The effects of early childhood intervention on child development and early skill formation. Evidence from a randomized experiment.

Malte Sandner *

Leibniz Universität Hannover
Wirtschaftswissenschaftliche Fakultät
March 1st, 2012

Abstract

This paper presents results from a randomized evaluation of a home visiting program for disadvantaged first time mothers and their families implemented in three German federal states. At the end of the first year of the program, children in home visited families perform significantly better than those in the control families by 0.18 standard deviations in the Mental Developmental Index. Examination of gender differences revealed that home visited girls scored 0.30 standard deviations higher than girls in the control families, whereas boys scored similar in both groups. Results indicate no differences in the scores of the Psychomotor Developmental Index and the birth outcomes, despite 0.28 standard deviations higher birth weight for boys in the home visited families compared to boys in the control families. We find evidence for skill self productivity but in different magnitude for boys and girls. Furthermore, we analyze possible monetary returns of the program.

JEL-Classification: J13, J12, I21, H52
Keywords: early childhood intervention, randomized experiment

*Financial support by the German Federal Ministry for Family, Seniors, Women and Youth (BMFSFJ) and the Saxony Social Ministry is gratefully acknowledged.

1 Introduction

In recent years interdisciplinary research has emphasized the negative impact of adverse early childhood conditions for lifelong human capital accumulation. This research elaborated the following aspects: Firstly, poor maternal health, dysfunctional families, adverse childhood environments and low parenting skills have detrimental effects for child development (see Almond and Currie, 2011, for a literature overview). Secondly, because of the dynamic nature of the skill formation process, the earlier these adverse childhood conditions occur the bigger the cumulative lifelong harm (Cunha and Heckman, 2007). Thirdly, to prevent these negative conditions, parents who play an essential role for child well-being must be targeted (Heckman, 2011). Therefore, policy interventions which concentrate on children from disadvantaged families, which start early enough in life, particularly prenatal, and which alter parenting behavior are supposed to have a lasting effect on children's life outcomes and can produce high cost-benefit ratios.

Home visiting is a type of early intervention which can fulfill these requirements. In the high quality versions of home visiting trained midwifes, nurses or social pedagogues visit disadvantaged families at their own home starting already during pregnancy. The home visitors typically interact with the parents to encourage and train them to raise their children. Evidence from meta-analyses including all varieties of home visiting, e.g. programs which start after birth, shows that home visiting is effective in improving child development, but in a modest size (Sweet and Appelbaum, 2004). High quality home visiting, concentrating on disadvantaged families and starting during pregnancy, appears to be more effective for child development (Olds et al., 1999; Gomby, 2005). The few existing studies on long-term effects show that the results on child development are stable over time (Eckenrode et al., 2010).

However, up to now only medical scientists or psychologists have investigated this promising type of early childhood intervention whereas economic research has so far neglected this topic. Therefore, previous research lacks to consider efficiency questions and to investigate the influence of home visiting on skill formation dynamics. Furthermore, the previous research on high quality home visiting mainly refers to the US. The outcomes could be different in continental European countries due to a higher degree of health insurance coverage, higher welfare payments and a system of mandatory doctor visits during pregnancy.

This paper provides an econometric analysis of the first randomized experiment on high quality home visiting conducted in Germany, the *Pro Kind* Project. The *Pro Kind* Project is a longitudinal study in which disadvantaged first time mothers in three federal states are randomly assigned to a treatment group with home visits during pregnancy and the following two years and a control group. The home visits are conducted by midwives, nurses or social pedagogues. Frequency of the home visits varies between weekly and bi-weekly. 755 mothers are involved in the project.

All of them receive welfare benefits or have other financial restrictions and they additionally possess a psychological risk characteristic. Trained research assistants conducted reliable video controlled mental and psychomotoric child development tests at the age of six and twelve months. Personal interviews and hospital data provide information about birth outcomes.

The *Pro Kind* data have been examined by a team of child development psychologists before. This analysis found that children in home visited families tend to have better birth outcomes and achieve higher mental development test scores (Jungmann et al., 2010). Past research primarily consists of comparisons of means. Little attention has been paid to potential threats to the validity of the experiment, to the longitudinal structure of the data or to the dynamic process of skill formation. Treatment effect heterogeneity by gender, the distribution of treatment effects and the efficiency of home visiting received no attention. Additionally, there were deviations from the ideal experimental design in the actual implementation of Project *Pro Kind*. First randomization was done at federal level and not at community level, although it was stratified for community level. Nevertheless, due to the high heterogeneity between communities in the same federal state bias could occur. Secondly, as in most longitudinal studies with disadvantaged participants, attrition is a common problem. One third of the infants whose mothers were randomized were missing in at least one development test. These limitations of the experiment have not been adequately addressed in previous work.

We find that the *Pro Kind* Project was effective in improving children's mental development. At the end of 12 months, children from home visited families performed significantly better than those in control families by 0.18 standard deviations (SD) in the Mental Developmental Index. This treatment effect is equal to 2.5 percentage points at the median of a normal distribution. We find no effect on mental development at six months. The *Pro Kind* Project fails to significantly improve the psychomotoric skills at any stage as well as the birth outcomes. The program has differential impacts on girls and boys. For girls we find significant effects on mental development with the size of 0.28 SD and 0.30 SD already after six months and after 12 months, respectively. In contrast, boys from home visited families do not score better on the mental development index than their counterparts from control families. However, boys from home visited families are born with a 0.28 SD higher birth weight. All results are robust to different model specifications, included control variables and including development tests where children refused certain tasks. Furthermore, this paper sheds light on the skill formation process in the first years of life. For this investigation the *Pro Kind* intervention gives a unique data base because all other studies about skill formation obtain data later in childhood or less frequently. We find that self-productivity of skills occurs already in the first year of life but in different degree for boys and girls.

The rest of the paper is organized as follows: Chapter 2 provides a description

of the *Pro Kind* Project. Section 3 describes the experimental design and data collection, and Section 4 discusses the randomization results and the impact of attrition on the internal validity. Section 5 presents results on the impact of the home visiting program on birth outcomes, mental and psychomotoric development. Section 6 analyses the dynamics of the skill development. Section 7 discusses aspects of the cost effectiveness of the home visiting program. Section 8 presents conclusions.

2 Background and Description of the *Pro Kind* Project

Pro Kind is a home visiting program for disadvantaged first time mothers and their families. The intervention starts during the 12th and 28th week of pregnancy and ends at the second birthday of the child. It runs in three German federal states, two in West- and one in East Germany. Families were affiliated between November 2006 and December 2009. Midwives, nurses or social pedagogues conduct the home visits alone or in a team. Frequency of the home visits varies between weekly, bi-weekly and monthly with highest frequency close before and after birth. Home visitors use teaching materials and a guidebook to structure the domain and the aim of each home visit. Anyhow, the home visitors are free to react flexibly on the demands of the mothers and their families. All home visitors get feedback, encouragement and support from especially trained supervisors regularly. *Pro Kind* is an adaption of the Nurse Family Partnership (NFP) Program, which provided instruction for home visitation frequency, employee selection, teaching materials and guidebooks.

Improving birth outcomes and child development are major goals of *Pro Kind*. For birth outcomes personal health of the mother during pregnancy is important, for child development parental skills, e.g. that parents understand signals of their child, play an important role. To generate a healthier environment issues like smoking and balanced diet are covered in the home visits. To enhance parental skills home visitors train the parents to perceive children's signals accurately and to answer them sensitively. In order to be successful in sensitive topics like smoking or parental behavior, *Pro Kind* reverts to different psychological theories like the ecologic theory, the attachment theory and the self-efficacy theory (Bronfenbrenner, 1992; Bowlby, 1969; Bandura, 1982, 1997). For example, one way to train the parental skills is Partners in Parenting Education (PIPE) where parents get feedback on the interaction with the child from the home visitors.

