308 (0.043)

**Note:** Sibling correlations for non-cognitive skill measures are presented. The models are estimated via REML. Standard errors of the sibling correlations are calculated via the delta method. All estimations control for fixed aged profiles (age and age squared) and a gender dummy as well as interactions of the gender dummy and polynomials of age. Samples are divided based on family income (sum of the father's and mother's average individual labor earnings as defined in section 3), with high-income families being above the median and maternal education indicating that the higher education of the mother corresponds to 12 or more years of education). All estimates are significant at least at the 5 percent level.

**Source:** SOEPv29.

Table 4: Decomposition of sibling correlations in non-cognitive skills

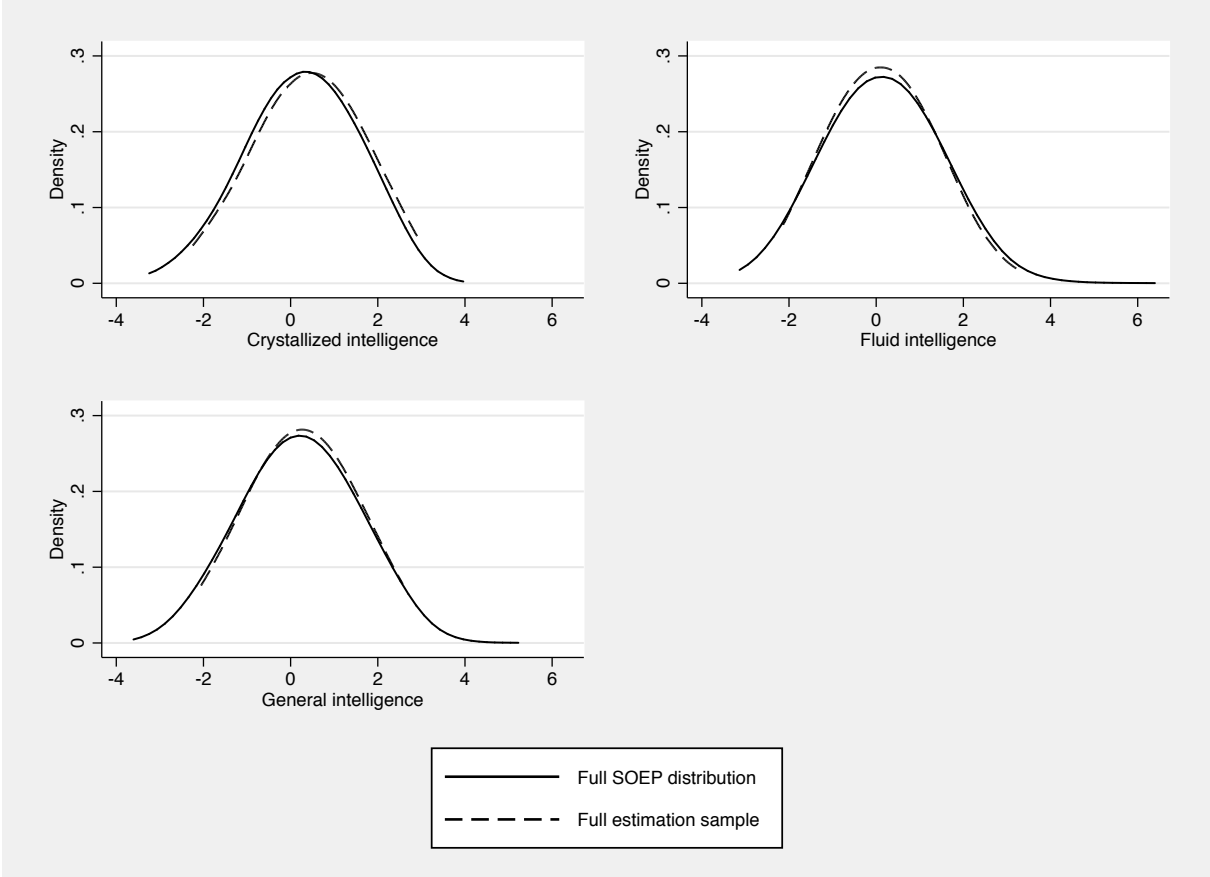
	Full sample (1)	Sample with par. char. (2)	controlling for			percentage reduction		
			parental skill (3)	parental education (4)	all par. char. (5)	parental skill (6)	parental education (7)	all par. char. (8)
Locus of Control	0.464 (0.042)	0.495 (0.049)	0.375 (0.055)	0.476 (0.050)	0.354 (0.057)	24%	4%	28%
Positive reciprocity	0.434 (0.063)	0.477 (0.081)	0.358 (0.093)	0.478 (0.081)	0.345 (0.095)	25%	0%	28%
Negative reciprocity	0.383 (0.044)	0.434 (0.051)	0.308 (0.057)	0.422 (0.052)	0.287 (0.058)	29%	3%	34%
B5: Openness	0.293 (0.037)	0.291 (0.044)	0.201 (0.046)	0.267 (0.045)	0.186 (0.046)	31%	8%	36%
B5: Conscientiousness	0.412 (0.041)	0.447 (0.047)	0.386 (0.050)	0.441 (0.047)	0.370 (0.051)	14%	1%	17%
B5: Extraversion	0.223 (0.037)	0.232 (0.044)	0.202 (0.045)	0.233 (0.044)	0.202 (0.045)	13%	0%	13%
B5: Agreeableness	0.349 (0.041)	0.337 (0.048)	0.281 (0.049)	0.338 (0.048)	0.262 (0.050)	17%	0%	22%
B5: Neuroticism	0.308 (0.043)	0.329 (0.051)	0.295 (0.052)	0.329 (0.051)	0.295 (0.052)	10%	0%	10%

