Determinants of personality and skill development in the Socio-emotional environment during childhood

Karsten Reuß

ZEW (Centre for European Economic Research)

2011

6. Determinants of personality and skill development in the Socio-emotional environment during childhood

Abstract:

This study investigates the importance of different socio-economic conditions on skill formation by using German data from a longitudinal study, the Mannheim Study of Children at Risk, starting at birth. A rich set of psychometric variables regarding the socio-emotional environment from birth until late childhood is assessed. The paper extends previous approaches by splitting up the information on the environment into several dimensions. The results could help policy makers to design educational interventions. Birth risk and the early mother-child interaction are the most important determinants in infancy. In middle childhood cognitive skills can be enhanced by parents who stimulate child play with appropriate play materials and by parental support in learning numbers, shapes or letters. Personality rather tends to be linked to a harmonious and motivational parent-child relationship, in particular a positive emotional climate and the stimulation of independence. Early investments are the most important, but should be complemented by investments in late childhood to unfold their benefits.

Keywords: cognitive skills, personality, multidimensional investments, socio-emotional environment, childhood, partial least squares

JEL-classification: I12, I21, J13

Acknowledgements:
I gratefully acknowledge support from the Leibniz Association, Bonn, through the grant "Noncognitive Skills: Acquisition and Economic Consequences". I thank Manfred Laucht from the Central Institute of Mental Health in Mannheim for the possibility to use the Mannheim Study of Children at Risk. I thank Dorothea Blomeyer, Manfred Laucht and Friedhelm Pfeiffer for the cooperation during the last five years and for many valuable discussions and remarks that improved my understanding of the childhood skill multipliers. The usual disclaimer applies. For helpful discussions, I thank Andrea Mühlenweg.

6.1 Introduction

In recent years several studies have examined the impact of environmental aspects on human capital formation (Borghans et al., 2008; Cunha and Heckman, 2009; Almond et al., 2009; Pfeiffer, 2010). They point out the shaping role of early childhood on skill formation and on socio-economic outcomes. An economic framework for analysing the relationships has been introduced by the technology of skill formation (Cunha and Heckman, 2007). Generally two major skill groups are distinguished: Cognitive skills including memory power, information processing speed, intellectual power, linguistic skills, motor skills as well as general problem solving abilities and noncognitive skills including motivation, persistence, activity level, social skills and emotional abilities, among others.

Even though most studies agree that the optimal timing of investments is in early childhood (Doyle et al., 2009), there is still much debate going on about how to optimally design investments. Heckman (2011) discusses different channels through which early intervention programs enhance noncognitive skills. The optimal assignments of private, governmental and non-governmental investment programs for different groups are analysed. Summing up skills and environmental aspects to a few aggregated scores might be tricky, because both are multidimensional. There may be many different ways to invest (e.g. material support vs. emotional support) and many different ways for individuals to profit from investments (e.g. in their discipline, mood or intellectual power). Understanding their interactions in more detail could help to better design investments and early intervention programs. Depending on the observed environment of each child at each age the lack of important investments could be monitored and remedied.

I use data from the Mannheim Study of Children at Risk (MARS), a longitudinal psychometric dataset following children from birth until early adulthood in Germany. The data is very rich in measurements of environmental conditions during early childhood as well as skills and socio-economic outcomes. A major part of environmental characteristics in the MARS is measured by the HOME score (Home Observation Measurement of the Environment, Bradly 1982), which assesses the parent-child interaction, living environment, play materials, activities and many other aspects by observations and interviews. The HOME consists of 25 to 101 items and 3 to 8 subscores, depending on age (for more details, see section 6.2).

The MARS data has been studied by Blomeyer et al. (2008, 2009), who look at particular measurements and find the measured abilities at preschool age as well as initial risk conditions at birth to be important for skills and performance later in life. Blomeyer et al. (2010) look at several selected environmental aspects of the HOME and the mother-child interaction and estimate the impact of early mother-child interaction on skill development until the age of 4.5 years. The mother-child interaction is as important as the early HOME for predicting the IQ in late childhood. Coneus et al. (2011) break up the skill dimension by using the complete information on eight different skill measurements to proxy three latent skills (cognitive, mental and emotional skills). In a second step the latent factors are used to estimate skill production functions. Mental and emotional skills are both noncognitive. The reason for the distinguishing two different noncognitive skills results from their great heterogeneity which can be illustrated with various statistical measures (cluster and factor analyses).

This paper combines the approach of Blomeyer et al. (2010) and Coneus et al. (2011) by splitting up the skill dimension and the environmental dimension simultaneously. The taxonomy of 3 different skill types introduced in Coneus et al. (2011) is adopted. The study extends previous approaches in three dimensions: Firstly, the HOME is split up into its aggregated subscores and into its single items. Secondly, the predictive power of additional variables of the environment (e.g. household composition, mother-child interaction, breasteeding, and childcare, among others) is examined. Thridly, a partial least squares regression (PLSR) is applied that is able to deal with the high number of correlated predictor variables.

Studies that estimate the effect of environmental aspects on skills have to consider a possible endogeneity bias. Two major sources of endogeneity are simultaneous causality and omitted variables. Not only investments could influence skills, but skills could also influence investments. A child that is smiling, motivated and curious might alter parental behaviour just like parental behaviour alters the child's personality (simultaneous causality). Several aspects might drive skill development, but are not observed such as genetic endowment, peers or the day care of grandparents (omitted variables).

I argue that simultaneous causality problems are small for several reasons. First, the correlation of the test result for one individual between interviewers amounts to 0.6 and 0.8 (Coneus et al., 2011). The respondent bias due to misunderstanding questionnaires is marginal, because the measures were commonly assessed by trained interviewers in different standardized surroundings. Also, the quality of the assessments is high, because trained interviewers observe children and their parents from birth

on leading to a high data quality. In addition, the possible bias is reduced by only considering lagged environmental characteristics in the model.

Possible omitted variables are tried to be captured by integrating the largest possible number of explanatory and dependant variables in the model. Even though some aspects are not directly observed (e.g. care of grandparents) the data may contain variables that are related (e.g. number of persons in household, parental age).

The rest of the paper is organized as follows. Section 6.2 provides information about the data and variables and shows descriptive statistics. Section 6.3 addresses the estimation strategy. Section 6.4 presents the results, section 6.5 concludes.

6.2 Data and descriptive statistics

The study uses data from the Mannheim Study of Children at Risk (MARS), a longitudinal epidemiological cohort study following infants at risk from birth to adulthood. The initial sample contains 382 first-born children (184 boys, 198 girls), born between February 1986 and February 1988. Medical and psychological examinations elevating environmental aspects, skills, personality and social outcomes were assessed in different research waves. They took place when the children were 3 months, 2, 4.5, 8, 11, 15 and 18 years old and are still going on. Participation rates between the seven waves are high, despite the extensive survey procedure, comprising a large number of medical and psychological examinations. The sample at the age of 11 years amounts to 360 observations. For a more detailed description of the dataset see Blomeyer et al. (2008, 2009).

6.2.1 Environmental variables

Many aspects of the home environment of the children between the ages of 3 months to 11 years are captured by the HOME. Bradly and Caldwell (1980) found a strong link between cognitive abilities and the HOME as a relevant measure for preparing and fostering abilities starting in early childhood. This study uses a modified version of the original HOME that is assessed by parent interviews and direct observations. The composition of HOME items changes with age as other factors become relevant. It consists of 25 items at the age of three months, 87 at the age of 2 years, 95 at the age of 4.5 years and 59 items at the ages of 8 years. The number of HOME items used at the

age of 2 and 4.5 years is extended compared to previous studies (Blomeyer et al., 2009; Coneus et al., 2011), who used 29 and 38 items at the ages of 2 and 4.5 years and Blomeyer et al. (2010), who used a selection of 40 and 47 items at the ages of 2 and 4.5 years[32]. Items are grouped into subscores, as table 6.1 shows.

Table 6.1: Description oft he HOME subscores and items

Subscore	Description	Items	Age when measured
Mother-child Interaction	Reactivity of the mother towards child, vocalisation, smooching, avoidance of punishment and aggression, integration of child during interview	11	3 months
Living Environment	Play materials, security, nursery, apartment appearance, yard quality, pet, etc…	6 (3 months) 20 (2-4.5 years)	3 months 2 years 4.5 years
Conversation process with parents and child	Clear speech, language, interest in interview, praising child, honesty, etc…	8	3 months 2 years 4.5 years
Stimulation of development and language	Allowance of child play, speech quality towards child, media use, types of playing (songs, colors, numbers, letters, etc...)	13	2 years 4.5 years
Avoidance of Restriction and Punishment	Allowing child to play during housework, avoidance of punishment, interesting activities for child during housework/interview	7	2 years 4.5 years
Promotion of Maturation and Autonomy	Praising child, promotion of autonomy with reasonable constraints, learning to tie shoes, to dress, to tidy up, to be polite, etc...	12	2 years 4.5 years
Play Materials	Toys to drive, to paint, to read, to build, to cuddle, to play music, to puzzle, to learn colors, numbers etc…	13 (2 years) 16 (4.5 years)	2 years 4.5 years
Emotional Climate	Integration of child during interview, smooching, avoidance of punishment and aggression, praising child, motivating the child, compassion	18 (2-4.5 years) 8 (8 years)	2 years 4.5 years 8 years
Emotional and Verbal Responsivity	Clear daily routine, praising child, motivating child, integration and support of child during interview, use of full sentences	10	8 years
Promotion of Social Maturity	Expecting child to tidy up, help in the household, to do homework, reasonable rules, consistency	6	8 years
Experiences and Materials promoting development	Radio and cassette recorder, music instruments, books, dictionaries, newspapers, visits to/from friends, etc…	8	8 years
Active Stimulation	Reasonable TV use, hobbies, variety of leisure activities, playground use, library card, museum visits, trips and travelling	8	8 years
Paternal engagement	Father (or equivalent person) engages in outdoor activities, sees child at least 4 days a week, participates at meals, etc…	4	8 years
Material environment	Appearance of house/apartment, order and cleanness, sufficient living space per person, acceptable noise level, secure environment for child, no smoking	7	8 years
Activities promoting development	Visits to friends or relatives, taking child to concerts, theatre, business trips, travelling, bike riding, roller skating, etc…	6	8 years

[32] Blomeyer et al. (2011) adjusted the selection in a way that it matches the original HOME as well as possible. In this study, however, all items are used for two reasons. Firstly, to address the problems of omitted variables as well as possible and secondly, because the estimator introduced in section 3 is robust to additional explanatory variables.

A major goal of this paper is to assess the predictive power of additional environmental variables. During the HOME interviews a general ***interviewer rating of the contact person*** was carried out at the ages from 3 months to 4.5 years at a 5-point scale rating. Elaborated aspects include the "perceived honesty" of the contact person during the interview varying from "continuously artificial" to "continuously sincere", the "acceptance of the child" during the interview varying from "not accepting child during the whole interview" to "child continuously accepted" and "parental reactivity "varying from "not reactive at all towards child" to "completely reactive towards child".

Additional variables on environmental aspects of the earliest stages of life are birth risks. A rating of organic risk was conducted based on the information of the maternal obstetrical and infant neonatal record. It is measured by the psychological and medical rating of several pre-, peri- or neonatal complications including premature birth, the EPH-gestosis of the mother, low birth weight, asphyxia, seizures, respiratory therapy, sepsis, etc. The variable "***low organic risk"*** denotes to the absence of organic risk factors. A rating of psychosocial risk was made based on the risk index developed by Rutter and Quinton (1977). It includes parental psychiatric disorders, broken home, delinquency, early parenthood, low quality partnership, unwanted pregnancy, disease and unemployment. "***Low psychosocial risk"*** denotes to the absence of psychosocial risk factors.

At the age of 3 months video-taped information on the ***mother-child interaction*** was rated by the MBS-MKI-S scale (Mannheimer Beurteilungsskala zur Erfassung der Mutter-Kind-Interaktion im Säuglingsalter, see Esser er al., 1989). Maternal behavior is broken down into eight dimensions: Emotion, tenderness, verbalization, verbal restrictions, congruity/authenticity, variability, reactivity/sensitivity and stimulation. Infant behavior is broken down into five dimensions: Emotion/facial expressions, verbalization, viewing direction, reactivity and the potential willingness to interact.

Another question this study will address is if the ***duration of breastfeeding*** provides additional predictive power on future skills. It was surveyed by an interview of the mothers at the age of 2 years and ranges from 0 to 104 weeks. It is split up into two variables in this paper: The amount of breastfeeding until the age of 3 month and the amount of breastfeeding until the age of 2 years.

Children do not only stay at home, especially during late childhood. For this reason, I also include the quality of ***neighborhood environment.*** It was assessed together with the HOME score at all ages between 3 months and 11 years**.** The house conditions in the neighborhood, the house type,

the infrastructure quality as well as nearby disturbances of traffic, noise, industry and bars were rated.

The data contain information on several characterstics of the parents. ***Parental education*** at the age of 3 months describes the highest graduation of the mother and father. Further variables include the ***income per capita, parental age,*** the ***number of persons in the household*** and ***single parenthood.*** Single parenthood is also part of psychosocial risk. It is separately included as psychosocial risk refers to the conditions before and during birth, but single parenthood was assessed in all waves. Information about ***external childcare*** at the age of 4.5 years was assessed retrospectively in an interview. A score that includes information on the institutional childcare of the last 6, 12 and 18 months (in daily and weekly dimensions), kindergarten use and nanny care is generated.