Affiliation criteria for participating in *Pro Kind* are an economic constraint of the family and that the mother owns at least one social risk factor. The economic constraint is defined as receiving social welfare, unemployment benefits, an income, which is as high as social welfare or overindebtedness. The considered social risk factors are: Low education, teenage pregnancy, isolation, experienced violence or health problems. Project partners, like gynecologists, job centers and youth welfare offices referred three quarters of the participants to *Pro Kind*. About one quarter of

the participants are women who registered to the program by themselves.

A process evaluation monitored the implementation of the *Pro Kind* program. For this reason, home visitors fill in a protocol for each home visit in which the duration and the covered domain is recorded. The process evaluation finds that in average a family got 33 home visits with a minimum of 0 and a maximum of 69 and a standard deviation of 19 home visits. During pregnancy the families got 9 home visits in average. We include the participants with 0 home visits because in the analyses below we estimate intention to treat effects. Considering only families where the intervention is conducted per protocol increases the average number of home visits to 47 with a minimum of 31. The duration of an average home visit is 82 minutes. 28% of the home visits are devoted to the domain maternal health, 20% to the domain parental skills and 10% to the domain healthy child environment (see Brand and Jungmann, 2010, for details of the process evaluation).

3 Experimental Design and Data Collection

All women who were referred or came forward to *Pro Kind* filled in a short screening questionnaire to check if affiliation criteria were fulfilled. If this was the case, participants, or if underage the parents, signed an informed consent for participating in research. Afterwards, participants answered a baseline questionnaire with socio demographic and psychological questions. We use these answers to identify the number of socio demographic and psychological risk factors of each mother. After answering the baseline questionnaire women got the results of the randomization which allocates them in a home visiting group (394 women) and a control group (361 women). A computer calculated the randomization, which is stratified for communities, immigration and being underage. After randomization the control group gets regular welfare state services, an address list with support services and feedback from research. The home visiting group includes slightly more participants than the control group because the first woman in each community was automatically allocated to the home visiting group.

After the randomization home visitation begins for the intervention group. For both groups research starts with a telephone interview and a personal interview during pregnancy. Telephone interviews continue in an interval of six months until child's third birthday. They contain questions about birth outcomes, labor market participation and other socio economic outcomes of the mother and the family. Personal interviews, including child development tests, are conducted at 6, 12 and 24 months after birth. At each personal interview cognitive abilities (IQ) are measured using the Mental Developmental Index (MDI) of the Bayley Scales of Infant Development (Bayley, 1969). The fine and gross motor abilities, called the motor quotient (MQ), are also assessed at each personal interview by the Psychomotor Developmental Index (PDI) of the Bayley Scales. All tests are video taped and conducted by

blinded testers. An important advantage of the Bayley Scales is that they provide observed data as opposed to parent-reported measures of child development.

The MDI and PDI test scores are normed on hundred with a SD of 15 by an average population. For our regression analysis we standardize the test scores and birth outcomes with a mean 0 and a SD of 1. The standardization allows to compare effects on birth outcomes and test scores and facilitates to compare them with other home visiting interventions. MDI and PDI tests consist of different tasks. If the refusal or interruption rate of these tasks in one test exceeds 20 percent the reliability of the test becomes problematic. Therefore, we exclude tests with a refusal or interruption rate higher than 20 percent and include them only for robustness checks of the results. We collect birth data outcomes at two times by the telephone interviews and the personal interviews. The birth outcome data is only used when the mothers give the identical information in both interviews. Additionally, some of the birth outcomes are checked by medical records which revealed a high reliability of the self statements.

4 Baseline Comparison and Attrition

4.1 Baseline Comparison

Randomly assigning families to the home visitation program ensures that assignment is independent of mothers' and families' characteristics that may be correlated with birth outcomes and child development. If this holds, any differences in outcomes between the two groups post-intervention can be causally attributed to the intervention. To check that mother and family characteristics were indeed similar between the two groups, we run regressions of baseline mother and family characteristics on treatment status, and then verify that changes in the sample due to attrition are also uncorrelated with treatment status.

We present the comparison of mother and family characteristics at baseline in table 1. Column 1 contains the average characteristics for the control group. Columns 2 and 3 present the estimated differences between the treatment and control groups for demographic characteristics and selected psychological risk characteristics. The results in column 2 do not include any controls, while those in column 3 control for community fixed effects, because randomization was conducted at state level.

The differences in average characteristics between the control and treatment group are all practically small and mostly statistically insignificant. Migration status, defined as women who have no German citizenship or who are not born in Germany, is the only demographic characteristic which is significantly different with a higher proportion of immigrants in the control group. None of the differences in psychological risk characteristics are statistically significant. Thus, overall, the randomization appears to have been successful in creating comparable treatment and control groups.

Table 1: Sample Balance Across Treatments

	Control Mean	Treatment Difference No Controls	Treatment Difference Community Fixed Effects
	(1)	(2)	(3)
Demographic characteristics			
Age in years	21,53	-0.263 (0.316)	-0.274(0.313)
Week in pregnancy	20,3	-0.540 (0.420)	-0.528 (0.423)
Underage	0.177	0.033 (0.029)	0.035 (0.028)
Migration	0.177	-0.053** (0.026)	-0.049* (0.025)
Education risk	0.748	0.054 (0.038)	0.055 (0.038)
Income risk	0.809	0.011 (0.028)	0.012 (0.028)
Employment risk	0.856	-0.036 (0.027)	-0.040 (0.027)
No partner	0.283	0.009 (0.033)	0.004 (0.033)
Living with parents	0.267	0.014 (0.033)	0.011 (0.033)
Selected psychological risk characteristics			
Unwanted pregnancy	0.166	0.014 (0.028)	0.012 (0.028)
Isolation	0.080	-0.019 (0.019)	-0.020 (0.019)
Foster care experience	0.194	0.039 (0.030)	0.041 (0.030)
Neglect experience	0.385	-0.009 (0.035)	-0.012 (0.036)
Lost experience	0.539	-0.045 (0.036)	-0.048 (0.036)
Violence ever	0.551	0.002 (0.036)	-0.001 (0.037)
Depression	0.133	-0.031 (0.023)	-0.031 (0.024)
Anxiety	0.177	-0.007 (0.028)	-0.008 (0.028)
Stress	0.288	0.027 (0.033)	0.028 (0.034)
Aggression	0.186	-0.041 (0.027)	-0.039 (0.027)
Sum risk factors	5.864	-0.131 (0.178)	0.035 (0.028)
Observations	361	755	755

Robust standard errors shown in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Analyzing the demographic and psychological characteristics of the participants reveals that women in both groups are highly disadvantaged. For example over one third of the mothers has experienced neglect in their life time and over half of the women lost an important person during childhood. Both is related to attachment problems with the own child (Olafson, 2004). Appendix A shows a comparison between *Pro Kind* participants and first time mothers from the German Socioeconomic Panel (GSOEP) which is a representative longitudinal panel study. In this study all new mothers were asked about their children and life circumstances with a special questionnaire. The average *Pro Kind* mother is around 7 years younger than the average GSOEP first time mother. Furthermore, in the GSOEP sample 80 percent of mothers lived their first 15 years in a two parent household compared to less than 40 percent in the *Pro Kind* sample. Age and family situation during childhood are just two examples of many characteristics which prove the disadvantage of the *Pro Kind* participants. Therefore, *Pro Kind* was successful in acquiring high burdened women and families who are the target population of the intervention.