**Note:** Sibling correlations for non-cognitive skill measures are presented. The models are estimated via REML. Standard errors of the sibling correlations are calculated via the delta method. All estimations control for fixed aged profiles (age and age squared), a gender dummy and interactions of the gender dummy and polynomials of age. All estimated correlations are significant at least at the 5 percent level. The sample size is 4,181 observations (2,778 individuals, 1,837 families) for the Big Five personality traits, 4,126 observations (2,866 individuals, 1,891 families) for locus of control, and 4,149 observations (2,877 individuals, 1,898 families) for reciprocity. A full list of parental characteristics can be found in Table 2.

**Source:** SOEPv29.

# Appendix

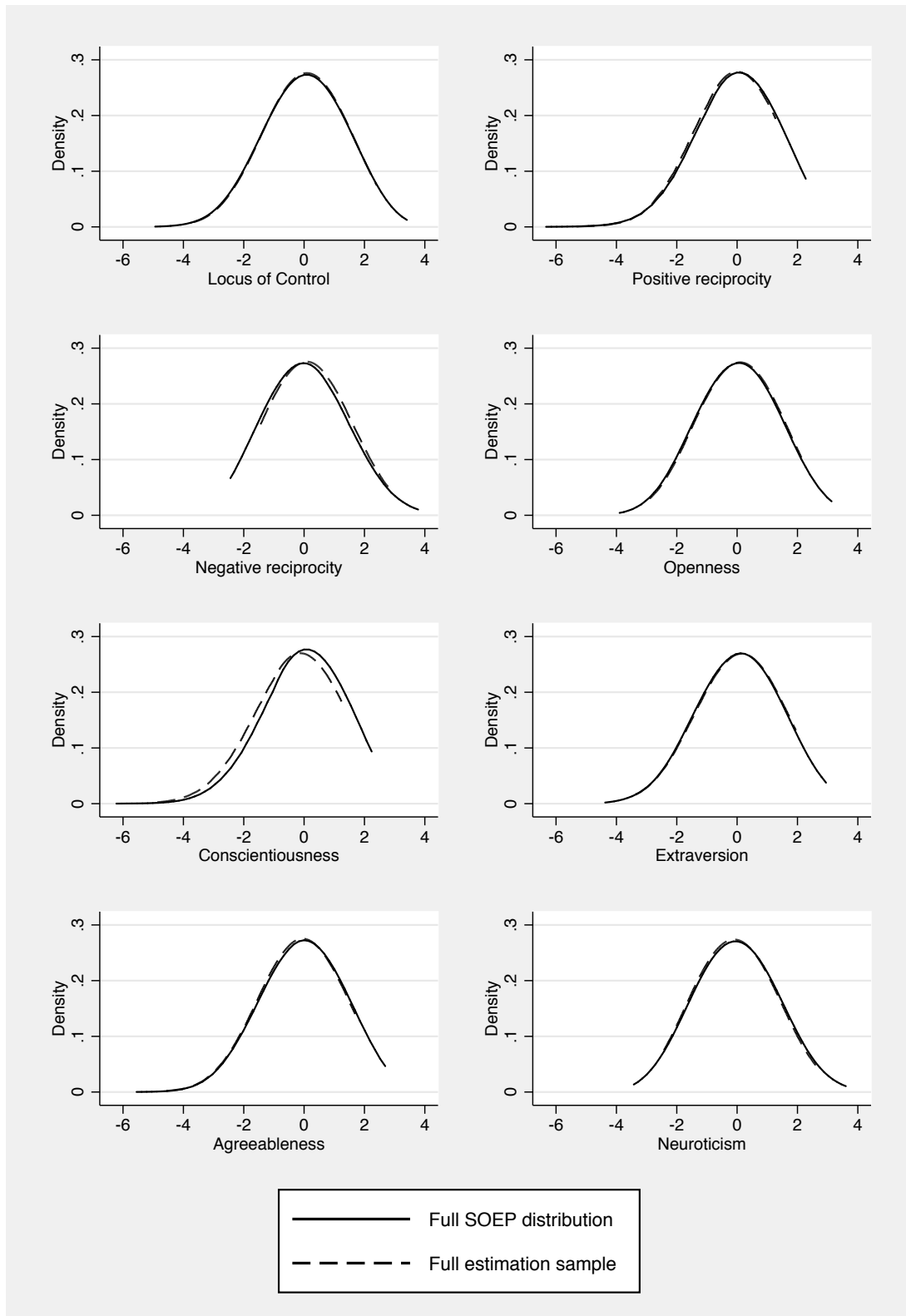
Figure A.1: Distribution of cognitive skills; full SOEP sample vs. estimation sample