6.2.2 Skill variables

The most prominent skill variable is the ***IQ*** (intelligence quotient) that measures cognitive abilities. It was assessed by the Mental Developmental Index (MDI) of the Bayley Scales of Infant Development (Bayley, 1969) at the ages of 3 months and 2 years, the Columbia Mental Maturity Scale (CMMS, Burgemeister et al., 1972) at the age of 4.5 years and the Culture Fair Test at the ages of 8 and 11 years (Cattell, 1960). Each test consists of a variety of subtests such as numeracy, memory, receptive and expressive language skills. The *IQ* was measured in a verbal ***(verbal IQ)*** as well as in a nonverbal dimension ***(nonverbal IQ)*** from the age of two years onwards.

The ***MQ*** (motor quotient) was assessed by the Psychomotor Developmental Index (PDI) of the Bayley Scales at the ages of 3 months and 2 years, the Test of Motor Abilities (MOT) 4-6 (Zimmer and Volkamer, 1984) at the age of 4.5 years and the Body coordination test for children (KTK) (Kiphard and Shilling, 1974). The MQ often relates to the IQ. For more detailed information on measuring the IQ and the MQ in the MARS, see Bloymeyer et al. (2009).

Evidence suggests that noncognitive skills are at least equally important (Duckworth et al. 2005). Besides the cognitive measures the data contain several personality traits that capture noncognitive abilities. They were surveyed within a standardized parent-interview and during structured direct observations in four standardized settings on two different days in both familiar (home) and unfamiliar (laboratory) surroundings. All ratings were assessed by trained judges on 5-point rating scales of five temperamental dimensions adapted from the New York Longitudinal Study NYLS

(Thomas et al., 1968). Personality taxonomies are usually based on parent interviews and direct observations instead of questionnaires like alternative personality measures such as the Big Five. The eight personality measures employed in this study are as follows: ***Activity*** describes the frequency and intensity of motor behaviour ranging from "being inactive and slow" to "being overactive and restless". ***Approach*** describes the initial reaction to new stimuli (e.g. strangers, new food, or unfamiliar surroundings) ranging from "withdrawal" to "approach". ***Adaptability*** denotes the length of time that is needed to get habituated to the new stimuli going from "very slow/not at all adapting" to "very quickly adapting". ***Mood*** describes the general tendency of the child to be in good or bad temper ranging from "negative mood" to "positive mood". ***Persistence*** refers to a child's ability to pursue a particular activity and its continuation in the face of obstacles varying from "very low" to "very high". ***Reactivity*** measures the vehemence of the child's expression of positive and negative emotions ranging from "apathetic" to "irritable/boisterous". ***Rhythmicity*** refers to the regularity of biological functions (e.g. sleep-wake-cycle, hunger, etc.) ranging from "unpredictable" to "totally regular (like clockwork"). Finally, ***responsiveness*** accounts for the sensitivity in the child's reaction to environmental changes or external stimuli (e.g. pain, parental frowning, food temperature or new food) going from "oversensitive" to "very insensitive".[33] For more detailed information on the skill measures see Coneus et al. (2011) and Blomeyer et al. (2011).

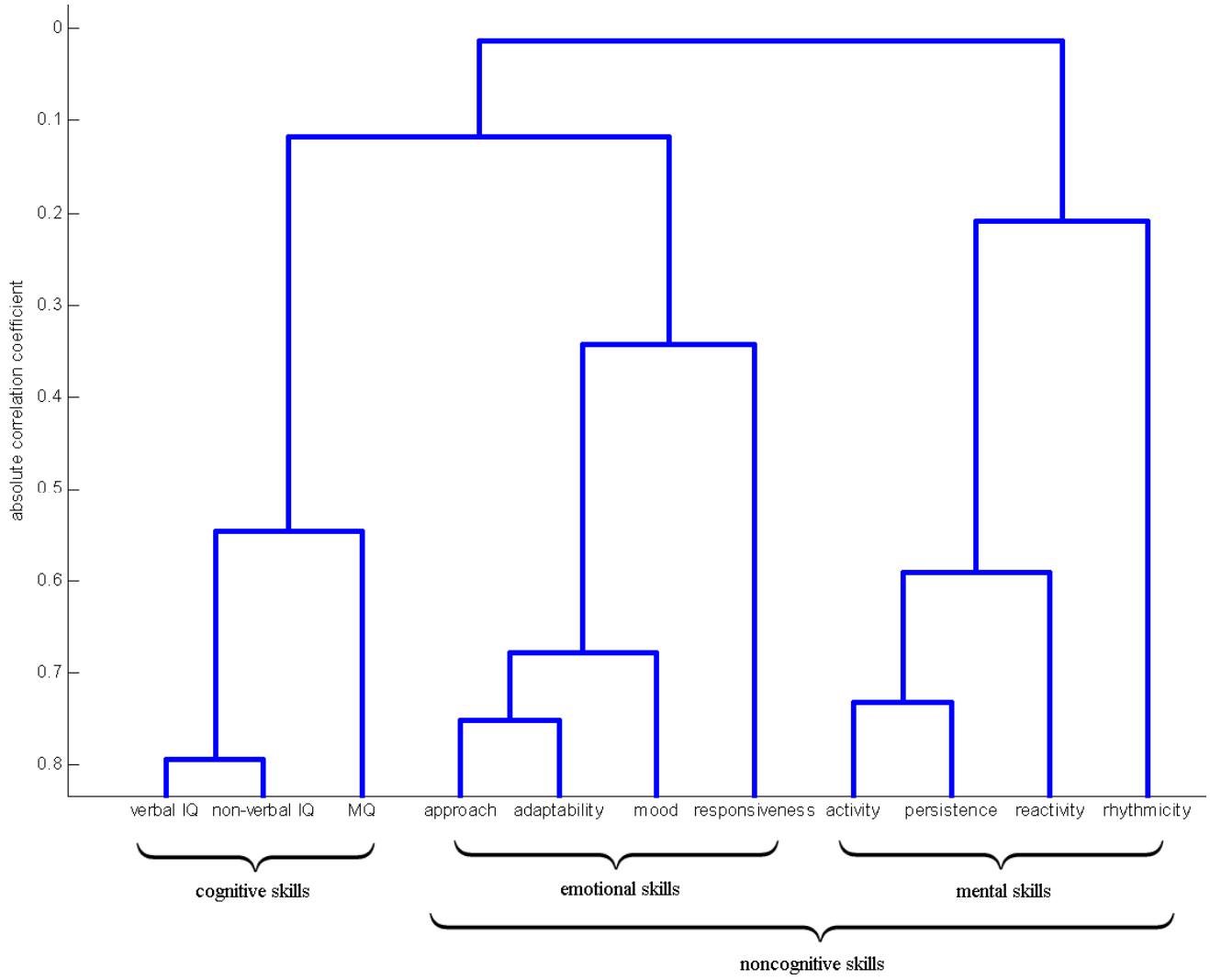
To obtain an overview of how the skills and personality measures (noncognitive skills) are related I apply hierarchical clustering (see figure 1). First, the absolute correlations are calculated and the pairs with the highest correlations are grouped. Next, the pairs that are in close proximity are linked using the information generated in the first step. As objects are paired into binary clusters, the newly formed clusters are grouped into larger clusters until a hierarchical tree is formed (Coneus et al., 2011). The link where groups in the tree presented in figure 1 connect always refers to the smallest correlation to the next cluster, e.g. the smallest absolute correlation between the measurement groups "IQ, MQ" and "approach, adaptability, mood, responsiveness" is 0.12. The Y-axis shows the absolute correlation, the X-axis the different skill measurements. Figure 6.1 shows that three major clusters exist: cognitive skills, mental skills and emotional skills. The cognitive group consists of the IQ (verbal IQ and nonverbal IQ) and the MQ. The emotional group consists of approach, adaptability and mood. All emotional measures are at least correlated by 0.67. Responsive-

[33] Usually high ratings can be associated with high noncognitive skills. In the case of activity, reactivity and sensitivity however the medium rating "3" is optimal. Hyperactivity, boisterous and very insensitive behavior can involve problems. Hence, those three temperamental measures are transformed such that the medium ratings come along with the highest score, "5", whereas the very high and very low ratings come along with the lowest score, "1".

ness is more distinct (its correlation with adaptability is 0.43). The mental group consists of activity, persistence and reactivity all being correlated by at least 0.61. Rhythmicity is only distantly related to the mental skill group. The results can be confirmed with a factor analysis (see Coneus et al., 2011).

To sum up, ***cognitive skills*** refer to memory power, information processing speed, intellectual power, linguistic skills and motor skills. ***Emotional skills*** describe the mood and the reaction and abilities to cope with new stimuli. ***Mental skills*** eventually refer to the ability to pursue certain goals and a reasonable activity level.

Figure 6.1: Correlations and clustering of skill measures



6.2.3 Variables on social outcomes

Skills generally lead to social achievements (see Pfeiffer and Reuß, 2008). Social outcomes give a glimpse on how individuals might perform later in life and on the labour market. In this study the child functional levels at the age of 11 years are studied. The functional level is a 7-point scale that condenses a multiplicity of social outcomes of the child (Marcus et al., 1993). It is broken down into five dimensions: The ***"functional level in the family"*** measures the role of the child in the family ranging on from "disintegrated, destructive behavior" to "positive engagement that improves the family environment". The ***"functional level in school"*** describes the child's achievements at school going from "huge problems at school/in attaining reading and math skills" to "very successful, high school outcomes, barely challenged". The quality of peer relations is described by the "***peer-functional level"*** from "unable to develop peer contacts" vs. "very popular, many peers, leading position" The ***"functional level of interests and leisure"*** ranges from "no interests, never inspired" to "multiple interests, high achievement in several leisure activities". Finally, the ***"functional level of autonomy"*** measures the level of independence on a scale from "dependent, not able to be out of home without assistance" to "able to travel alone, very autonomous".

6.3 Method

6.3.1 Estimation strategy

The major goal of this study is to examine the predictive power different kinds of investments have on future skills. For this purpose, the technology of skill formation is employed (Cunha and Heckman, 2007):

$$S_t^c = f_t(S_{t-1}^c, S_{t-1}^m, S_{t-1}^e, E_{t-1}) \qquad (6.1)$$

S_t^c, S_t^m and S_t^e denote cognitive, mental and emotional skills in period *t* and E_{t-1} denotes the environmental conditions that can be interpreted as investments in the child's skills for each period *t.* I try to reduce possible problems of simultaneous causality by only inserting lagged environmental aspects into equation 6.2.

In order to decrease the omitted variable bias as much as possible, all of the 11 skill variables and all of the measured environmental aspects (in some waves more than 100 items) are integrated in the model. To do so, a large set of predictor variables (environmental aspects) is regressed on the relevant response variables (skills and social outcomes). Using many correlated explanatory variables in traditional ordinary least squares regression is likely to cause multicollinearity problems (Wooldridge, 2003). It often makes the original regressor matrix almost singular and leads to identification problems. To overcome this problem several techniques have been developed such as ridge regression (Hoerl and Kennard, 1970). In ridge regression the matrix of the original regressors is modified such that it remains non-singular. For the application of this technique, however, a lot of computation is required if the number of variables is large (Aswani and Bickel, 2011). Alternatively one could implement principal component or factor analysis (Jöreskog, 1967; Cunha et al., 2010; Coneus et al., 2011). Those techniques aim at generating a small number of principal components (or latent factors) that comprise as much variation of the original variables as possible. As they can be restricted to be orthogonal in the regression, multicollinearity problems are eliminated even though a lot of information is preserved.

While factor analysis and principal component regression (PCR) are useful tools to reduce the multiplicity of response variables, their application involves some problems if one aims at reducing the multiplicity of inputs of predictor variables. Ideally the latent factors of the predictor variables should be chosen by taking into account how well they are able to predict the response variables. Choosing them independently from their responses could lead to an over-specification of the model. Additional latent predictor scores might produce unnecessary bias. Considering the principle that "it can scarcely be denied that the supreme goal of all theory is to make the irreducible basic elements as simple and as few as possible without having to surrender the adequate representation of a single datum of experience" (Einstein, 1934), the selection of latent predictor scores should lead to a model as precise as necessary but as simple and parsimonious as possible.

A regression technique suited to deal with numerous correlated predictor variables is partial least squares regression (PLSR). Aswani and Bickel (2011) perform predictions from highly correlated variables and find PLSR to perform significantly better than PCR and ridge regession. Wold (1966) did a pioneering work on PLSR in the field of econometrics. Since then PLSR has been popular among chemometricians and chemical engineers (Helland, 1980; Wold et al. 2001), but has also been used in economics (Dijkstra, 1983; Knight, 2008).

The basic framework of PLSR in this paper consists of an *n*-by-*q* matrix S of *q* skills and *n* observations and an *n*-by-*p* matrix E with *p* environmental aspects. To regress skill variables in S on the environment E, PLSR tries to find latent factors that play the same role as E (Rao et al., 2008; Boulesteix and Strimmer, 2007). The basic framework of PLS consists of two equations 6.2 and 6.3:

$$E = TP^T + \varepsilon_E \tag{6.2}$$

$$S = TQ^T + \varepsilon_S \tag{6.3}$$

T is an *n*-by-*c* matrix giving *c* latent components for *n* observations, P is a *p*-by-*c* matrix and Q a *q*-by-*c* matrix of coefficients. ε_E and ε_S contain the random errors. In a first step all variables are standardized to a mean value of 0 and a variance of 1 as uncentered basic data is assumed (Rao et al., 2008). Before starting with the identification, *c* has to be specified exogenously. Each latent factor is a linear combination of $E_1,...,E_p$:

$$T = EW \tag{6.4}$$

W is a *p*-by-*c* matrix of weights. The aim of PLSR is to capture as much information of E as possible in order to predict $S_1,...,S_q$ while reducing the dimensionality of the regression problem by using fewer components than *p*. In this paper W is identified by the SIMPLS algorithm (for a detailed description, see de Jong, 1993). In SIMPLS the following maximization is solved for each w_i in W with i=1,...,*c*:

$$w_i = \arg\max_w w^T E^T S S^T E w \tag{6.5}$$

subject to

$w_i^T w_i = 1$ and $t_i^T t_j = w_i^T E^T E w_j = 0$ for $j = 1,...,i-1$.