4.2 Attrition

Although, the baseline comparisons presented in table 1 show that the treatment and control groups were similar at baseline, it is possible that non random attrition

Table 2: Sample Composition

	Control	**Homevisited**	**Total**
Allocated to treatment	361	394	755
Completed 3 months Telephone Interview	265 (73.4%)	299 (75.9%)	564
Boys	*116*	*136*	*252*
Girls	*146*	*154*	*300*
Completed 6 months Cognitive-Test	237 (65.7%)	265 (67.3%)	502
Boys	*110*	*125*	*235*
Girls	*127*	*140*	*267*
Completed 12 months Cognitive-Test	205 (56.8%)	225 (57.1%)	430
Boys	*94*	*105*	*199*
Girls	*111*	*120*	*231*

from the two groups between the baseline and follow up surveys may have rendered the two groups incomparable. Table 2 shows the attrition rates for both groups and child genders. There are no statistically significant differences between the attrition rates for the control and treatment groups and between genders. In both groups 25 percent of the birth outcomes are not available. The attrition rate for the 6 months test is around one third of the baseline participants, for the twelve months test 45 percent are missing.

Table 3: Selective Attrition

	Treatment Difference MDI 6 months (1)	Treatment Difference MDI 12 months (2)
Demographic characteristics		
Age in years	0.085 (0.416)	0.173 (0.465)
Week in pregnancy	-0.205 (0.538)	-0.306 (0.583)
Underage	-0.004 (0.035)	0.004 (0.037)
Migration	-0.050 (0.032)	-0.055 (0.035)
Education risk	0.060 (0.054)	0.034 (0.045)
Income risk	0.022 (0.038)	0.017 (0.043)
Employment risk	-0.029 (0.037)	-0.051 (0.041)
No partner	0.016 (0.042)	0.053 (0.046)
Living with parents	-0.016 (0.041)	0.017 (0.045)
Selected psychological risk characteristics		
Unwanted pregnancy	0.013 (0.034)	0.048 (0.038)
Isolation	-0.013 (0.023)	0.003 (0.026)
Foster care exper.	0.015 (0.036)	0.029 (0.039)
Neglect experience	-0.003 (0.045)	0.002 (0.050)
Lost experience	-0.052 (0.046)	-0.079 (0.051)
Violence ever	-0.030 (0.046)	-0.053 (0.051)
Depression	-0.014 (0.028)	0.019 (0.032)
Anxiety	0.025 (0.033)	0.031 (0.038)
Stress	0.034 (0.043)	0.047 (0.048)
Aggression	-0.062* (0.034)	-0.085** (0.038)
Sum risk factors	-0.178 (0.220)	-0.099 (0.241)
Observations	466	388

Robust standard errors shown in parentheses. Estimates include community fixed effects.
* p < 0.1, ** p < 0.05, *** p < 0.01

Even though the attrition rates were similar for both groups, the characteristics of the attritors and non attritors could have still differed between the two groups.

We check this possibility in table 3 for the six months tests in column 1 and twelve months tests in column 2, respectively. Again we run regressions of mother and family characteristics from the baseline survey on treatment status just with the mothers and families who did not attrite. All of the differences are statistically insignificant with the exception of the proportion of mothers with aggression risk. The difference in mothers with immigration background gets insignificant at the six and twelve months interview. We, therefore, conclude that the comparability of the control and home visited families has been sustained throughout the follow up tests.

5 Estimating Program Effects

5.1 Descriptive Data

In order to allow a better interpretation of the intervention outcomes table 4 gives a combined overview of the birth outcomes and test results for treatment and control group members. A comparison of the *Pro Kind* birth outcomes with the first-borns from the GSOEP reveals that birth weight and height are similar in both samples. Hence, head circumference is statistically smaller in the *Pro Kind* sample than in GSOEP data (T=5.6). The gender difference in birth outcomes is similar to the average population. Looking at the developmental test scores reveals that the *Pro Kind* average is below the population norm of 100 points in all tests. As expected the *Pro Kind* eligibility criteria seem to be negatively related with test score results. After 12 months all test scores are closer to the norm of 100 points. Girls score significantly better than boys in MDI 6 months (T=2.1). Using the Levene-Test, variance of test scores is not significantly different between the genders at any point. Additionally, appendices B and C show density graphs of birth outcomes and child development tests scores by gender.

Table 4: Descriptive statistic

	Whole Sample		Boys		Girls	
	Mean	N	Mean	N	Mean	N
Birth Outcomes Pro Kind						
Weight in grams	3281 (554.2)	559	3384 (526.4)	252	3199 (554.3)	300
Height in cm	50.57 (3.15)	558	50.97 (3.00)	252	50.25 (3.18)	299
Head Circumference in cm	34.31 (1.93)	520	34.55 (1.82)	233	34.12 (1.95)	281
Birth Outcomes GSOEP						
Weight in grams	3253 (597.3)	825	3303 (613.7)	417	3203 (576.4)	408
Height in cm	50.86 (3.21)	824	51.20 (2.81)	417	50.51 (2.81)	407
Head Circumference in cm	35.11 (3.22)	765	35.26 (3.28)	386	34.95 (3.14)	379
6 Months Test Scores Pro Kind						
MDI	92.77 (7.93)	466	91.95 (8.43)	219	93.51 (7.38)	247
PDI	82.35 (12.37)	482	81.99 (12.90)	223	82.66 (11.91)	259
12 Months Test Scores Pro Kind						
MDI	94.22 (12.64)	393	93.90 (12.58)	187	94.50 (12.71)	206
PDI	92.57 (16.09)	375	92.68 (16.25)	169	92.48 (16.00)	206

Standard deviation in parentheses

5.2 Specification Model for Estimating Treatment Effects

We estimate the *Pro Kind* effects on child development by OLS-regression analysis using equation 1:

$$Y_{ic} = \beta_0 + \beta_1 HV_{ic} + \beta_2 h_{ic} + \alpha_c + \epsilon_{ic}, \tag{1}$$

where Y_{ic} is the outcome of child i in community c. HV_{ic} is a dummy variable indicating whether the child's family is home visited. h_{ic} is a vector of demographic and psychological family characteristics at base line. We also include a dummy variable α_c for each community to absorb the community effects. The outcomes of interest are the normalized birth weight, birth height and birth head circumference as well as the normalized MDI and PDI test scores at six and twelve months. The coefficient of interest is β_1 indicating the size of the causal effect of *Pro Kind*. The first model in each analysis includes no controls. The second model is estimated with community fixed effects. In the third model we use community fixed effects and control for age, aggression, indication of psychic problems, immigration status and total number of risk factors. The results are also robust for including more control variables up to all available base line characteristics.

We separately run regressions for boys and girls because gender is a child characteristic which is unlikely to be correlated to any family characteristic. Therefore, different intervention effects between boys and girls can be fully attributed to gender. Furthermore, reevaluations of preschool programs suggests these programs benefit girls but not boys (Anderson, 2008). Such gender reevaluations are absent for home visiting programs so far.

5.3 Impact of *Pro Kind* on Birth Outcomes

Table 5: Impact of Home Visiting on Children's Birth Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Birth Weight			Birth Height			Birth Head Circumference		
Home visiting	0.108	0.113	0.114	0.047	0.051	0.053	0.038	0.045	0.074
	(0.084)	(0.084)	(0.084)	(0.085)	(0.082)	(0.082)	(0.089)	(0.088)	(0.089)
Community fixed effects	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Household Controls	No	No	Yes	No	No	Yes	No	No	Yes
Observations	559	559	559	558	558	558	520	520	520
R^2	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.02

Robust standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

We do not find any significant effect of *Pro Kind* on birth outcomes for the whole sample. Nevertheless, the home visiting coefficient has a positive sign for all outcomes. The size of the coefficients does not vary with the model specifications, which shows that control variables are independent of the home visiting variable (table 5). Splitting the sample by gender reveals that boys in the home visiting group have a significant higher birth weight of 0.28 SD which is an increase of 4.8 percent. Birth Height and Birth Head Circumference are also increased, but not in every specification significantly (table 6). Additionally, appendix D presents density graphs of birth outcomes in treatment and control group. Appendix E shows density graphs only for boys in treatment and control group.