Source: SOEPv29.



Figure A.2: Distribution of non-cognitive skills; full SOEP sample vs. estimation sample



Source: SOEPv29.

Table A.1: Sibling correlations in cognitive skills – basic estimates

	Crystallized intelligence	Fluid intelligence	General intelligence
<b>Sibling correlation</b> (s.e.)	<b>0.607</b> (0.062)	<b>0.545</b> (0.076)	<b>0.578</b> (0.067)
Family component (s.e.)	0.412 (0.058)	0.449 (0.078)	0.458 (0.070)
Individual component (s.e.)	0.266 (0.041)	0.374 (0.060)	0.335 (0.052)
Transitory component (s.e.)	-	-	-

**Note:** Sibling correlations for cognitive skill measures are presented with standard errors in parentheses. The models are estimated via REML. Standard errors of the sibling correlations are calculated via the delta method. All estimations control for fixed aged profiles (age and age squared) and a gender dummy as well as interactions of the gender dummy and polynomials of age. All estimated correlations are significant at least at the 5 percent level. The sample size is 443 observations/individuals from 364 families.

**Source:** SOEPv29.

Table A.2: Sibling correlations in non-cognitive skills – basic estimates

	Locus of Control	Positive reciprocity	Negative reciprocity	Big Five O	Big Five C	Big Five E	Big Five A	Big Five N
<b>Sibling correlation</b> (s.e.)	<b>0.464</b> (0.042)	<b>0.434</b> (0.063)	<b>0.383</b> (0.044)	<b>0.293</b> (0.037)	<b>0.412</b> (0.041)	<b>0.223</b> (0.037)	<b>0.349</b> (0.041)	<b>0.308</b> (0.043)
Family component (s.e.)	0.210 (0.021)	0.132 (0.019)	0.161 (0.020)	0.162 (0.021)	0.215 (0.023)	0.142 (0.024)	0.168 (0.021)	0.149 (0.022)
Individual component (s.e.)	0.243 (0.022)	0.172 (0.024)	0.260 (0.023)	0.390 (0.024)	0.307 (0.025)	0.496 (0.028)	0.313 (0.023)	0.333 (0.024)
Transitory component (s.e.)	0.466 (0.015)	0.611 (0.019)	0.481 (0.015)	0.360 (0.011)	0.499 (0.015)	0.371 (0.011)	0.444 (0.013)	0.415 (0.012)

**Note:** Sibling correlations for non-cognitive skill measures are presented. The models are estimated via REML. Standard errors of the sibling correlations are calculated via the delta method. All estimations control for fixed aged profiles (age and age squared), a survey year dummy, and a gender dummy as well as interactions of the gender dummy and polynomials of age. All estimated correlations are significant at least at the 5 percent level.