If we assume *q*>1 the term on the right hand side of equation 6.5 is the sum of the squared empirical covariances between the latent environmental factors, T, and the measured skills, $S_{1,...,}$ S_q.:

$$w^T E^T S S^T E w = ((Ew)^T S)^T ((Ew)^T S) = n^2 \cdot \sum_{j=1}^{q} Cov(T, S_j)^2 . \tag{6.6}$$

After identifying W, the latent components can be computed using equation 6.4. Those are then used for prediction in place of the original environmental variables. Q^T is obtained as the least squares solution of (6.2):

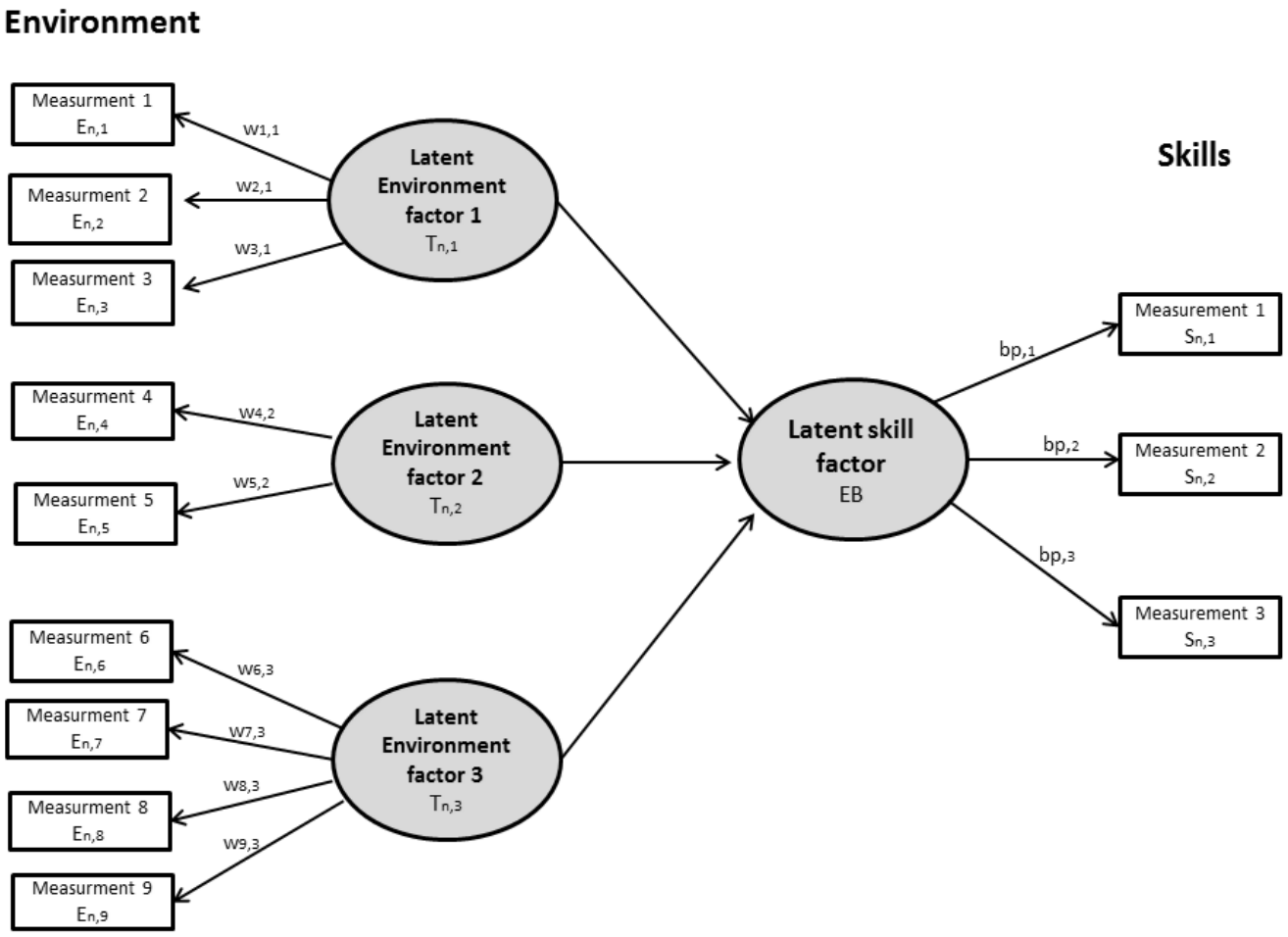
$$Q^T = (T^T T)^{-1} T^T Y \tag{6.7}$$

Obtaining P^T is analogous. Finally the *p*-by-*q* matrix B of regression coefficients for the model S=EB+ ε_S is given as:

$$B = WQ^T \tag{6.8}$$

Figure 6.2 illustrates the principles of PLSR for *c*=3, q=3 and p=9. For simplification only some of the w and the corresponding arrows are shown.

Figure 6.2: Example of partial least squares regression



PLSR offers several advantages. If $n \leq p$, traditional regression techniques such al OLS cannot be applied because the p-by-p covariance matrix E^TE is singular. In contrast, PLSR may be applied (Garthwaite, 1994). The precision and reliability of PLSR can be increased by either increasing n or p or even both. As the number of 382 observations in the first period is relatively low compared to other data while the number of variables is relatively high, this is a useful feature. By performing Monte Carlo simulations, Cassel et al. (1999) show PLSR to be robust with regard to skewness, multicollinearity, misspecification and that the latent variable scores conform to the true values. For these reasons PLSR can come in operation with quasi-metric (e.g. Likert-scales, see Vinzi et al., 2009), metric or dichotomous data. This is advantageous when analyzing MARS as different skills and environmental aspects are measured on different scales with different ratings.

The interpretation of the PLSR coefficients may be difficult as the causal relationship is only estimated for the latent factors. In case of many omitted variables the significant coefficients have to be interpreted as predictors and signals of underlying latent factors that may have causal relationships. The reliablitiy of the coefficients increases with the inclusion of additional variables into the model.

To estimate the technology of skill formation specified in equation 6.1. In line with Coneus et al. (2011) I restrict the number latent skill factors to 3. The resulting factors $\hat{S}$ correspond to cognitive, mental and emotional skills. For the age of 3 month (0.25 years) the technology of skill formation is estimated by a PLSR with

$$S^{0.25} = \begin{pmatrix} s_{1,1}^{0.25} & \cdots & s_{1,q}^{0.25} \\ \vdots & \ddots & \vdots \\ s_{n,1}^{0.25} & \cdots & s_{n,q}^{0.25} \end{pmatrix}. \tag{6.9}$$

$s_{1,1}^{0.25}$ describes the first measured skill for the first individual at the age of 0.25 years, $s_{n,1}^{0.25}$ is the first measured skill for the n^{th} individual and $s_{1,q}^{0.25}$ the q^{th} measured skill for the first individual. E for each period consists of the environmental aspects $e_{i,j}$. Only lagged aspects $e_{i,j}^{t-1}$ are considered. That means for estimating skills at the age of 3 months, $E^{0.25}$ includes only the birth risks (age 0) and a few parental characteristics.

$$E^{0.25} = \begin{pmatrix} e^0_{1,1} & \cdots & e^0_{1,p} \\ \vdots & \ddots & \vdots \\ e^0_{n,1} & \cdots & e^0_{n,p} \end{pmatrix} \tag{6.10}$$

For the ages of t=2, 4.5, 8 and 11 years E is extended by the three latent skill factors k=3 of the previous period, $\hat{S}^{t-1}_{n,k}$.

$$S^t = \begin{pmatrix} s^t_{1,1} & \cdots & s^t_{1,q} \\ \vdots & \ddots & \vdots \\ s^t_{n,1} & \cdots & s^t_{n,q} \end{pmatrix} \tag{6.11}$$

and

$$E^t = \begin{pmatrix} e^{t-1}_{1,1} & \cdots & e^{t-1}_{1,p-3} & \hat{S}^{t-1}_{1,emotional} & \hat{S}^{t-1}_{1,cognitive} & \hat{S}^{t-1}_{1,mental} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ e^{t-1}_{n,1} & \cdots & e^{t-1}_{n,p-3} & \hat{S}^{t-1}_{n,emotional} & \hat{S}^{t-1}_{n,cognitive} & \hat{S}^{t-1}_{n,mental} \end{pmatrix} \tag{6.12}$$

The model is then estimated according to (2) - (8).

Before starting the estimation, the number of latent components c needs to be specified. With each additional component c the fraction of variance explained in S is increased, but a too large number of components might lead to overfitting. For $\lim_{c \to p} B$ the coefficients of PLSR become similar to OLS involving possible problems of multicollinearity. Hence, c should be sufficiently large to capture enough variation of S, but as small as possible. A useful tool for finding the optimal number of components is the value of c that minimizes the root-mean-square error of cross-validation (RMSECV), which is a measure of a model's ability to predict new samples (Rao et. al., 2008):

$$RMSECV_c = \sqrt{\frac{\sum_{i=1}^{n} (s_{c,i} - \widetilde{s_{c,i}})^2}{n}} \tag{6.13}$$

Given a specific number of latent components c, $s_{c,i}$ are the predicted skills of the sample included in the model formation and $\widetilde{s_{c,i}}$ are predictions for samples not included in the model formation. For

the purpose of cross validation I partition the total sample into 10 subsamples. As the standard errors in PLSR cannot be derived directly from the formal structure, bootstrapping with 10,000 repetitions is used.

6.3.2 Method Illustration

Suppose we want to estimate a simplified specification of the technology of skill formation for the IQ at the age of 8 years:

$$S_8^c = f_8(S_{4.5}^c, I_{4.5}) \tag{6.14}$$

S^c in equation 6.14 is assumed to be solely measured by the IQ[34]. Assume $I_{4.5}$ to contain multiple environmental aspects at the age of 4.5 years, including the HOME subscores and others (see table 6.1 and table 6.2). Some of those variables are correlated by more than 0.6, e.g. the correlation of "avoidance of punishment and restriction" and "emotional climate" is 0.62. The results for estimations with OLS, PCR and PLSR are presented in table 6.2.

The first column (1) shows the OLS regression for the previous period IQ and the aggregated HOME. This resembles studies estimating the technology of skill formation for the MARS (Blomeyer et al., 2009; Coneus et al., 2011) with the difference that noncognitive skills are missing in equation 6.14. The self-productivity is estimated with a coefficient of 0.83. The coefficient of the HOME score at the age of 4.5 years is 0.19 and significant. An increase in the HOME at the age of 4.5 years by 0.19 standard deviations (which, for example, corresponds to an increase of the 40th to the 89th percentile in the HOME distribution) leads to an increase of the IQ at the age of 8 years by 2.85 points[35]. The second column (2) shows the OLS regression for the case when the HOME is split up into 7 different subscores and several additional environmental variables are added. Several coefficients are significant and positive, such as "play materials", the parental age, living with biological parents and the number of persons in the household, while others such as "stimulation of development and language", "single parenthood" and "external childcare" are negative. Some of the results seem counterintuitive. They might result from the high correlations among predictors. Column (3) presents the results for principal component regression.