Table 6: Impact of Home Visiting on Children's Birth Outcomes by Gender in SD

	Boys			Girls		
	(1)	(2)	(3)	(4)	(5)	(6)
	Birth Weight					
Home visiting	0.284**	0.267**	0.290**	-0.0358	-0.0198	-0.0192
	(0.120)	(0.122)	(0.127)	(0.115)	(0.114)	(0.112)
Community fixed effects	No	Yes	Yes	No	Yes	Yes
Household Controls	No	No	Yes	No	No	Yes
Observations	252	252	252	300	300	300
R^2	0.02	0.02	0.04	0.00	0.00	0.02
	Birth Height					
Home visiting	0.219*	0.175	0.199	-0.0897	-0.0496	-0.0420
	(0.122)	(0.119)	(0.125)	(0.116)	(0.113)	(0.110)
Community fixed effects	No	Yes	Yes	No	Yes	Yes
Household Controls	No	No	Yes	No	No	Yes
Observations	252	252	252	299	299	299
R^2	0.01	0.01	0.05	0.00	0.00	0.01
	Birth Head Circumference					
Home visiting	0.148	0.136	0.213*	-0.0272	-0.0274	-0.0119
	(0.126)	(0.122)	(0.117)	(0.122)	(0.126)	(0.129)
Community fixed effects	No	Yes	Yes	No	Yes	Yes
Household Controls	No	No	Yes	No	No	Yes
Observations	233	233	233	281	281	281
R^2	0.01	0.01	0.05	0.00	0.00	0.02

Robust standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.4 Impact of *Pro Kind* on Child Development

Our analysis of home visiting effects on cognitive abilities (MDI) or fine and gross motor abilities (PDI) begins with the whole sample at the age of six months (table 7). We do not find significant positive effects of the intervention, nevertheless all coefficients are positive. The coefficients have the same size for MDI and PDI. At 12 months the coefficient on MDI becomes significant. Within the three specifications in column 4 to 6 the coefficients vary slightly in a range between 0.18 and 0.22 SD, confirming the validity of the randomization. The coefficient on PDI is still positive but insignificant and smaller than at 6 months.

Table 7: Impact of Home Visiting on Children's Development in SD

	6 Months			**12 Months**		
	(1)	(2)	(3)	(4)	(5)	(6)
	A. Mental Developmental Index (MDI)					
Home visiting	0.128	0.119	0.137	0.180*	0.208**	0.228**
	(0.093)	(0.091)	(0.094)	(0.101)	(0.098)	(0.098)
Community fixed effects	No	Yes	Yes	No	Yes	Yes
Household Controls	No	No	Yes	No	No	Yes
Observations	466	466	466	393	393	393
R^2	0.00	0.00	0.01	0.01	0.01	0.02
	B. Psychomotor Developmental Index (PDI)					
Home visiting	0.0960	0.103	0.111	0.071	0.076	0.082
	(0.091)	(0.091)	(0.092)	(0.103)	(0.104)	(0.104)
Community fixed effects	No	Yes	Yes	No	Yes	Yes
Household Controls	No	No	Yes	No	No	Yes
Observations	482	482	482	375	375	375
R^2	0.00	0.00	0.03	0.00	0.00	0.04

Robust standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Splitting the sample by gender reveals that at 6 months the coefficient of home visiting for boys is close to zero on MDI and PDI (table 8, column 1 to 3). On the other hand, girls benefit strongly, depending on the specification, by 0.28 SD to 0.25 SD. The PDI effect for girls is higher than for the whole sample, but not significant (table 8, column 4 to 6). At 12 months the home visiting coefficient on MDI for boys raises but remains insignificant, the coefficient on PDI stays close to zero (table 9, column 1 to 3). The coefficient on MDI for girls reaches a value between 0.30 and 0.34 SD. The coefficient on PDI gets larger but does not reach significance. However, the heterogeneity between the specifications is quite large which needs further investigation (table 9, column 4 to 6). Appendix F shows the density graphs for MDI at 6 and 12 Months in treatment and control group separated by gender.

Appendices G-I show estimates in which we include the tests with a task refusal or interruption rate higher than 20 percent. The newly included observations increase the sample size by 36 and 37 at six months and at 12 months, respectively. The coefficients of home visiting do not change much, indicating that refusals or interruptions are not related to treatment. Only for girls the home visiting coefficients

Table 8: Impact of Home Visiting on Children's Development by Gender at Six Months in SD

	6 Months					
	Boys			Girls		
	(1)	(2)	(3)	(4)	(5)	(6)
	A. Mental Developmental Index (MDI)					
Home visiting	-0.031	-0.036	-0.041	0.275**	0.249**	0.248**
	(0.145)	(0.136)	(0.143)	(0.118)	(0.120)	(0.120)
Community fixed effects	No	Yes	Yes	No	Yes	Yes
Household Controls	No	No	Yes	No	No	Yes
Observations	219	219	219	247	247	247
R^2	0.00	0.00	0.03	0.02	0.02	0.05
	B. Psychomotor Developmental Index (PDI)					
Home visiting	0.017	-0.012	-0.019	0.164	0.190	0.200
	(0.141)	(0.138)	(0.142)	(0.120)	(0.120)	(0.124)
Community fixed effects	No	Yes	Yes	No	Yes	Yes
Household Controls	No	No	Yes	No	No	Yes
Observations	223	223	223	259	259	259
R^2	0.00	0.00	0.04	0.01	0.01	0.02

Robust standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

are slightly smaller when including the additional observations.

Table 9: Impact of Home Visiting on Children's Development by Gender at twelve Months in SD

	12 Months					
	Boys			Girls		
	(1)	(2)	(3)	(4)	(5)	(6)
	A. Mental Developmental Index (MDI)					
Home visiting	0.0491	0.0948	0.128	0.300**	0.320**	0.342**
	(0.147)	(0.145)	(0.145)	(0.139)	(0.136)	(0.135)
Community fixed effects	No	Yes	Yes	No	Yes	Yes
Household Controls	No	No	Yes	No	No	Yes
Observations	187	187	187	206	206	206
R^2	0.00	0.00	0.06	0.02	0.03	0.03
	B. Psychomotor Developmental Index (PDI)					
Home visiting	-0.031	-0.036	-0.042	0.154	0.088	0.108
	(0.154)	(0.154)	(0.156)	(0.140)	(0.141)	(0.143)
Community fixed effects	No	Yes	Yes	No	Yes	Yes
Household Controls	No	No	Yes	No	No	Yes
Observations	169	169	169	206	206	206
R^2	0.00	0.00	0.07	0.01	0.00	0.03

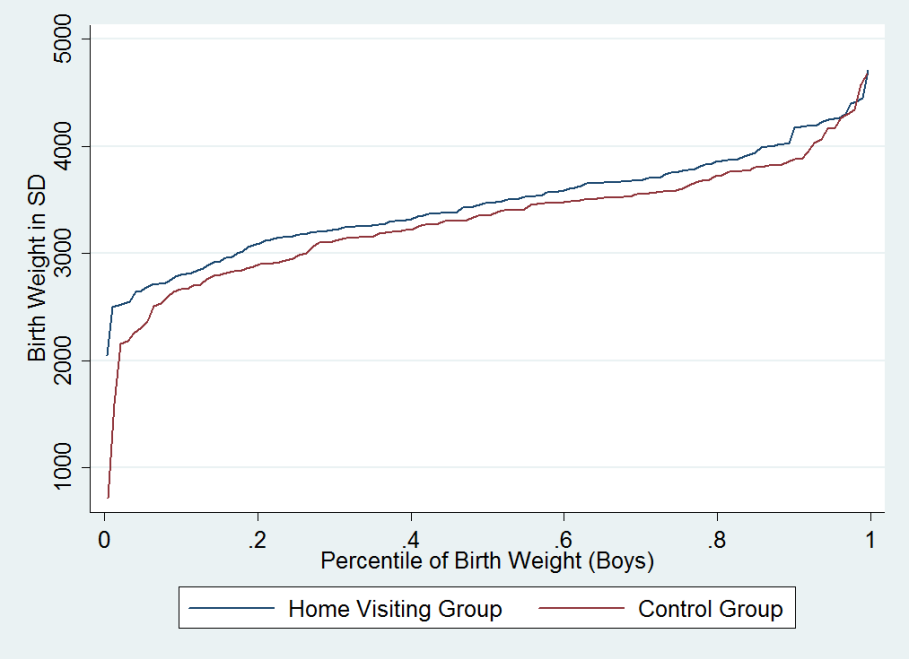
Robust standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.5 Distribution of Treatment Effects