**Source:** SOEPv29.

Table A.3: Decomposition of sibling correlations in non-cognitive skills - by gender of parent

	Full sample	Sample with par. char.	controlling for			percentage reduction		
			all maternal char.	all paternal char.	all parental char.	all maternal char.	all paternal char.	all parental char.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Locus of Control	0.464 (0.042)	0.495 (0.049)	0.386 (0.055)	0.402 (0.054)	0.354 (0.057)	22%	19%	28%
Positive reciprocity	0.434 (0.063)	0.477 (0.081)	0.360 (0.092)	0.408 (0.088)	0.345 (0.095)	24%	14%	28%
Negative reciprocity	0.383 (0.044)	0.434 (0.051)	0.325 (0.056)	0.341 (0.055)	0.287 (0.058)	25%	21%	34%
B5: Openness	0.293 (0.037)	0.291 (0.044)	0.218 (0.045)	0.227 (0.045)	0.186 (0.046)	25%	22%	36%
B5: Conscientiousness	0.412 (0.041)	0.447 (0.047)	0.397 (0.050)	0.393 (0.050)	0.370 (0.051)	11%	12%	17%
B5: Extraversion	0.223 (0.037)	0.232 (0.044)	0.199 (0.045)	0.223 (0.044)	0.202 (0.045)	14%	4%	13%
B5: Agreeableness	0.349 (0.041)	0.337 (0.048)	0.292 (0.049)	0.272 (0.050)	0.262 (0.050)	13%	19%	22%
B5: Neuroticism	0.308 (0.043)	0.329 (0.051)	0.305 (0.052)	0.302 (0.052)	0.295 (0.052)	7%	8%	10%

**Note:** Sibling correlations for non-cognitive skill measures are presented. The models are estimated via REML. Standard errors of the sibling correlations are calculated via the delta method. All estimations control for fixed aged profiles (age and age squared), a gender dummy and interactions of the gender dummy and polynomials of age. All estimated correlations are significant at least at the 5 percent level. The sample size is 4,181 observations (2,778 individuals, 1,837 families) for the Big Five personality traits, 4,126 observations (2,866 individuals, 1,891 families) for locus of control, and 4,149 observations (2,877 individuals, 1,898 families) for reciprocity. A full list of parental characteristics can be found in Table 2.

**Source:** SOEPv29.

Table A.4: Sibling correlations in non-cognitive skills – coefficients including (both waves) and excluding (first wave, second wave) the transitory component

	<b>First wave 2005 (1)</b>	<b>Second wave 2009/10 (2)</b>	<b>Both waves balanced sample (3)</b>	<b>Both waves Full sample (4)</b>
Locus of Control	0.225 (0.046)	0.203 (0.042)	0.407 (0.065)	0.464 (0.042)
Positive reciprocity	0.185 (0.044)	0.222 (0.042)	0.438 (0.087)	0.434 (0.063)
Negative reciprocity	0.231 (0.043)	0.209 (0.042)	0.379 (0.065)	0.383 (0.044)
B5: Openness	0.185 (0.041)	0.148 (0.042)	0.273 (0.053)	0.293 (0.037)
B5: Conscientiousness	0.180 (0.041)	0.228 (0.042)	0.373 (0.060)	0.412 (0.041)
B5: Extraversion	0.132 (0.045)	0.109 (0.040)	0.195 (0.056)	0.223 (0.037)
B5: Agreeableness	0.168 (0.041)	0.152 (0.043)	0.305 (0.060)	0.349 (0.041)
B5: Neuroticism	0.165 (0.043)	0.200 (0.045)	0.333 (0.062)	0.308 (0.043)

**Note:** Estimates in columns (1) - (3) are based on a balanced sample of individuals with observations in both waves. Sibling correlations for single-year observations in (1) and (2) are estimated without a transitory component, and estimates in (3) include a transitory component in the model. **Source:** SOEPv29.