[34] In the rest oft he paper it is assumed the cognitive skills are additionally measured by the verbal IQ, nonverbal IQ and the MQ.
[35] The IQ scale is normalized to a mean of 100 and a standard deviation of 15. Hence the coefficient values simply can be multiplied by 15 to see the gain in IQ points (0.19*15=2.85)

Table 6.2: Estimation of the IQ at the age of 8 years by OLS, PCR and PLSR

	(1) **OLS**	(2) **OLS**	(3) **PCR**	(4) **PLSR**
HOME SCORE (total)	**0.19 ***** (.065)			
Interviewer rating of contact person		0.02 (.041)	-0.06 (.079)	-0.05 (.035)
HOME: Conversation process with parents		0 (.05)	0.05 (.088)	0.03 (.037)
HOME: Stimulation of development and language		**-0.10 **** (.049)	0 (.061)	-0.05 (.036)
HOME: Living Environment		-0.04 (.051)	0.07 (.086)	0.02 (.041)
HOME: Avoidance and Restriction and Punishment		0.1 (.065)	0.05 (.047)	**0.05 *** (.03)
HOME: Emotional Climate		0.05 (.071)	0.01 (.056)	0 (.03)
HOME: Promotion of Maturation and Autonomy		0.02 (.057)	0.05 (.045)	0.03 (.036)
HOME: Play Materials		**0.10 *** (.058)	**0.24 ***** (.057)	**0.21 ***** (.041)
Neighborhood Environment		0.03 (.048)	-0.05 (.108)	0 (.055)
Income		0 (.041)	0 (.071)	0.01 (.039)
Single Parenthood		**-0.13 **** (.059)	-0.04 (.04)	**-0.08 **** (.036)
Average Parental Age		**0.08 *** (.048)	0.08 (.084)	**0.10 ***** (.037)
Number of Persons in Household		**0.11 **** (.057)	-0.09 (.048)	-0.02 (.041)
Biological Parents		**0.12 **** (.058)	**0.11 *** (.056)	**0.09 *** (.047)
External Childcare		**-0.07 *** (.037)	-0.13 (.074)	-0.1 (.054)
IQ	**0.83 ***** (.042)	**0.80 ***** (.044)	**0.48 ***** (.145)	**0.69 ***** (.053)
Adjusted R^2	0.5796	0.5946		
MSE	0.6597	0.6362		
Number of components			6	2
Percent of explained variance in S			0.3276	0.588
Estimated MSE Prediction Error			0.5442	0.472

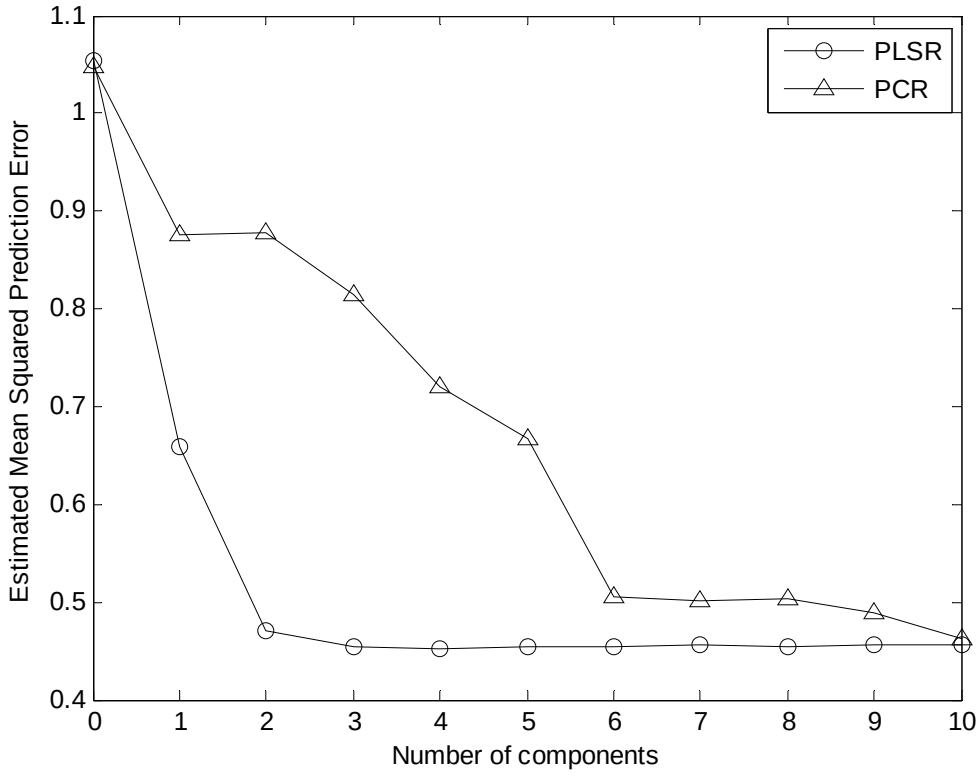
Source: Mannheim Study of Children at Risk. 360 observations. Own calculations. Standard errors are in parentheses: ***significant at 1% level, ** significant at 5 % level, * significant at 10 % level.

As figure 6.3 indicates, a low level of mean standard errors of PCR is achieved by using at least 6 components. The explained variance in S, 32.7 percent, is relatively small and smaller than the adjusted R^2 in column (2).This results from the fact that fewer predictor variables (6 components) are used. On the other hand, imposing a latent factor structure on E leads to uncorrelated predictor variables[36] and overfitting is avoided. The estimated MSE is smaller than in OLS. “Play Materials” and the fact of “living with biological” parents have a significant positive impact on the IQ of the next period according to PCR. The estimated self-productivity is relatively low (0.48). To sum up, it seems PCR can reduce multicollinearity problems, but may have problems in explaining enough variance in S.

Figure 6.3 shows that the estimated MSE for PLSR indicates that two components does about as good a job as possible. On the other hand, PCR requires 6 to 10 components to get similar prediction accuracy. In fact, the second component in PCR even increases the prediction error slightly, suggesting that the combination of predictor variables contained in that component is not strongly

correlated with S. Again, that is because PCR constructs components to explain variation in E, but not in S. The PLSR model is much more parsimonious than the PCR model. Column (4) in table 6.2 shows the estimation results of PLSR. While the estimated MSE is lower than in PCR even though only 2 components are used and the explained variance in S is higher and similar to the adjusted R^2 of OLS. In difference to OLS, PLSR avoids multicollinearity problems by choosing orthogonal latent predictors. In difference to PCR, they are chosen such that they can efficiently explain S. According to the PLSR of the simplified technology of skill formation "avoidance of punishment and restriction", "play materials", the "parental age" and "living with biological parents" have a positive effect on the future IQ. Single parenthood seems to have adverse effects. Self-productivity is estimated by 0.69, being lower than in the OLS case, but higher than in the PCR model.

Figure 6.3: Estimated mean sqaured prediction error of PCR and PLSR depending on the number of latent components



[36] PCR as well as PLSR assumes orthogonal latent factors in this study.

6.4 Results

6.4.1 Estimating the role of environmental aspects and HOME-subsores

This section estimates the technology of skill formation specified in equation 6.1 with cognitive, mental and emotional skills by PLSR for different stages of childhood. Measurements were assigned to the emotional, cognitive and mental group as specified in section 6.2.2. Table 6.3 shows the results for the PLSR of the environmental aspects at the age of 3 months on skills measures at the age of 2 years.

Table 6.3: Estimation of Mood, the IQ and the activity level at the age of 2 years based on environmental conditions until the age of 3 month

	(1) **Mood**		(2) **IQ**		(3) **Persistence**	
Interviewer rating of contact person	0	(.029)	0.01	(.029)	-0.03	(.038)
HOME Conversation process with mother and child	**0.05** *	(.03)	**0.06** **	(.029)	0.01	(.039)
HOME Mother Child Interaction	-0.01	(.028)	0.02	(.026)	0.01	(.04)
HOME Living Environment	0.01	(.025)	**0.06** **	(.024)	0.03	(.042)
Low Organic Risk	0.03	(.037)	**0.16** ***	(.036)	0.01	(.044)
Low Psychosocial Risk	**0.05** *	(.028)	**0.10** ***	(.028)	**0.12** ***	(.037)
Breastfeeding	-0.03	(.029)	0.04	(.023)	**0.07** ***	(.028)
Mother-Child Interaction (Video)	**0.09** ***	(.033)	**0.11** ***	(.031)	0.08	(.047)
Neighborhood Environment	-0.05	(.03)	-0.01	(.026)	0.03	(.036)
Income	-0.01	(.028)	**0.06** ***	(.022)	0.01	(.029)
Single Parenthood	0.04	(.031)	0.04	(.03)	0	(.036)
Parental Age	-0.04	(.03)	0.04	(.03)	0.01	(.036)
Number of Persons in Household	0.02	(.033)	-0.01	(.033)	-0.01	(.042)
Living with biological parents	**0.06** *	(.035)	0.03	(.025)	0.04	(.049)
Parental Education	-0.05	(.034)	**0.09** ***	(.026)	**0.09** ***	(.034)
Emotional Skills	**0.09** ***	(.032)	**0.07** **	(.035)	0.01	(.043)
Cognitive Skills	**0.07** **	(.037)	**0.20** ***	(.044)	**0.09** **	(.047)
Mental Skills	0.05	(.04)	0.03	(.029)	-0.02	(.045)
Number of components	2		2		2	
RMSECV	1.88		1.29		1.14	

Source: Mannheim Study of Children at Risk. 360 observations. Own calculations. Standard errors are in parentheses: ***significant at 1% level, ** significant at 5 % level, * significant at 10 % level.

At the age of 2 years the malleability of the IQ is high, with cognitive skills of the previous period accounting for about 20 percent (evidence for low self-productivity). Emotional skills of the previous period also have a positive influence on the IQ (0.07). Birth risks are the most important determinants (0.16 and 0.10) followed by the quality of mother-child Interaction (0.11), parental education (0.09), income (0.06) and two HOME subscores, "living environment" (0.06) and the "conversation process with mother and child" (0.06). A decrease in organic risk by one standard devia-

tion relates to 0.16*15=2.4 additional IQ points at the age of 2 years. A decrease in psychosocial risk by a standard deviation is linked to 1.6 additional IQ points. The mother-child interaction is equally important. Remarkably, the parental education has an additional strong positive coefficient on the IQ, even though it is closely related to the aspects covered by psychosocial risk. The mother-child interaction has a significant positive coefficient on mood (0.09), besides the HOME subscore "Conversation process with mother and child" (0.05) and a low psychosocial risk (0.05). What seems even more important for mood is "living with biological parents" (0.09). This variable may capture several aspects at the very young age. Children not living with their biological parents in infancy have usually been adopted and may have faced difficulties regarding their biological parents. Not all of the issues involved with not living with biological are captured by the HOME or the mother-child interaction. The development of persistence is also driven by a low psychosocial risk (0.12), parental education (0.09) and breastfeeding (0.07), supporting studies that find beneficial effects from breastfeeding (Goodhall et al., 2007).

The results show that the HOME at the age of 3 months captures only a fraction of the relevant determinants. The the mother-child interaction is equally important, a result in line with Blomeyer et al. (2010). The birth risks tend to have even more predictive power on skills at the age of 2 years. Organic risk seems to be mainly related to lower cognitive skills. Psychosocial risk affects all skills, but mental skills (persistence) the most. Material aspects such as income and the living environment are more closely linked to the IQ.

While interpreting the results, it has to be kept in mind that multiple indirect effects of the environment may additionally exist as "skill begets skill" (Cunha et al., 2006). Even cognitive skills at the age of 3 months have already a strong significant coefficient on mood (0.11) at the age of 2 years. So anything that improves the IQ could eventually also improve mood.

Table 6.4 shows results of regressing of the environmental conditions at the age of 2 years on the skills at the age of 4.5 years. A big difference to the results shown in table 6.3 is the strong increase in self-productivity for all skills.[37] For the IQ cognitive skills of the previous period have a coefficient of 0.49, the impact of emotional skills of the previous period on mood amounts to 0.29 and

[37] Note that the coefficients are not exactly equal to the self-productivity as they refer to the latent skills. The latent skills are generally highly correlated with the measurements (>0.8), so a high coefficient gives evidence of a high self-productivity. Alternatively the sole measurements could be used in the regression, but integrating all of them would make the interpretation of the results more diffuse, integrating only few would drop a lot of relevant information on previous period skills.

the impact of previous period mental skills on persistence to 0.28. These results suggest that skills start to become more stable at the age of 4.5 years. Direct-complementarities start to emerge among skills: Cognitive skills have a large coefficient on noncognitive skills. The relationship also goes into the other direction, but to a much smaller extent.

The HOME still captures roughly only half of the determinants that have significant coefficients on the skills even though it includes 87 items at the age of 2 years (compared to 25 at the age of 3 months, see section 6.2.2). By far the most important and sole significant HOME subscore is "play materials" with a coefficient of 0.16 on the IQ, 0.09 on persistence and 0.07 on mood. The coefficient of 0.16 means that an increase in the subscore "play materials" by one standard deviation (which corresponds to a move from the 50^{th} to the 87^{th} percentile) increases the IQ at the age of 4.5 years by 2.4 points. Income and "living with biological parents" are other determinants with a positive, significant coefficient. Especially the latter variable gives evidence that not all aspects of the child-parent relationship are covered by the HOME. The coefficient is even higher for persistence (0.14), which is also influenced by breastfeeding (0.06). Single parenthood negatively affects emotional skills.

All in all, several other variables seem to account for the skill development and outperform the HOME with respect to predictive power on future skills. Table 6.5 shows the results of regression of the environmental conditions at the age of 4.5 years on the skill measurements at the age of 8 years.

The evidence suggests self-productivity to further increase as the coefficient of cognitive skills on the IQ (0.67) rises. It remains relatively low for mental (0.3) and emotional skills (0.27). Cognitive skills foster the noncognitive skills of the subsequent periods to a great extent (the coefficient of cognitive skills on persistence amounts to 0.52, on mood it is 0.11). Like in table 6.4, the HOME subscore "play materials" is the most relevant subscore again (a coefficient of 0.21 on the IQ and 0.16 on persistence). That means an increase of one standard deviation in "play materials" (corresponding to an increase of the 30^{th} to the 91^{th} percentile) is related to an increase of the future IQ by 3.15 points. Additionally, persistence is enhanced, which may in turn positively affect the IQ again and vice versa.