This section investigates the treatment effects dependent on the distribution of the observed outcome. Therefore, we simply plot the observed development outcome at each percentile for control and treatment group in one figure. We only present figures for the outcomes and genders in which effects occur. Appendix J shows that home visited girls reach higher MDI test scores at 6 and 12 months at every percentile of the respective test. However, the home visiting effect on boys' birth weight seems unevenly distributed. The effect seems strongest in the first percentile (see figure 1). The number of boys with low birth weight (birth weight < 2500 grams) affirms this observation. In the home visiting group one of the boys (0,7% of the sample) is born with low birth weight and seven boys (5,8% of the sample) in the control group, respectively.

Figure 1: Distribution of treatment effects for boys birth weight



6 Skill Formation Dynamics

The *Pro Kind* experiment gives the unique possibility to analyze the skill formation process in children's first two years. The data is unique in the respect that all other studies about skill formation, which we are aware of, collect data later in children's lives or less frequently in the first two years. The knowledge we get about early skill formation can shed light on the mechanisms how home visiting generates effects and why these effects occur with girls but not with boys. Furthermore, investigating early skill formation can show if it is possible that effects of home visiting are sustained later in life.

In accordance with Cunha and Heckman (2007) skill formation is a dynamic process in which self productivity as well as direct and dynamic complementaries are important factors. Equation 2 presents the skill production function, where S_t denotes the vector of skills acquired at stage t.

$$S_{t+1} = f_t(h, S_t, HV) \tag{2}$$

Like in equation 1, h is defined as demographic and psychological family characteristic at baseline. Cunha and Heckman (2007) propose to include family investment in the production function. We use the home visiting variable HV as a proxy for family investment. Self productivity in the skill formation process arises when

$$\frac{\partial S_{t+1}}{\partial S_t} = \frac{\partial f_t(h, S_t, HV)}{\partial S_t} > 0, \tag{3}$$

i.e., when higher stocks of skills in one period create higher stocks of skills in the next period. In accordance with self-productivity, direct complementarities apply if one set of skills is productive for the formation of other skills in previous periods and vice versa. In order to have a longer observation period we also include data from the 24 months test in this chapter. Nevertheless, we use the 24 months data carefully, because the data collection has not been completed for this test, yet. The investigation methods base on Blomeyer et al. (2009) and Coneus et al. (2011) who also analyzed early childhood skill formation in the German context.

We use four stages in our approach. Our basic estimation equation for all four stages is a linear representation of the skill production function described in equation 2. In equation 4 $S^k_{t,i}$ denotes the skill indicator in t, $S^k_{t+1,i}$ denotes skills k acquired in a next period. At stage t_1 birth weight is the measure for $S^k_{t,i}$, at stage t_2,t_3 and t_4 we use 6, 12 and 24 months MDI and PDI test scores as measure for $S^k_{t,i}$

$$S^k_{t+1,i} = \gamma S^k_{t,i} + \phi HV + \eta h + \epsilon_{i,t} \tag{4}$$

Our coefficients of interest are γ indicating self productivity or direct complementaries and ϕ indicating the effects of the home visiting investment. All variables are standardized as explained in chapter 3.

Appendix K presents estimations for MDI at stage 3, where we estimate if skills acquired at 6 months are related to skills at 12 months. The specification in column 1 includes $S^k_{t,i}$ and HV only. In column 2 the specification additionally includes the demographic characteristics of h. Column 3 uses the variables of the specification in column 1 and only psychological characteristics. The specification in Column 4 includes $S^k_{t,i}$, HV and all characteristics of h. The column 5 specification excludes the $S^k_{t,i}$. Hence, the HV coefficient in column 5 and in table 7 column 6 are similar in size. Table 10 shows only the coefficients of γ and ϕ with the specification in column 4 from Appendix I. IQ indicates the use of the MDI test and MQ the use of the PDI test, respectively.

For the whole sample we find self productivity for MDI and PDI at any stage. The coefficients for self productivity rise gradually indicating that skills get more stable with age. Direct complementaries appear only at stage 3, where IQ at 24 months increases by 0.14 SD if MQ increases by one SD at 12 months. If we separate the sample by gender, the picture changes. For boys we find no self productivity for IQ at stage 2 and no self productivity for MQ at stage 3. Instead of self productivity we find direct complementarity of 6 months MQ for 12 months IQ. For girls self productivity is sustained in all stages, but additionally direct complementaries occur. The HV coefficients show the net impact of home visiting in each stage, because the estimates are controlled for the impact of home visiting in previous stages. All net impact coefficients on MQ and IQ are smaller than estimated in tables 7 to 9 with the exception of the coefficient for boys on IQ at stage 2. This coefficient is larger because HV has a negative value in stage 1.

The results of self productivity are in line with our expectations. Coneus et al. (2011) find significant self productivity for IQ from 3 months to 2 years with a coefficient of 0.3 with a slightly higher value for boys than for girls. The increase of self productivity is also in line with previous studies finding values of 0.9 at the age of 8 years which represent a high stability of skills (Cunha and Heckman, 2008). The size of the self productivity coefficients documents the relevance of early interventions. On the one hand, in each stage the skills are related to the skills from the previous stage. On the other hand, skills seem to be malleable because self-productivity does not have a coefficient close to one. The direct complementarities are surprisingly low. Further research has to reveal the reasons and the consequences for home visiting. The coefficient of HV shows that the main reason for the insufficient effect for boys lays in the first six months of home visiting. At 12 months the net effect is comparable with the girls' effect. In further steps we investigate in more detail dynamic complementaries, if home visiting effects self productivity and why home visiting has no effect on boys especially in the first six months.

Table 10: Estimates of the skill production function with two skill factors and home visiting as investment

	Whole Sample			Boys			Girls		
	IQ t-1	MQ t-1	HV	IQ t-1	MQ t-1	HV	IQ t-1	MQ t-1	HV
t = 24 Months									
IQ	0,42***	0,14*	0,19	0,53***	0,05	0,17	0,18*	0,27***	0,36*
MQ	0,01	0,39***	0,13	0,27	0,05	0,05	-0,25	0,41***	0,30
t = 12 Months									
IQ	0,27***	0,07	0,19*	0,10	0,18**	0,24	0,36***	-0,01	0,20*
MQ	0,08	0,41***	-0,02	-0,05	0,44***	-0,02	0,20**	0,42***	0,11
	Birth Weight		HV	Birth Weight		HV	Birth Weight		HV
t = 6 Months									
IQ	0,18**		0,10	0,25**		-0,15	0,18*		0,25**
MQ	0,26**		0,05	0,25*		-0,15	0,26**		0,19

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

All estimates conducted with robust standard errors and include community fixed effects

7 Cost-Benefit Analysis

We use three approaches to value the impact of *Pro Kind* on child development. The first approach benchmarks the impact of *Pro Kind* directly to NFP. The second one analyzes the relationship between increased cognitive ability at 12 months age and the probability to attain the highest German school degree. The third one examines the impact of increased birth weight on lifelong outcomes.

Since *Pro Kind* is an adaption of the NFP Program we use the results of NFP as a benchmark. An interdisciplinary research team evaluated the NFP Program in three different trails. The first trail started in Elmira in the early 1980's, the second in Memphis in 1990 and the third in Denver in 1995. All three trails used the Mental Developmental Index (MDI) of the Bayley Scales of Infant Development to assess child development. However, the tests in Memphis and Denver were conducted only at the age of 24 months and in Elmira only at the age of 12 months. Neither in the Elmira trail nor in the Memphis trail home visited infants scored significantly better at MDI (Olds et al., 1986; Kitzman et al., 1997). Only the Denver trail revealed an impact of NFP on the child mental development (Olds et al., 2002). Home visited children scored 4 points higher on a scale with the population mean of 100, which is higher than the *Pro Kind* effect (2.3 points) at 12 months age. Nevertheless, in all three trails other program effects occurred like lower childhood injuries or fewer subsequent pregnancies.

The time period for follow up research is different for each trail. In the Elmira trail data is available for 19 years, in Memphis for 12 years and in Denver only for 4 years. In Elmira home visiting reduces reported serious antisocial behavior and emergent use of substances for the home visited adolescents at age of 15 and 19. (Olds et al., 1998; Eckenrode et al., 2010). The only measure for school success was high school graduation at age 19, where the intervention caused no effect (Eckenrode et al., 2010). In contrast, the program not only reduced antisocial behavior but also

improved the academic achievement of children at age 12 in Memphis. The four years follow up in Denver showed that home visited children scored better in a series of cognitive tasks focusing primarily on the children's capacity for sustained attention and inhibitory control (Olds et al., 2010). Therefore, it appears that the effect on child development is lasting at least for a time span of four years.

Comparing the costs of the interventions *Pro Kind* and NFP are in a similar range. The NFP program in the Denver trail costs about \$11.511 in 2006 (Olds et al., 2010) and *Pro Kind* costs approximately €8.790 in 2010 which is \$11.866 assuming an exchange rate of 1.35 €/\$. The monetary benefits caused by the Elmira and Memphis trail are higher than the program costs whereas the major part of the benefits occurred by changes in maternal life course. Only in Elmira less anti social behavior of the home visited children also played a role. Since MDI was not significantly changed at age one it seems that other domains influence anti social behavior more strongly. In contrast, MDI could strongly influence school success. Then, the missing change in MDI could explain the weak results on school performance in the Elmira trail. To estimate the impact of MDI on school success and to consider the German setting of *Pro Kind* another data base is needed.

Therefore, in the second approach we use data from the German Mannheim Risk Study (MARS). MARS is a longitudinal epidemiological cohort study following infants at risk from birth to adulthood. The initial sample contains 382 children born between February 1986 and February 1988. The MDI of the Bayley Scales was used to assess children's cognitive development at age of three months and 24 months. This gives the unique possibility to analyze the relationship between early cognitive development and later school success in a German context. Analyzing the MARS data (Coneus et al., 2011) show that an increase in cognitive development by one SD at 24 months increases the probability of attaining a high school degree by 13 percentage points. To assess the economic relevance of the increase we continue with a numerical calculation.

In Germany the life earning premium for attaining a high school degree is €230.548 (Fritschi and Oesch, 2008). The Net Present Value of this amount is €118.837 assuming a discount rate of 1.5 percent, an entry age of 25 and 40 years workforce participation. Furthermore, we assume that the 0.18 SD effect on cognitive development remains stable until the age of 24 months. Then the *Pro Kind* effect on cognitive development increases the probability to attain high school by 2.34 percent. This means on average a higher life time income of €2.780,79 for the home visited, which is 32 percent of the intervention costs. Until now we have no data about how increased cognitive development is related to the probability of dropping out or class repetition. These adverse school events could have a high relevance in the *Pro Kind* sample. Hence, they are in the focus of our next investigations.

In the third approach to benchmark the results of *Pro Kind* we investigate the life long impact of the increased birth weight for boys. Although, the home visiting

has not strongly affected boys cognitive and psychomotoric development yet, birth weight could affect later life through channels which are not presented here, for example an improved health status. Because birth weight is an outcome which is easy to obtain, many studies examine the relation between birth weight and later life outcomes. In a study including 33,366 twin pairs Black et al. (2007) find that a 10% increase in birth weight increases probability of high school completion by 1.2% and increases IQ for males at the age of 20 by 1.1%. Other studies also using twin samples (Almond et al., 2005; Behrman and Mark R. Rosenzweig, 2004) find causal links between birth weight and health outcomes. Overall, evidence shows that the effects on birth weight and especially low birth weight can contribute to the monetary returns of home visiting.

8 Conclusion

Home visiting for disadvantaged families is proven to be effective for child development in the US. The analysis of the *Pro Kind* project using the bailey scales of infant development and birth outcomes as measure for child development shows that this is also true for continental Europe. We find a better cognitive development at the age of 12 months. However, program effects on cognitive development are concentrated on girls. Girls in the treatment group achieve higher test scores at six and twelve months than their counterparts in the control group, whereas boys show no difference. However, the intervention is effective to increase birth weight only for boys. This effect occurs mainly due to a reduced number of boys with low birth weight in the home vising group. Our findings of gender differences in cognitive development are in line with reevaluations of other early childhood interventions like the Perry Preschool program, where the intervention is also exclusively effective for girls.

The effects of *Pro Kind* on child development are robust to several specifications. We use community fixed effects estimations because randomization was done on state level. Furthermore, we estimate models with different family baseline characteristics as controls. The home visiting effect is hardly influenced by any specification. Since more than 20% refusals or interruptions of tasks within one test affect reliability of the overall test result we excluded these tests from the main estimations. If we include these tests, effects of home visiting on child development are still robust with marginal changes.

We also investigated the dynamic nature of the skill formation process because of its importance for the interpretation of the effect size. We showed that self productivity is present at all stages. We do not find direct complementarities between MDI and PDI. After estimating separate models for boys and girls we find strong differences in the skill formation process, which could explain some of the gender differences in the effectiveness of *Pro Kind* for cognitive development. Furthermore, we analyzed heterogeneous treatment effects and we revealed that *Pro Kind* is ef-

fective to raise the birth weight in the lowest percentile for boys. No heterogeneous treatment effects are found for cognitive development.

Considering the question if the size of the *Pro Kind* effect on child development is meaningful, psychological and behavioral literature would claim that effect sizes below 0.2 SD are small. Nevertheless, the *Pro Kind* effect at 12 months could have large lifelong impact because of the dynamic nature of the skill formation process. For example the cognitive development at 24 months is strongly related to high school graduation, which is a strong indicator for life income. Especially for girls the effect size, which is classified as moderate, could have a lifelong impact. The meaning of the effect size is also enlarged because the home visitors do not directly interact with the child but with the mothers. Thus, it is likely that the mother uses the acquired skills also in other domains, like her own or her child's health or in planning her own life course. Furthermore, there could be spill-over effects of the acquired skills for the second child because *Pro Kind* just affiliates first time mothers.

References

Almond, D., K. Chay, and D. Lee (2005). The costs of low birth weight. *The Quarterly Journal of Economics 120*(3), 1031–1084.

Almond, D. and J. Currie (2011). *Human capital development before age five*, Volume 4b of *Handbook of Labor Economics*. Elsevier.

Anderson, M. L. (2008). Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association 103*(484), 1481–1495.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Journal of Psychology* (37), 122–147.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W.H. Freeman.

Bayley, N. (1969). *Bayley Scales of Infant Development*. New York: Psychological Cooperation.

Behrman, J. R. and Mark R. Rosenzweig (2004). Returns to birthweight. *The Review of Economics and Statistics 86*(2), 586–601.

Black, S., P. Devereux, and K. Salvanes (2007). From the cradle to the labor market? : the effect of birth weight on adult outcomes. *The Quarterly Journal of Economics 122*(1), 409–439.

Blomeyer, D., K. Coneus, M. Laucht, and F. Pfeiffer (2009). Initial risk matrix, home resources, ability development, and children. *European Economic Association: Journal of the European Economic Association 7*(2), 638–648.