Table 6.4: Estimation of Mood, the IQ and the activity level at the age of 4.5 years from environmental conditions at the age of 2 years

	(1) **Mood**		(2) **IQ**		(3) **Persistence**	
Interviewer rating of contact person	**0.06 ***	(.034)	-0.04	(.028)	-0.03	(.032)
HOME Conversation process with parents	0	(.03)	-0.01	(.026)	-0.04	(.035)
HOME Stimulation of development and language	-0.02	(.034)	0.04	(.026)	0.02	(.029)
HOME Living Environment	0.02	(.035)	0.05	(.034)	0	(.044)
HOME Avoidance and Restriction and Punishment	-0.04	(.03)	-0.04	(.025)	-0.02	(.03)
HOME Emotional Climate	0.03	(.025)	-0.01	(.02)	0.02	(.022)
HOME Promotion of Maturation and Autonomy	0.04	(.031)	0	(.023)	0.02	(.03)
HOME Play Materials	**0.07 ****	(.033)	**0.16 *****	(.035)	**0.09 *****	(.035)
Neighborhood Environment	0	(.034)	0.02	(.028)	-0.05	(.034)
Income	-0.01	(.034)	**0.07 *****	(.02)	0.02	(.039)
Single Parenthood	**-0.07 ****	(.036)	0	(.034)	0	(.038)
Parental Age	-0.04	(.035)	0.05	(.033)	0.03	(.038)
Number of Persons in Household	-0.05	(.04)	0	(.033)	0.02	(.043)
Living with biological parents	0.06	(.037)	**0.06 ***	(.037)	**0.14 *****	(.039)
Breastfeeding	-0.06	(.043)	0.02	(.028)	**0.06 ***	(.033)
Emotional Skills	**0.29 *****	(.051)	**0.10 *****	(.034)	0.02	(.043)
Cognitive Skills	**0.16 *****	(.041)	**0.49 *****	(.043)	**0.33 *****	(.05)
Mental Skills	0.05	(.044)	**0.11 ****	(.051)	**0.28 *****	(.048)
Number of components	2		2		2	
RMSECV	1.75		1.07		0.86	

Source: Mannheim Study of Children at Risk. 360 observations. Own calculations. Standard errors are in parentheses: ***significant at 1% level, ** significant at 5 % level, * significant at 10 % level.

Table 6.5: Estimation of Mood, the IQ and the activity level at the age of 8 years from environmental conditions at the age of 4.5 years

	(1) **Mood**		(2) **IQ**		(3) **Persistence**	
Interviewer rating of contact person	0.01	(.026)	-0.04	(.034)	-0.07	(.041)
HOME Conversation process with parents	**0.04 ***	(.023)	0.04	(.036)	0.01	(.045)
HOME Stimulation of development and language	-0.02	(.025)	-0.05	(.033)	-0.07	(.039)
HOME Living Environment	0.02	(.026)	0.03	(.041)	-0.06	(.042)
HOME Avoidance and Restriction and Punishment	0.03	(.02)	0.04	(.03)	-0.01	(.042)
HOME Emotional Climate	0.01	(.021)	-0.01	(.03)	0	(.032)
HOME Promotion of Maturation and Autonomy	0.04	(.027)	0.04	(.034)	**0.11 *****	(.047)
HOME Play Materials	0.02	(.023)	**0.21 *****	(.041)	**0.16 *****	(.037)
Neighborhood Environment	-0.01	(.029)	-0.02	(.049)	-0.01	(.05)
Income	0.03	(.024)	0	(.039)	0.02	(.037)
Single Parenthood	-0.04	(.026)	**-0.07 ****	(.036)	-0.01	(.045)
Average Parental Age	**-0.06 ****	(.027)	**0.07 ***	(.039)	0.03	(.04)
Number of Persons in Household	0	(.025)	-0.03	(.039)	0	(.045)
Living with biological Parents	0	(.024)	**0.09 ****	(.046)	0.04	(.044)
External Childcare	0.03	(.027)	-0.1	(.052)	-0.03	(.049)
Emotional Skills	**0.27 *****	(.04)	**0.08 ***	(.047)	0.01	(.051)
Cognitive Skills	**0.11 *****	(.031)	**0.67 *****	(.06)	**0.52 *****	(.064)
Mental Skills	0.04	(.033)	0.06	(.058)	**0.30 *****	(.067)
Number of components	2		2		2	
RMSECV	1.7		1.05		1.02	

Source: Mannheim Study of Children at Risk. 360 observations. Own calculations. Standard errors are in parentheses: ***significant at 1% level, ** significant at 5 % level, * significant at 10 % level.

Interestingly, all other HOME subscores do not seem very relevant for the IQ. “Living with biological parents” again provides additional predictive power (0.09) as well as single parenthood (-0.07) and parental age (0.07). Astonishingly, the coefficient of “parental age” on emotional skills points just to the opposite direction (-0.06) and is significant. This suggests that children of older parents may turn out smarter, but children of younger parents may turn out happier. It is an example that investments not necessarily go into one direction only and reveals a possible disadvantage of large aggregates. Mood at the age of 8 years is related to the “conversation process with the parents” (0.04). Another relevant HOME subscore is the “promotion of maturation and autonomy”. While having a positive coefficient on all skills it is significant only for persistence (0.11). Finally, table 6.6 shows the results of regression of the environmental conditions at the age of 8 years on the skills at the age of 11 years.

Again, self-productivity of noncognitive skills only slightly increases (0.37 for emotional and 0.33 for mental skills). In contrast, cognitive skills are relatively stable at the age of 11 years (0.84). Direct complementarities exist particularly among cognitive and mental skills. Those results are in line with previous studies (Blomeyer et al., 2009; Coneus et al., 2011). A difference is that several HOME aspects at the age of 8 years are still linked to the skills at the age of 11 years, suggesting some degree of plasticity. While the fundament of cognitive skill is set in early childhood continuous practise may help in maintaining a higher IQ. The age of 8 years is the only period, in which the HOME alone accounts for all of the relevant determinants. The most important subscores are “materials and experiences promoting development” (0.16 on cognitive skills, 0.06 on mental skills) and “active stimulation” (0.04 on mood, 0.08 on the IQ and 0.08 on persistence). Upon that “emotional climate” plays a role for mental skills (0.07) and “activities promoting development” for emotional skills (0.05).

All in all, the results suggest that cognitive skills stabilize faster than noncognitive skills. This is in line with studies that suggest noncognitive skills to be more malleable (Cunha and Heckman, 2007). On the other hand, if the HOME is not aggregated, certain subscores still play a role at the age of 8 years. This does not contradict the fact that early childhood is very important as the results related to the birth risks and mother-child interaction show. But the results suggest that investments in infancy should be complemented by investments in later childhood to yield success.

To test the robustness of the results and to look at the long-term consequences of investments during infancy, the skills of the previous periods are substituted by earlier environmental conditions

and regressed on the skills at the age of 11 years. For this, S_4 are substituted by E_3 and S_3 in equation 1 for t=5. The S_3 are then substituted by E_2 and S_2. S_2 are finally substited by S_1 and E_1 and S_1 by the initial birth risk, E_0. Equations 6.2 to 6.8 are applied. Table 6.7. shows the results.

Table 6.6: Estimation of Mood, the IQ and the activity level at the age of 11 years from environmental conditions at the age of 8 years

	(1) **Mood**		(2) **IQ**		(3) **Persistence**	
HOME Paternal Engagement	-0.03	(.028)	0	(.045)	-0.06	(.035)
HOME Emotional and verbal responsitivty	0.03	(.027)	0	(.038)	0.02	(.031)
HOME Active Stimulation	**0.04 ***	(.023)	**0.08 ****	(.04)	**0.08 *****	(.03)
HOME Material Environment	0.03	(.026)	0	(.038)	-0.04	(.031)
HOME Activities promoting development	**0.05 ****	(.023)	-0.01	(.04)	0	(.032)
HOME Emotional Climate	0.02	(.028)	-0.07	(.039)	**0.07 ****	(.031)
HOME Promotion of Social Maturity	**0.05 ***	(.026)	0.01	(.044)	0.01	(.036)
HOME Materials and Experiences promoting development	0.02	(.023)	**0.16 *****	(.037)	**0.06 ****	(.03)
Neighborhood Environment	0.03	(.034)	0.02	(.051)	0.06	(.048)
Income	0.04	(.031)	0.02	(.034)	0.02	(.03)
Single Parenthood	0	(.029)	-0.02	(.049)	-0.07	(.036)
Average Parental Age	-0.03	(.03)	0.01	(.051)	0.02	(.037)
Number of Persons in Household	0	(.034)	-0.06	(.052)	-0.02	(.052)
Living with biological Parents	0.03	(.031)	0.06	(.048)	**0.10 *****	(.04)
Emotional Skills	**0.37 *****	(.051)	**0.18 *****	(.045)	**0.09 ****	(.045)
Cognitive Skills	0.03	(.054)	**0.84 *****	(.056)	**0.40 *****	(.048)
Mental Skills	-0.06	(.05)	**0.18 *****	(.057)	**0.33 *****	(.037)
Number of components	3		3		3	
RMSECV	1.65		1.13		0.8	

Source: Mannheim Study of Children at Risk. 360 observations. Own calculations. Standard errors are in parentheses: ***significant at 1% level, ** significant at 5 % level, * significant at 10 % level.

There exist several differences to the results previously presented in this section. Psychosocial risk has no significant coefficient anymore, but organic risk is even more harmful. It seems that organic risk has stronger adverse long-term effects relative to psychosocial risk, which is more detrimental for skills at the age of 2 years. Psychosocial risk is rather related to other adverse conditions later in childhood and its coefficient is absorbed by those aspects. Parental education is a better predictor for persistence and the IQ. The mother-child interaction has a long-term predictive power on mood and the IQ at the age of 11 years. Breastfeeding has a significant positive coefficient on the IQ and persistence in late childhood. The subscore “play materials” has significant coefficients at the age of 2 years and 4.5 years on both, the IQ and persistence. This confirms the results presented earlier in this section. At the age of 4.5 years single parenthood is inversely related with skills. Even at the age of 8 years several subscores have predictive power, in particular “materials and experiences promoting development” and “active stimulation”.

Table 6.7: Estimation of skills at the age of 11 years from all environmental conditions

		(1) Mood		(2) IQ		(3) Persistence	
3 months	Interviewer rating of contact person	-0.01	(.023)	-0.06	(.035)	0	(.032)
	HOME Conversation process with mother and child	**0.04 ***	(.023)	0.02	(.043)	-0.05	(.04)
	HOME Mother Child Interaction	-0.02	(.023)	0	(.035)	0	(.036)
	HOME Living Environment	0.01	(.021)	0	(.034)	0	(.039)
	Low Organic Risk	0.03	(.03)	**0.26 *** **	(.056)	**0.14 *** **	(.041)
	Low Psychosocial Risk	-0.02	(.022)	0.03	(.033)	0.05	(.034)
	Mother-Child Interaction (Video)	**0.06 ** **	(.025)	**0.13 *** **	(.047)	0.05	(.046)
	Neighborhood Environment	0.01	(.018)	0	(.033)	-0.06	(.033)
	Income	-0.01	(.014)	0	(.028)	0.03	(.03)
	Single Parenthood	0.02	(.019)	-0.04	(.035)	0.01	(.033)
	Parental Age	-0.03	(.016)	0.01	(.019)	0	(.016)
	Number of Persons in Household	-0.03	(.026)	-0.03	(.042)	0.03	(.043)
	Living with biological parents	**0.03 ***	(.02)	-0.06	(.034)	-0.01	(.029)
	Parental Education	-0.01	(.02)	**0.10 *** **	(.036)	**0.08 *** **	(.033)
2 years	Interviewer rating of contact person	0	(.018)	-0.02	(.029)	-0.04	(.024)
	HOME Conversation process with parents	**0.05 ** **	(.021)	-0.03	(.032)	-0.04	(.035)
	HOME Stimulation of development and language	0	(.016)	0	(.026)	-0.03	(.025)
	HOME Living Environment	0.02	(.021)	0.03	(.044)	0.05	(.041)
	HOME Avoidance and Restriction and Punishment	0	(.017)	-0.05	(.027)	-0.02	(.03)
	HOME Emotional Climate	0.01	(.017)	-0.05	(.027)	0.02	(.029)
	HOME Promotion of Maturation and Autonomy	0	(.018)	-0.04	(.032)	-0.04	(.035)
	HOME Play Materials	0.02	(.022)	**0.14 ** **	(.065)	**0.09 ** **	(.045)
	Neighborhood Environment	-0.01	(.021)	-0.01	(.029)	0	(.034)
	Income	-0.01	(.014)	0.03	(.023)	0.02	(.025)
	Single Parenthood	**-0.06 *** **	(.022)	-0.04	(.038)	-0.02	(.038)
	Parental Age	-0.03	(.016)	0.01	(.019)	0	(.016)
	Number of Persons in Household	-0.05 **	(.022)	-0.01	(.041)	-0.07	(.036)
	Living with biological parents	0.03	(.019)	-0.01	(.038)	0.04	(.028)
	Breastfeeding	-0.04	(.024)	**0.06 ** **	(.028)	**0.06 ** **	(.029)
4.5 years	Interviewer rating of contact person	0	(.02)	0.02	(.034)	-0.04	(.028)
	HOME Conversation process with parents	**0.08 *** **	(.023)	0.01	(.038)	0.03	(.038)
	HOME Stimulation of development and language	-0.01	(.015)	0	(.026)	-0.03	(.024)
	HOME Living Environment	0.02	(.021)	0.03	(.037)	-0.05	(.038)
	HOME Avoidance and Restriction and Punishment	0.03	(.021)	0.01	(.03)	0.01	(.032)
	HOME Emotional Climate	**0.03 ***	(.019)	-0.02	(.031)	0.02	(.033)
	HOME Promotion of Maturation and Autonomy	**0.05 *** **	(.02)	0.04	(.033)	0.05	(.035)
	HOME Play Materials	0.01	(.02)	**0.18 *** **	(.054)	**0.18 *** **	(.047)
	Neighborhood Environment	-0.02	(.023)	-0.01	(.046)	0.01	(.047)
	Income	0	(.019)	-0.02	(.032)	-0.05	(.029)
	Single Parenthood	-0.02	(.019)	**-0.09 *** **	(.039)	**-0.07 ** **	(.035)
	Average Parental Age	-0.03	(.016)	0.01	(.019)	0	(.016)
	Number of Persons in Household	-0.04	(.019)	-0.06	(.032)	-0.02	(.027)
	Living with biological Parents	0.02	(.02)	0.01	(.037)	0.04	(.029)
	External Childcare	**0.06 *** **	(.025)	-0.06	(.053)	-0.06	(.043)
8 years	HOME Paternal Engagement	-0.03	(.021)	0.05	(.05)	-0.01	(.036)
	HOME Emotional and verbal responsitivty	0.01	(.022)	0.03	(.028)	**0.07 *** **	(.031)
	HOME Active Stimulation	0.02	(.022)	**0.13 *** **	(.041)	**0.10 *** **	(.033)
	HOME Material Environment	0.01	(.017)	-0.05	(.035)	-0.02	(.034)
	HOME Activities promoting development	0.02	(.017)	0.03	(.034)	0.01	(.037)
	HOME Emotional Climate	0	(.023)	-0.03	(.039)	**0.12 *** **	(.039)
	HOME Promotion of Social Maturity	0.03	(.019)	0.02	(.041)	0.05	(.037)
	HOME Materials and Experiences promoting development	0	(.019)	**0.18 *** **	(.046)	**0.10 *** **	(.035)
	Neighborhood Environment	0.01	(.024)	0.04	(.045)	0.07	(.05)
	Income	0.02	(.021)	0.02	(.028)	0.01	(.028)
	Single Parenthood	0	(.023)	0.04	(.057)	-0.01	(.04)
	Average Parental Age	-0.03	(.016)	0.01	(.019)	0	(.016)
	Number of Persons in Household	0	(.023)	-0.04	(.042)	0.02	(.042)
	Living with biological Parents	0.01	(.02)	**0.10 ** **	(.049)	**0.09 *** **	(.034)
Number of components		3		3		3	
RMSECV		2.08		1.83		1.13	

Source: Mannheim Study of Children at Risk. 360 observations. Own calculations. Standard errors are in parentheses: ***significant at 1% level, ** significant at 5 % level, * significant at 10 % level.