Bowlby, J. (1969). *Attachment and loss: Vol. 1 Attachment*. New York: Basic Books.

Brand, T. and T. Jungmann (2010). Abschlussbericht der Implementationsforschung zum Modellprojekt Pro Kind.

Bronfenbrenner (1992). *The ecology of human development: Experiments by nature and design*. Cambridge and Mass.: Havard University Press.

Coneus, K., M. Laucht, and K. Reuß (2011). The role of parental investments for cognitive and noncognitive skill formation—Evidence for the first 11 years of life. *Economics & Human Biology*, in press.

Cunha, F. and J. Heckman (2008). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources 43*(4), 738–782.

Cunha, F. and J. J. Heckman (2007). The technology of skill formation. *The American Economic Review 97*(2), 31–47.

Eckenrode, J., M. Campa, D. Luckey, Ch. Henderson, R. Cole, H. Kitzman, E. Anson, K. Sidora-Arcoleo, J. Powers, and D. Olds (2010). Long-term Effects of Prenatal and Infancy Nurse Home Visitation on the Life Course of Youths. *Archives of Pediatrics and Adolescent Medicine 164*(1), 9–15.

Fritschi, T. and T. Oesch (2008). *Volkswirtschaftlicher Nutzen von frühkindlicher Bildung in Deutschland: Eine ökonomische Bewertung langfristiger Bildungseffekte bei Krippenkindern*.

Gomby, D. S. (2005). *Home Visitation in 2005: Outcomes for Children and Parents*. Washington DC.

Heckman, J. J. (2011). *The American family in black and white: a post-racial strategy for improving skills to promote equality*, Volume 5495 of *IZA Discussion Paper*. Bonn.

Jungmann, T., V. Kurtz, T. Brand, S. Sierau, and K. von Klitzing (2010). Präventionsziel Kindergesundheit im Rahmen des Modellprojektes Pro Kind: Vorläufige Befunde einer längsschnittlichen, randomisierten Kontrollgruppenstudie. *Bundesgesundheutsblatt 53*, 1180–1187.

Kitzman, H., D. Olds, C. Henderson, C. Hanks, R. Cole, and R. Tantelbaum (1997). Effects of prenatal and infancy home visitation by nurses on pregnancy outcomes, childhood injuries, and repeated childbearing: A randomized controlled trial. *The Journal of the American Medical Association* (278), 644–652.

Olds, D., Ch. Henderson, R. Cole, J. Eckenrode, H. Kitzman, D. Luckey, L. Pettitt, K. Sidora, P. Morris, and J. Powers (1998). Long-term Effects of Nurse Home Visitation on Children's Criminal and Antisocial Behavior: 15-Year Follow-up of a Randomized Controlled Trial. *Journal of the American Medical Association 280*(14), 1238–1244.

Olds, D., C. Henderson, H. Kitzman, J. C. R. Eckenrode, and Tatelbaum R.C (1999). Prenatal and infancy home visitation by nurses: Recent findings. *The Future of Children* (9), 44–63.

Olds, D., Henderson C., Chamberlin R., and Tatelbaum R. (1986). Preventing child abuse and neglect: A randomized trial of nurse home visitation. *Pediatrics 78*, 65–78.

Olds, D., H. Kitzman, R. Cole, C. Hanks, and K. Arcoleo (2010). Enduring Effects of Prenatal and Infancy Home Visiting by Nurses on Maternal Life Course and Government Spending: Follow-up of a Randomized Trial Among Children at Age 12 Years. *Arch Pediatr Adolesc Med. 2010;164(5):419-424 164*(5), 419–424.

Olds, D., J. Robinson, R. O'Brien, D. Luckey, L. Pettitt, C. Henderson, R. Ng, K. Sheff, J. Korfmacher, S. Hiatt, and A. Talmi (2002). Home Visiting by Paraprofessionals and by Nurses: A Randomized, Controlled Trial. *Pediatrics 110*(3), 486–496.

Sweet, M. A. and M. I. Appelbaum (2004). Is Home Visiting an Effective Strategy? A Meta-Analytic Review of Home Visiting Programs for Families With Young Children. *Child Development 75*(5), 1435–1456.

A Comparison between *Pro Kind* Mothers and SOEP first time mothers

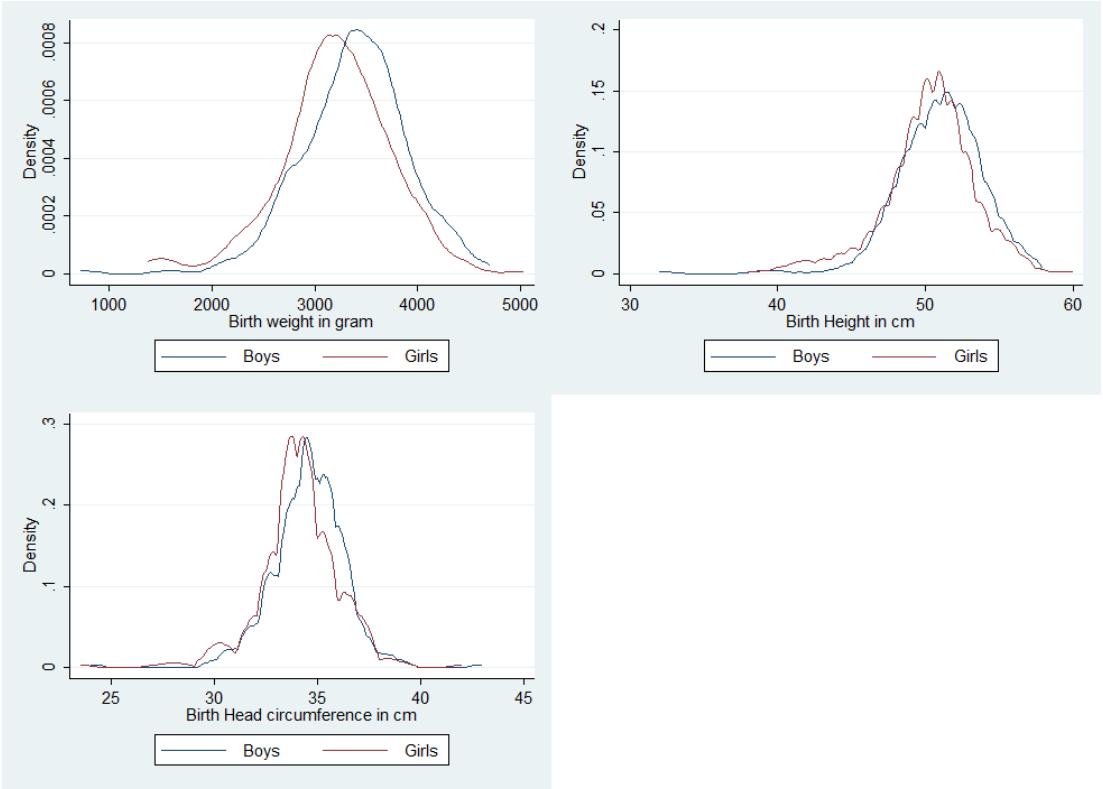
Figure 2: Comparison between *Pro Kind* Mothers and SOEP first time mothers