A striking result is the close similarity to the predictive power of the variables on the IQ and persistence until the age of 4.5 years. The results remind the coefficients in tables 6.3 to 6.6 that predict the IQ. The reason for this could result from the high coefficients of IQ on persistence (direct complementarity). As shown in the earlier regressions, the IQ has a great predictive power on persistence. Direct complementarities could also explain why environmental aspects at the age of 4.5 years are relevant for mood at the age of 11 years, even though they have not been relevant for mood at the age of 8 years (see table 6.5).

To sum up, this section suggests that other factors besides the HOME play a role in predicting skills of future periods - especially until the age of 4.5 years. The analysis confirms results of previous studies that investigate the predictive power of HOME subscores on later academic performance and find play materials to be the most relevant subscore (Bradley and Caldwell, 1984; Wolfgang and Stakenas, 1985).

The predictive power of the HOME score could be stronger if it is not aggregated to subscores. Section 6.4.2 splits off the HOME into its single items and preserves all the contained information. This also sheds some more light on the meaning of the different HOME subscores that have been discussed in this section.

6.4.2 Estimating the role of environmental aspects and HOME items

Instead of using the aggregated HOME subscores as predictor variables of the skills, this section integrates all single items in the model. As a result the matrix E contains up to 125 environmental variables *(*at age of 4.5 years). As described in section 6.3, the PLSR tends to profit from additional variables even if they are highly collinear. On the other hand, presenting the coefficient of each single item would boost the size of the tables. Thus, I always present the five highest absolute coefficients of the environmental aspects on the respective skills. Skills of the previous periods are integrated in the model in an identical way as in section 6.4.1, but are not discussed here. Table 6.8 shows the results of the regression of the environmental aspects at the age of 3 months on skill measures at the age of 2 years. Even though the available data increases, the RMSECV can be reduced from 1.29 to 1.16 for the prediction of the IQ and from 1.14 to 1.13 for the prediction of persistence. For mood it is slightly higher. The results presented in table 6.3 are confirmed.

The IQ at the age of 2 years is mainly linked to the birth risks, with both risks being approximately equally important, followed by the mother-child interaction and income. The integration of all the HOME items slightly reduces the contribution of organic risk (from 0.16 to 0.10), but not of psychosocial risk (0.10 to 0.11) and breastfeeding (0.06). No HOME items are found among the strongest predictors of the IQ, but several have positive coefficients on noncognitive skills. Adequate living conditions and praising the child are linked to patterns that drive persistence. Maternal behaviour during the conversation process, unforcefulness, eloquence during the interview and maternal interest have positive predictive power on mood. Organic risk and income are more closely related to the IQ. The mother-child interaction and psychosocial risk have a strong predictive power on all skills. Table 6.9 shows the results for the regression of the environmental aspects at the age of 2 years on skill measures at the age of 4.5 years.

Table 6.8: Estimation of Mood, the IQ and the activity level at the age of 2 years by environmental conditions until the age of 3 month, five largest environmental coefficients

Mood	Mother Child Interaction (Video)		**0.08*** **	0.03
	Living with biological parents		0.05	0.032
	Low Psychosocial Risk		**0.05***	0.026
	Mother talks unforcedly, is eloquent	Conversation process	**0.05****	0.021
	Mother is active during the interview (asks questions, etc.)	Conversation process	**0.04****	0.022
	Number of components		2	
	RMSECV		2.03	
IQ	Low Psychoscoial Risk		**0.11*** **	0.035
	Low Organic Risk		**0.10*** **	0.039
	Mother-Child Interaction (Video)		**0.10*** **	0.032
	Income		**0.09*** **	0.026
	Breastfeeding		**0.06***	0.037
	Number of components		3	
	RMSECV		1.16	
Persistence	Low Psychosocial Risk		**0.12*** **	0.031
	Parental Education		**0.10*** **	0.03
	Adequate living conditions (not too dark, loud, narrow)	Living Environment	**0.09***	0.045
	Mother Child Interaction (Video)		**0.08***	0.042
	During visit mother praises child at least once	Conversation process	**0.08***	0.044
	Number of components		2	
	RMSECV		1.13	

Source: Mannheim Study of Children at Risk. 360 observations. Own calculations. Standard errors are in parentheses: ***significant at 1% level, ** significant at 5 % level, * significant at 10 % level. HOME items in *italics*

Again, the integration of all the distinct HOME items can reduce the RMSECV from 1.05 to 0.83 for cognitive skills even though more information is included in the model. For noncognitive skills it slightly increases.

In difference to table 6.8 the results in table 6.9 are dominated by HOME items. One reason may result from the fact that the HOME at the age of 2 years contains 87 items (instead of 25 at the age of 3 months). Another reason is the additional information gained by disaggregating the HOME (compare with table 6.4). The items of the subscore "play materials" are the most common among the relevant predictors – especially for the IQ. Owning a table and a chair suited for children (e.g. a place to sit and practise) as well as equipment to paint is linked to patterns that improve cognitive skills. Toys to drive are likely to be beneficial for the cognitive abilities, too. The other two relevant items rather stem from the patterns of conversation between the parents and the child, answering the child's questions and helping it to improve. This shows an important issue: The single items should not be understood as the sole relevant aspects educators and parents should adress. A variety of play materials in combination with parents that take care of the child's verbal abitlites indicates that it is important for the IQ to have parents that actively help the child in training various tasks. A lot of toys highlight that parents take care to improve their child's abilities.

Table 6.9: Estimation of Mood, the IQ and the activity level at the age of 4.5 years by environmental conditions until the age of 2 years, five largest environmental coefficients

Mood	*Child has toys that requiring hand movement (e.g. coloring books)*	Play Material	**0.10*** **	0.033
	Child has toy that requires free movement	Play Material	**0.07** **	0.034
	Parents praise child at least twice during interview	Emotional Climate	**0.07*** **	0.03
	No disturbances by traffic, commerce in neighborhood	Neighborhood environment	**0.07** **	0.028
	Parents motivate child to do something idependently	Promotion of Maturation and Autonomy	**0.06***	0.034
	Number of components		3	
	RMSECV		1.98	
IQ	*Child has a child table and chair or equivalent*	Play Material	**0.09*** **	0.029
	Family has painting equipent or material requiring hand movement	Play Material	**0.07*** **	0.024
	Parents verbally react to questions and wishes of the child	Emotional Climate	**0.06*** **	0.027
	Child has a toy to drive	Play Material	**0.06** **	0.027
	Parents expand verbalizations of the child (to full sentences)	Stimulation of development and language	**0.05*** **	0.021
	Number of components		3	
	RMSECV		0.83	
Persistence	Living with biological parents		**0.10*** **	0.027
	Parents set reasonable limits for child	Promotion of Maturation and Autonomy	**0.08*** **	0.03
	Child has toy that requires free movement	Play Material	**0.07*** **	0.029
	No excessive interference in childs' actions during interview	Stimulation of development and language	**0.07** **	0.031
	If child calls for help, parents motivate child to help itself	Emotional Climate	**0.06*** **	0.025
	Number of components		3	
	RMSECV		1.14	

Source: Mannheim Study of Children at Risk. 360 observations. Own calculations. Standard errors are in parentheses: ***significant at 1% level, ** significant at 5 % level, * significant at 10 % level. HOME items in *italics*

However, play materials also predict mood. Aspects such as praising and motivating the child seem relevant, too. Disturbances in the neighbourhood such as traffic are inversely related. With respect to persistence, play materials have one significant item (toys that require free hand movement). The parent-child interaction tends to be more important. Parents, who set reasonable limits, motivate

the child and do not excessively interrupt the child's actions, are related to a higher persistence. This indicates that parents who have a very harmonious and stimulating relationship with their children can enhance its noncognitive skills.

The coefficient of "living with biological parents" is smaller than in table 6.4 (0.1 instead of 0.14), but still might capture additional important patterns. Table 6.10 shows the results for the regression of the environmental aspects at the age of 4.5 years on skill measures at the age of 8 years.

Table 6.10: Estimation of Mood, the IQ and the activity level at the age of 8 years by environmental conditions until the age of 4.5 years, five largest environmental coefficients

Mood	*Child can express negative emotions without having to expect sanctions*	Avoidance and Restriction and Punishme	**0.06*** **	0.027
	Family has a pet	Living Environment	**0.06** **	0.027
	Acceptance of child	Contact Person	**0.05** **	0.025
	Parents set reasonable limits for child	Promotion of Maturation and Autonomy	**0.05***	0.026
	Parents do not continously patronize child	Promotion of Maturation and Autonomy	**0.05** **	0.024
	Number of components		3	
	RMSECV		1.89	
IQ	*Parents teach the child numbers and letters*	Stimulation of development and language	**0.07*** **	0.019
	Child has toys that prepare learning numbers	Play Material	**0.07*** **	0.018
	Parents teach the child colors and shapes	Stimulation of development and language	**0.05** **	0.024
	Child has toy to drive	Play Material	**0.05***	0.028
	Child has at least three puzzles	Play Material	**0.05***	0.024
	Number of components		3	
	RMSECV		0.97	
Persistence	*Child has toys that prepare learning numbers*	Play Material	**0.07*** **	0.018
	Parents teach the child colors and shapes	Stimulation of development and language	**0.06*** **	0.019
	Child has toys that prepare learning letters	Play Material	**0.05*** **	0.018
	Acceptance of child	Contact Person	**0.04*** **	0.018
	Parents teach child to be polite	Promotion of Maturation and Autonomy	**0.04***	0.022
	Number of components		3	
	RMSECV		1.01	

Source: Mannheim Study of Children at Risk. 360 observations. Own calculations. Standard errors are in parentheses: ***significant at 1% level, ** significant at 5 % level, * significant at 10 % level. HOME items in *italics*

Among the environmental aspects at the age of 4.5 years the home items dominate in terms of predictive power. Other variables besides the HOME are less relevant compared to the results shown in table 6.5. This indicates that an integration of 101 separate HOME items provides additional information. Also the RMSECV of the IQ and persistence slightly decrease (from 1.05 to 0.97 and from 1.02 to 1.01, respectively).

Among the predictors of the IQ at the age of 8 years items of the "play materials" subscore still dominate. In difference to the age of 2 years (table 6.8), when toys that improve coordination and verbal interaction with the child were beneficial, the child should get into contact with puzzles,

numbers, letters and shapes at the age of 4.5 years. Again, the toys alone are likely not to help much. They should be complemented by adequate parental support.

Persistence has similar predictors (toys preparing to learn numbers and toys to drive). In addition the "general acceptance of the child" that was observed during the interview (not being part of the HOME score) and learning to be polite are linked to persistence. While the predictors of mood were still closely related to some of the IQ predictors at earlier ages they tend to be more distinct at the age of 4.5 years. Being able to express emotions without having to expect sanctions, owning a pet, parents setting reasonable limits and avoidance of continuous patronization are strong indicators of relevant patterns. Just like for persistence the general acceptance of the child during the interview is important.

To sum up, at the age of 4.5 years HOME items dominate among the relevant variables. While teaching the child numbers, shapes and letters (with toys) enhances the IQ. A harmonious and stimulating parent-child interaction is beneficial for mood. Persistence lies in between.

Table 6.11 shows that even at the age of 8 years several items have predictive power on future skills (at the age of 11 years). The relevant predictors of cognitive skills move away from play materials to parental support of reading, participating at courses or library visits. This confirms the presumptions made for earlier periods: Not the presence of toys and activies alone enhances skills, but the combination of possibilities to practise various tasks with adaequate parental support. The external environment now gains importance relative to the family environment. For both, the cognitive and the mental skills, teaching the child to perform certain tasks in the household (and tidying up) can be beneficial.

In late childhood the most important variables for mood shift towards other dimensions such as trips or reasonable TV use. Just like in the previous periods, emotional skills depend less on training, but more on a harmonious parent-child interaction. Motivating the child, not being depressed, setting reasonable limits and trips are related to factors enhancing the mood.

With respect to policy implications, the results should not be understood in a way that only the HOME items with the largest significant coefficients need to be addressed by interventions. They can be understood as strong signals of underlying latent factors that are crucial. Thus, a variety of play materials is a signal that the parents are keen on provding possibilities to their child that train

motor or reading abilities. Providing toys to children is important, but alone it will not help to improve their cognition. It needs to be complemented by parents assisting the child play and taking care of training their child's abilties. If children are rather low on noncognitive skills, interventions should improve the parent-child relationship in other directions. A harmonious and stimulating parent-child relationship that promotes motivation, independence and autonomy (in the household) seems particularly important.

Table 6.11: Estimation of Mood, the IQ and the activity level at the age of 11 years by environmental conditions until the age of 8 years, five largest environmental coefficients

Mood	*Parents have not been depressed in presence of child during last week*	Emotional Climate	**0.06****	0.026
	Parents motivate child to participate in interview	Emotional and verbal responsitivty	**0.05***	0.029
	Family uses TV in a reasonable way	Active Stimulation	**0.04*****	0.019
	Parents set reasonable limits for child	Promotion of Social Maturity	**0.04*****	0.017
	Parents took child on a trip of at least 50 km recently	Activities promoting development	**0.04***	0.022
	Number of components		3	
	RMSECV		1.59	
IQ	*Child is encouraged to read*	Emotional and verbal responsitivty	**0.11*****	0.034
	Family owns a dictionary and motivates child to use it	Materials, Experiences promoting developmer	**0.09*****	0.024
	Parents expect child to indepedantly do routines in household	Promotion of Social Maturity	**0.06****	0.025
	Parents promote talents of the child (in courses,institutions)	Active Stimulation	**0.05****	0.026
	Child owns a library card and parents support libary visits	Active Stimulation	**0.05*****	0.022
	Number of components		3	
	RMSECV		0.77	
Persistence	*Parents promote talents of the child (in courses,institutions)*	Active Stimulation	**0.08*****	0.019
	Family motivates and enables child to pursue a hobby	Active Stimulation	**0.07*****	0.021
	Parents didn't lose control while dealing with child during last week	Emotional Climate	**0.07*****	0.021
	Parents have not been depressed in front of child during last week	Emotional Climate	**0.05*****	0.023
	Parents expect child to tidy up	Promotion of Social Maturity	**0.05*****	0.021
	Number of components		3	
	RMSECV		0.95	

Source: Mannheim Study of Children at Risk. 360 observations. Own calculations. Standard errors are in parentheses: ***significant at 1% level, ** significant at 5 % level, * significant at 10 % level. HOME items in *italics*

6.4.3 Estimating the role of environmental aspects for social outcomes

Skills that are produced by investments determine social outcomes. Depending on the task, combinations of different skills may be necessary. It is known that cognitive skills enhance economic outcomes during the life cycle (Hanushek and Wößman, 2008). Coneus et al. (2011) show that cognitive skills are the most important of the three skills for school outcomes, mental skills are in second place and emotional skills in third place. Combined the two latter are approximately as important as cognitive skills. Thus, these aspects are not examined again. In contrast, as the study focuses on the optimal composition of various investments for the child functional levels. The PLSR offers the opportunity to estimate response variables choosing from a variety of correlated predictors, so all environmental aspects of the respective previous periods are integrated. To use

the child functional levels as responses rather than skills is also a useful way to check the robustness of the results.

Table 6.12 shows the results of regressing environmental aspects of the ages between 3 months and 8 years on the child functional levels in the family, among peers, regarding interests, in school and the child's autonomy (see section 6.2.3). Like in section 6.4.2., birth risks play an important role for predicting the outcomes at the age of 11 years. In particular, organic risk tends to have adverse long-term effects, which is in line with the results presented in table 6.7. The early HOME is not relevant, but the mother-child interaction has a significant predictive power on the child's autonomy. Parental education and breastfeeding are better predictors for school outcomes and interests. This indicates that not a stimulating environment alone promotes academic success, but also parents who may give more attention to education. Interestingly, a low parental age during all periods increases the level of autonomy while the number of persons in the household has a positive relationship in several periods.

Play materials at the ages of 2 and 4.5 years are strong predictors for all functional levels in the same way they are for skills. This confirms the relevance of parents promoting stimulation through play materials. This is in line with previous studies that examined the predictive power of HOME subscores on later academic performance (Bradley and Caldwell, 1984; Wolfgang and Stakenas, 1985). "Avoidance of punishment and restriction" does not seem to be beneficial regarding the functional level in the family, in interests and at school: The regression coefficients are significantly negative.

A certain degree of rules and restriction could enhance social progress. Single parenthood mainly has negative consequences for the functional level among peers. This again indicates possible problems that may arise due to aggregation of all HOME items. The HOME subscores at the age of 8 years all relate to high functional levels except for autonomy. Emotional climate is particularly relevant for family and peer relationships, active stimulation is the most important for interests, school outcomes and peer relations, promotion of social maturity helps for all of the five functional levels.

The high significance of HOME subscores at the age of 8 years confirms the prior results of this study: Advantageous conditions during early childhood should be followed by further investments during middle to late childhood to eventually yield positive outcomes.

Table 6.12: Estimation of the child functional levels at the age of 11 years by environmental conditions until the age of 8 years

		(1) Family		(2) Peers		(3) Leisure		(4) School		(5) Autonomy	
3 months	Interviewer rating of contact person	-0.02	(.022)	-0.03	(.023)	-0.01	(.024)	-0.03	(.021)	-0.01	(.026)
	HOME Conversation process with mother and child	-0.03	(.028)	-0.02	(.025)	-0.03	(.02)	0	(.028)	0.01	(.029)
	HOME Mother Child Interaction	-0.04	(.025)	0	(.021)	0.02	(.026)	-0.03	(.022)	-0.03	(.028)
	HOME Living Environment	0	(.024)	0.03	(.023)	0.03	(.032)	-0.01	(.022)	0	(.023)
	Low Organic Risk	**0.05** *	(.032)	**0.08** ***	(.031)	**0.11** ***	(.032)	**0.15** ***	(.032)	**0.10** ***	(.038)
	Low Psychosocial Risk	**0.06** ***	(.024)	0.03	(.022)	**0.04** **	(.026)	0.03	(.019)	0.01	(.026)
	Mother-Child Interaction (Video)	0.01	(.026)	0.03	(.024)	0.02	(.026)	0.03	(.025)	**0.07** ***	(.03)
	Neighborhood Environment	-0.03	(.024)	-0.02	(.022)	-0.02	(.022)	0	(.019)	-0.03	(.026)
	Income	-0.02	(.022)	-0.01	(.017)	0.03	(.016)	**0.04** ***	(.015)	**0.03** *	(.017)
	Single Parenthood	0.01	(.027)	-0.01	(.021)	0	(.015)	0	(.021)	0	(.025)
	Parental Age	-0.02	(.019)	-0.03	(.017)	0	(.016)	0.01	(.017)	**-0.06** ***	(.018)
	Number of Persons in Household	0.05	(.03)	0.05	(.028)	0	(.023)	0.01	(.025)	**0.05** **	(.027)
	Living with biological parents	0.02	(.024)	0.01	(.017)	0.02	(.015)	-0.01	(.018)	0	(.023)
	Parental Education	0.03	(.02)	0.02	(.02)	**0.08** ***	(.025)	**0.10** ***	(.022)	0.02	(.024)
2 years	Interviewer rating of contact person	-0.02	(.02)	-0.01	(.021)	-0.02	(.021)	0	(.019)	0	(.026)
	HOME Conversation process with parents	0.01	(.022)	-0.01	(.019)	0.01	(.021)	-0.02	(.021)	-0.01	(.023)
	HOME Stimulation of development and language	-0.03	(.02)	0	(.018)	-0.01	(.021)	-0.02	(.019)	0.01	(.023)
	HOME Living Environment	0.02	(.025)	0.02	(.024)	0.02	(.026)	0.02	(.027)	-0.02	(.03)
	HOME Avoidance and Restriction and Punishment	**-0.05** **	(.022)	0	(.016)	**-0.04** **	(.018)	**-0.04** ***	(.016)	-0.01	(.025)
	HOME Emotional Climate	0	(.019)	-0.01	(.016)	-0.02	(.017)	0	(.017)	-0.01	(.021)
	HOME Promotion of Maturation and Autonomy	-0.01	(.02)	0.01	(.018)	-0.01	(.023)	-0.02	(.018)	0.02	(.022)
	HOME Play Materials	**0.06** ***	(.026)	**0.10** ***	(.03)	**0.07** ***	(.024)	**0.08** ***	(.029)	**0.06** *	(.037)
	Neighborhood Environment	0	(.026)	-0.01	(.021)	0	(.021)	0	(.02)	-0.03	(.027)
	Income	0	(.02)	-0.03	(.018)	0.01	(.018)	0.02	(.013)	0.01	(.017)
	Single Parenthood	-0.04	(.029)	**-0.06** ***	(.024)	-0.05	(.017)	0	(.026)	0	(.028)
	Parental Age	-0.02	(.019)	-0.03	(.017)	0	(.016)	0.01	(.017)	**-0.06** ***	(.018)
	Number of Persons in Household	**-0.07** ***	(.027)	-0.05	(.024)	-0.01	(.023)	-0.01	(.026)	0.04	(.03)
	Living with biological parents	**0.04** *	(.022)	0.01	(.018)	0.02	(.013)	0.02	(.021)	**0.04** *	(.022)
	Breastfeeding	0	(.023)	0.01	(.019)	**0.06** **	(.026)	**0.05** ***	(.018)	**0.04** *	(.026)
4.5 years	Interviewer rating of contact person	-0.02	(.023)	0.01	(.023)	-0.01	(.022)	0	(.019)	0.02	(.027)
	HOME Conversation process with parents	0.04	(.029)	0.03	(.026)	0.04	(.028)	0.04	(.024)	0.04	(.027)
	HOME Stimulation of development and language	**-0.04** **	(.02)	0	(.018)	-0.01	(.02)	-0.02	(.018)	0.01	(.023)
	HOME Living Environment	0.04	(.024)	**0.04** *	(.021)	0.01	(.026)	0.02	(.02)	-0.01	(.028)
	HOME Avoidance and Restriction and Punishment	0.02	(.021)	0.01	(.022)	0	(.018)	0.01	(.018)	-0.02	(.025)
	HOME Emotional Climate	0.01	(.021)	0	(.02)	-0.03	(.023)	-0.01	(.018)	0.01	(.02)
	HOME Promotion of Maturation and Autonomy	**0.07** ***	(.022)	**0.05** ***	(.016)	0.02	(.025)	0.03	(.02)	0.02	(.025)
	HOME Play Materials	**0.06** ***	(.018)	**0.08** ***	(.02)	**0.06** ***	(.021)	**0.07** ***	(.021)	**0.06** *	(.031)
	Neighborhood Environment	-0.01	(.029)	0	(.028)	0.01	(.021)	0.02	(.027)	0.01	(.031)
	Income	-0.01	(.019)	-0.03	(.02)	0	(.019)	-0.02	(.019)	0.01	(.024)
	Single Parenthood	-0.02	(.028)	**-0.05** **	(.022)	-0.03	(.019)	-0.03	(.022)	0	(.026)
	Average Parental Age	-0.02	(.019)	-0.03	(.017)	0	(.016)	0.01	(.017)	**-0.06** ***	(.018)
	Number of Persons in Household	-0.04	(.023)	-0.03	(.02)	-0.02	(.02)	-0.01	(.021)	0.04	(.024)
	Living with biological parents	**0.04** **	(.021)	0.03	(.02)	**0.04** *	(.013)	0.02	(.02)	0.03	(.022)
	External Childcare	-0.03	(.037)	0.02	(.032)	-0.04	(.029)	-0.04	(.031)	-0.02	(.036)
8 years	HOME Paternal Engagement	0.02	(.028)	0.04	(.027)	**0.06** ***	(.026)	0.02	(.027)	-0.02	(.03)
	HOME Emotional and verbal responsitivty	**0.08** ***	(.022)	**0.07** ***	(.019)	**0.08** ***	(.023)	**0.04** *	(.019)	0.02	(.02)
	HOME Active Stimulation	**0.06** ***	(.021)	**0.09** ***	(.021)	**0.12** ***	(.022)	**0.10** ***	(.023)	0.04	(.028)
	HOME Material Environment	0	(.026)	**0.04** **	(.017)	0.01	(.026)	-0.01	(.022)	0	(.024)
	HOME Activities promoting development	0.02	(.02)	**0.04** **	(.017)	**0.06** ***	(.02)	0.03	(.02)	0.02	(.023)
	HOME Emotional Climate	**0.14** ***	(.028)	**0.07** ***	(.02)	**0.04** *	(.024)	0.02	(.023)	-0.04	(.027)
	HOME Promotion of Social Maturity	**0.06** ***	(.024)	**0.06** ***	(.021)	**0.08** ***	(.026)	**0.07** ***	(.023)	**0.07** **	(.03)
	HOME Materials and Experiences promoting development	**0.08** ***	(.023)	**0.08** ***	(.022)	**0.11** ***	(.019)	**0.10** ***	(.023)	0.03	(.027)
	Neighborhood Environment	-0.01	(.022)	0.01	(.025)	0.02	(.023)	0.03	(.024)	-0.01	(.028)
	Income	0.03	(.02)	0.02	(.019)	**0.04** *	(.019)	0.02	(.019)	0	(.026)
	Single Parenthood	0	(.031)	0	(.028)	0.02	(.024)	0.01	(.028)	0.01	(.032)
	Average Parental Age	-0.02	(.019)	-0.03	(.017)	0	(.016)	0.01	(.017)	**-0.06** ***	(.018)
	Number of Persons in Household	-0.04	(.025)	-0.03	(.024)	0.01	(.022)	0.03	(.027)	**0.05** **	(.025)
	Living with biological parents	**0.05** **	(.021)	**0.04** *	(.021)	**0.05** ***	(.014)	**0.04** *	(.021)	**0.04** *	(.027)
Number of components		3		3		3		3		3	
RMSECV		0.58		0.56		0.49		0.53		0.67	

Source: Mannheim Study of Children at Risk. 360 observations. Own calculations. Standard errors are in parentheses: ***significant at 1% level, ** significant at 5 % level, * significant at 10 % level. HOME items in *italics*

To sum up, the child functional levels can significantly be improved by avoiding organic risk. Autonomy is enhanced by the mother-child interaction. Play materials are strong predictors for functional levels just like they are for skills. Also the HOME subscore "avoidance of punishment and restriction", single parenthood and living with biological parents can play a role. At the age of 8 years a multiplicity of patterns exists. Emotional climate seems particularly relevant for family and peer relationships, active stimulation and emotional and verbal responsivity are most important for the interests, school outcomes and peer relations and promotion of social maturity enhances all of the five functional levels.

6.5 Conclusion

The socio-emotional environment is crucial for life cycle skill formation. Even though it is known that early childhood shapes the development of cognitive and noncognitive skills to a great degree, there is still a research gap on how investments could ideally be designed at different ages. This paper therefore investigates the importance of multiple socio-economic conditions on cognitive, mental and emotional skill development. This could help to develop practical toolkits for educators and politicians.

Children are followed regarding their skills and environments from in utero conditions until adolescence in the Mannheim Study of Children at Risk (MARS). To measure investments the HOME score (Home Observation Measurement of the Environment) is employed. It consists of items measuring the quality of the home environment, such as the maternal involvement with the child, acceptance of the child, play materials, variety in daily stimulation and others.

This paper provides additional evidence to the existing literature by splitting up the HOME into several dimensions. The second novel feature is the consideration of multiple additional variables apart from the HOME that contain information on the birth risks, the mother-child interaction at the age of 3 months, breastfeeding, parental characteristics, income, the neighbourhood environment and others. Third, a PLSR (partial least squares regression) is presented that can deal with many correlated predictor variables.

In a first step, the aggregated HOME subscores are analysed, in a second step the HOME is divided into many different items. I find other environmental aspects apart from the HOME to be relevant for predicting skills of future periods - especially until the age of 4.5 years. During infancy the birth risks and the mother-child interaction outperform the HOME with respect to predictive power.

The results have several implications for policy makers and educators on how educational investments could be designed. The earliest investments should aim at avoiding birth risks. Organic risk factors could be reduced by policies that improve maternal health during pregnancy. Psychosocial risk factors could be addressed by providing assistance to pregnant mothers in their socio-economic environment. Organic risk has stronger adverse long-term effects relative to psychosocial risk, which is more detrimental for skills at the age of 2 years. Thus priority should be given to avoiding organic risk as the consequences of psychosocial risk can still be weakend by later investments. After birth, policies that help to improve the mother-child interaction are important. Breastfeeding seems to be advantageous, too. Even though early investments are crucial, they need to be complemented by later investments to produce high outcomes.

For cognitive skills parents, who actively support the child in training various tasks are important. This can be achieved by providing a stimulating material environment. It changes during childhood from toys that promote motor abilities to toys that teach shapes, colours, numbers and letters. In later childhood it refers to parents enabling the child to visit courses, to attend the library or to play a music instrument. The material environment may also include other related aspects that were not mentioned, but are linked to the same latent factor. If the family does not provide an adequate support, several of those factors could be substituted by policy makers.

For noncognitive skills a harmonious and motivational parent-child relationship is important. Reasonable investments tend to be linked to the emotional climate, the reasonable integration of the child in the household and the promotion of autonomy. Children that independently take over certain tasks in the household may indicate such a relationship.

My results suggest to favour the formation of cognitive skills first as they have a strong influence on mental skills and are likelier to yield positive long-term effects. During later childhood efforts should rather aim at improving noncognitive skills, because cognitive skills stabilize faster. Exceptions should be made for children, who have either particularly low cognitive or low noncognitive skills.

The relationship of investments and the child functional levels in the family, among peers, in school, regarding interests and the child's autonomy confirms the results. In particular, the autonomy of the child is rather driven by the infant environment. Conditions during later childhood can still significantly alter the child functional level in the family or among peers. School outcomes are more closely linked to the education of the parents, suggesting that parents who give more attention to education are relevant for their child's academic performance - apart from the HOME.

As the MARS still goes on, new variables that cover socioeconomic outcomes during adulthood could be regressed on a large set of environmental variables from birth until adolescence by PLSR in the future.

6.6 References for Chapter 6

Almond, D., Edlund, L., Palme, M., 2009. Chernobyl's Subclinical Legacy: Prenatal Exposure to Radioactive Fallout and School Outcomes in Sweden. The Quarterly Journal of Economics 124, 1729-1772.

Aswani A., Bickel P., 2011. Regression on manifolds: Estimation of the exterior derivative. The Annals of Statistics 39(1), 48–81.

Bayley, N., 1969. Bayley Scales of Infant Development. Psychological Corporation, New York.

Blomeyer, D., Coneus, K., Laucht, M., Pfeiffer, F., 2008. Self-Productivity and Complementarities in Human Development: Evidence from the Mannheim Study of Children at Risk. IZA Discussion Paper 3734.

Blomeyer, D., Coneus, K., Laucht, M. Pfeiffer, F., 2009. Initial Risk Matrix, Home Resources, Ability Development and Children's Achievement. Journal of the European Economic Association, Papers and Proceedings 7(2-3), 638-648.

Blomeyer, D., Laucht, M., Pfeiffer, F., Reuß, K., 2010. Mutter-Kind-Interaktion im Säuglingsalter, Familienumgebung und Entwicklung früher kognitiver und nicht-kognitiver Fähigkeiten: eine prospektive Studie, DIW Vierteljahreshefte für Wirtschaftsforschung 79(3), 11-26.

Blomeyer, D., Laucht, M., Mühlenweg, A., 2011. Effects of Age at School Entry (ASE) on the Development of Non-Cognitive Skills: Evidence from Psychometric Data, ZEW Discussion Paper No. 11-017, Mannheim.

Borghans, L., Duckworth, A.L., Heckman, J.J., ter Weel, B., 2008. The Economics and Psychology of Cognitive and Non-Cognitive Traits, Journal of Human Resources, 43(4), 972-1059.

Boulesteix A.L., Strimmer, K., 2007. Partial Least Squares: a Versatile Tool for the Analysis of High-Dimensional Genomic Data. Briefings in Bioinformatics (8), 32-44.

Bradley, R.H., Caldwell, B.M., 1984. The Relation of Infants' Home Environments to Achievement Test Performance in First Grade: A Follow Up Study. Child Development 55(3), 803-809.

Burgemeister, B., Blum, L., Lorge, I., 1972. Columbia Mental Maturity Scale, Harcourt Brace Jovanovich, New York.

Cassel, C. M., Hackl, P.,Westlund, A. H., 1999. Robustness of partial least-squares method for estimating latent variable quality structures. Journal of Applied Statistics, 26, 435-446.

Cattell, R. B., 1960. Culture Fair Intelligence Test, Scale 1 (Handbook). 3rd ed., IPAT, Champaign, Ill.

Coneus, K., Laucht, M., Reuß, K., 2011. The Role of Parental Investments for Cognitive and Noncognitive Skill Development – Evidence for the First 11 Years of Life. Economics and Human Biology (in press).

Cunha, F., Heckman, J.J., Lochner, L., Masterov, D.V., 2006. Interpreting the Evidence on Life Cycle Skill Formation, in: E.A. Hanushek and F. Welsch (eds.), Handbook of the Economics of Education, vol. 1, Amsterdam, 697-804.

Cunha, F., Heckman, J.J., 2007. The Technology of Skill Formation. American Economic Review 97(2), 31-47.

Cunha, F., Heckman, J.J., 2009. The economics and psychology of inequality and human development. Journal of the European Economic Association 7(2-3).

Cunha, F., Heckman, J.J., Schennach, S.M., 2010. Estimating the technology of cognitive and noncognitive skill formation. Econometrica 78(3), 883-931.

De Jong, S., 1993. SIMPLS: An Alternative Approach to Partial Least Squares Regression. Chemometrics and Intelligent Laboratory Systems (18), 251–263.

Dijkstra, T., 1983. Some Comments on Maximum Likelihood and Partial Least Squares Methods, Journal of Econometrics, 22, 67–90.

Doyle, O., Harmon, C.P., Heckman, J.J., Tremblay, R.E., 2009. Investing in early human development: Timing and economic efficiency. Economics and Human Biology 7(1), 1-6.

Duckworth, A.L., Seligman, M.E.P., 2005. Self-Discipline Outdoes IQ in Predicting Academic Performance, Psychological Science 16 (12), 939-944.

Einstein, A., 1934, On the Method of Theoretical Physics, Philosophy of Science 1(2), 163-169, 165.

Esser, G., Scheven, A., Petrova, A., Laucht, M., Schmidt, M.H., 1989. Mannheimer Beurteilungsskala zur Erfassung der Mutter-Kind-Interaktion im Säuglingsalter (MBS-MKI-S), Zeitschrift für Kinder- und Jugendpsychiatrie, 17, 185-193.

Garthwaite, P. H., 1994. An interpretation of partial least squares. Journal of the American Statistical Association 89, 122–127.

Goodhall, S., Gunnell, D.J., Martin, R.M., Smith, G.D., 2007. Breastfeeding in infancy and social mobility: 60 year follow-up of the Boyd Orr cohort, Archives of Disease in Childhood 92, 317-321.

Hanushek, E. A., Wößmann, L., 2008. The Role of Cognitive Skills in Economic Development, Journal of Economic Literature 46 (3), 607-668.

Heckman, J., 2011. Effective Child Development Strategies. In E. Zigler, W. Gilliam, and W. S. Barnett (Eds.), The Pre-K Debates: Current Controversies and Issues. Forthcoming.

Helland, I. S., 1980. On the structure of partial least squares regression. Communications in Statistics, Simulations and Computation 17, 581-607.

Hoerl, A.E., Kennard, R.W., 1970. Ridge regression: Biased estimation for nonorthogonal problems. Technometrics 42 (1), 55-67.

Jöreskog, K.G., 1967. Some Contributions to Maximum Likelihood Factor Analysis. Psychometrika 32, 443-482.

Kiphardt, E. J., Schilling, F., 1974. Körperkoordinationstest für Kinder (KTK), Beltz, Weinheim.

Knight, K., 2008. Shrinkage Estimation for Nearly Singular Designs, Econometric Theory 24. 323-337

Marcus, A., Blanz G., Esser, G., Niemeyer, J., Schmidt, M.H., 1993. Beurteilung des Funktionsniveaus bei Kindern und Jugendlichen mit psychischen Störungen, Kindheit und Entwicklung 2, 166-172.

Pfeiffer, F., Reuß, K., 2008. Age-Dependent Skill Formation and Returns to Education, Labour Economics 15 (4), 631-646.

Pfeiffer, F., 2010. Entwicklung und Ungleichheit von Fähigkeiten: Anmerkungen aus ökonomischer Sicht, in: Krüger, H., U. Rabe-Kleberg, R. Kramer and J. Bude (eds.), Bildungsungleichheit revisited. Bildung und soziale Ungleichheit vom Kindergarten bis zur Hochschule, Studien zur Schul- und Bildungsforschung 30, Wiesbaden, 25-44.

Rao, C.R., Toutenberg, H., Shalabh, Heumann, D., Schomaker, M., 2008. Linear Models and Generalizations: Least Squares and Alternatives, 3rd ed. Springer, New York.

Rutter, M. Quinton, D., 1977. Psychiatric disorder – ecological factors and concepts of causation. In: Ecological factors in human development, edited by H. McGurk, North Holland, Amsterdam, 173-187.

Thomas A., Chess, S., Birch, H., 1968. Temperament and behavior disorders. New York: University Press.

Vinzi, V.E., Chin, W.W., Henseler, J. Wang, H., 2009. Handbook of Partial Least Squares: Concepts, Methods and Applications. 1st Ed. Springer. Berlin. Heidelberg.

Wold, H., 1966. Estimation of principal components and related models by iterative least squares. In: Krishnaiah PR, editor. Multivariate Analysis. New York: Academic Press, 391-420.

Wold, S., Sjöström M., Eriksson, M., 2001. PLS-regression: a basic tool of chemometrics, Chemometrics and Intelligent Laboratory Systems, 58, 109–130.

Wolfgang, C. H., Stakenas, R. G., 1985. An exploration of toy content of preschool children's home environments as a predictor of cognitive development, Early Child Development and Care, 19 (4), 291-307.

Wooldridge, J., 2003. Introductory Econometrics: A Modern Approach. Mason: Thomson South-Western.

Zimmer, R., Volkamer, M., 1984. Motoriktest für 4-6jährige Kinder (MOT 4-6), Beltz, Weinheim.