	TG		CG		SOEP	
	n	%	n	%	n	%
Economic sitituation during pregnancy:						
- Alg II-Transfer	246	74,8%	223	74,6%		
- Alg-Transfer	8	2,4%	6	2,0%	68[1]	10,2%
- HLU-Transfer	16	4,9%	11	3,7%		
- Overindebtedness	5	1,5%	6	2,0%	9	1,3%
- Little income	22	6,7%	18	6,0%	14	1,8%
- Participant in training	18	5,5%	26	8,7%	34	5,1%
- No economic risk factor noticed	14	4,3%	9	3,0%	553	81,4%
- ***Total***	***329***		***299***		***664***	
Age:						
- in years, at start of the intervention	21,48 (4,36)		21,57 (4,39)		28,3 (5,50)	
Total	***343***		***314***		***662***	
Country of birth						
- Germany	307	89,5%	261	83,1%	568	88,8%
- Turkey	4	1,2%	2	0,6%	13	2,0%
- East-Europe	12	4,1%	23	7,3%	41	6,5%
- Others	20	5,2%	28	9,0%	18	2,7%
Total	***343***		***314***		***640***	
Living situation within the first 15 years of life						
- With both parents	131	38,4%	130	41,8%	333	81,0%
- At least one year with just one parent	160	46,9%	142	45,7%	67	16,3%
- At least one year in foster care or with foster parents	50	14,7%	39	12,5%	11	2,7%
Total	***341***		***311***		***411***	
Family status						
- Unmarried	294	85,7%	273	87,5%	179	33,9%
- Married	41	12,0%	33	10,6%	329	62,3%
- Divorced/widowed	8	2,3%	6	1,9%	20	3,8%
Total	***343***		***312***		***528***	
School graduation in the year of pregnancy						
- Upper-track secondary qualifying for university entry (Fachhochschulreife/Abitur)	19	7,5%	23	10,3%	238	42,2%
- Intermediate-track secondary (Realschule)	79	31,2%	63	28,1%	217	38,7%
- Lower-track secondary (Hauptschule)	94	37,2%	88	39,3%	66	11,8%
- School for special needs/ other graduation	15	5,9%	9	4,0%	14	2,5%
- Left school without graduating	26	10,3%	20	8,9%	3	0,5%
- Graduation in a foreign country	10	4,0%	11	4,9%	22	3,9%
- Still attending school	10	4,0%	10	4,5%	2	0,4%
Total	***253***		***224***		***561***	

TG=Treatment group, CG= Control group

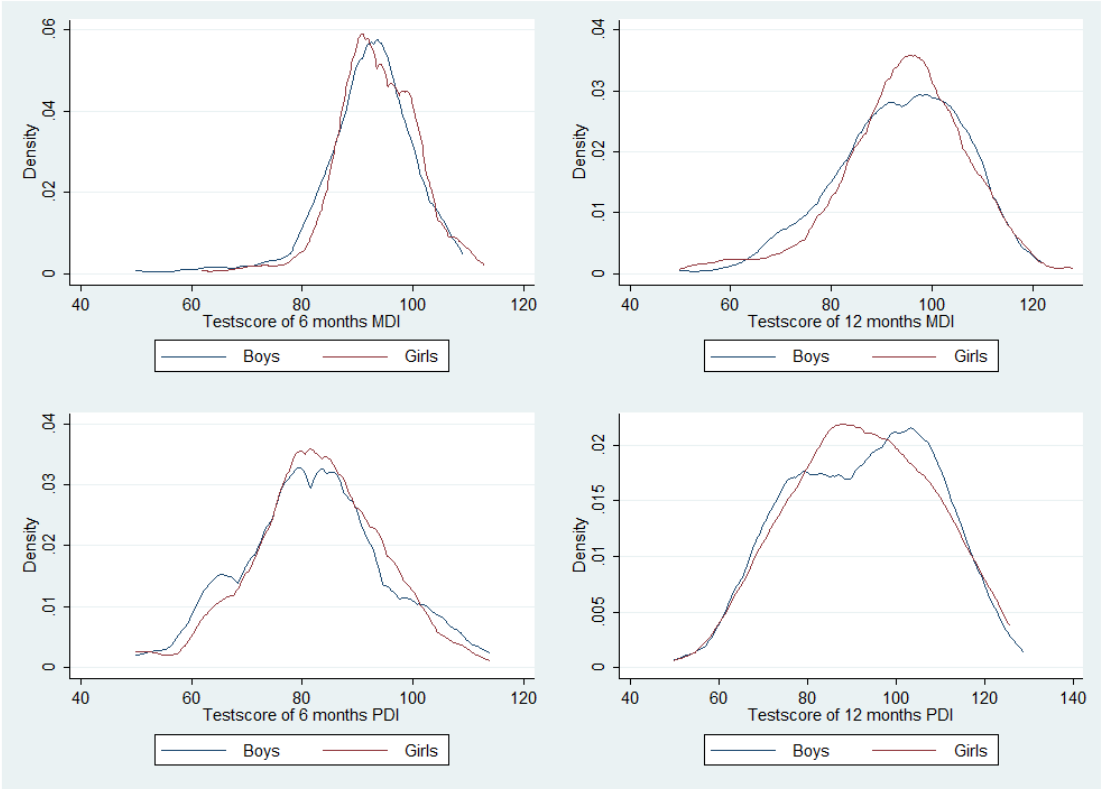
B Density Birth Outcomes for Boys and Girls

Figure 3



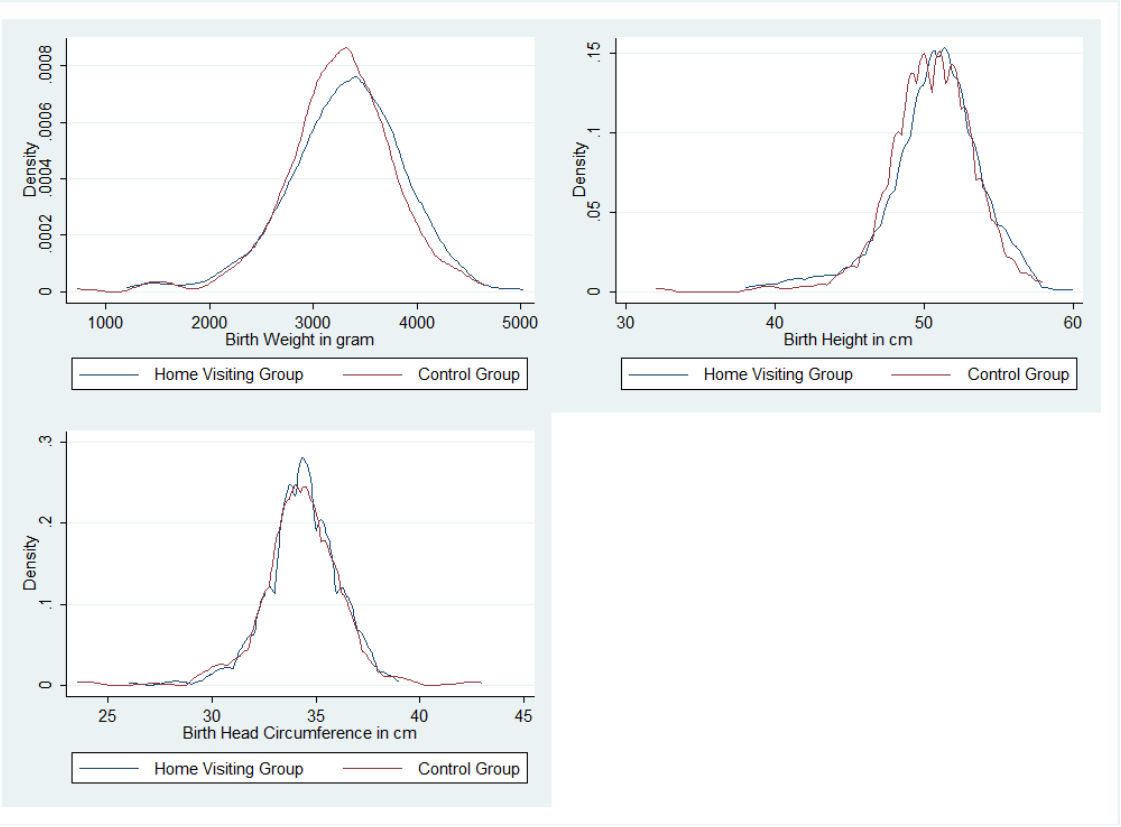
C Density Bayley Scales for Boys and Girls

Figure 4



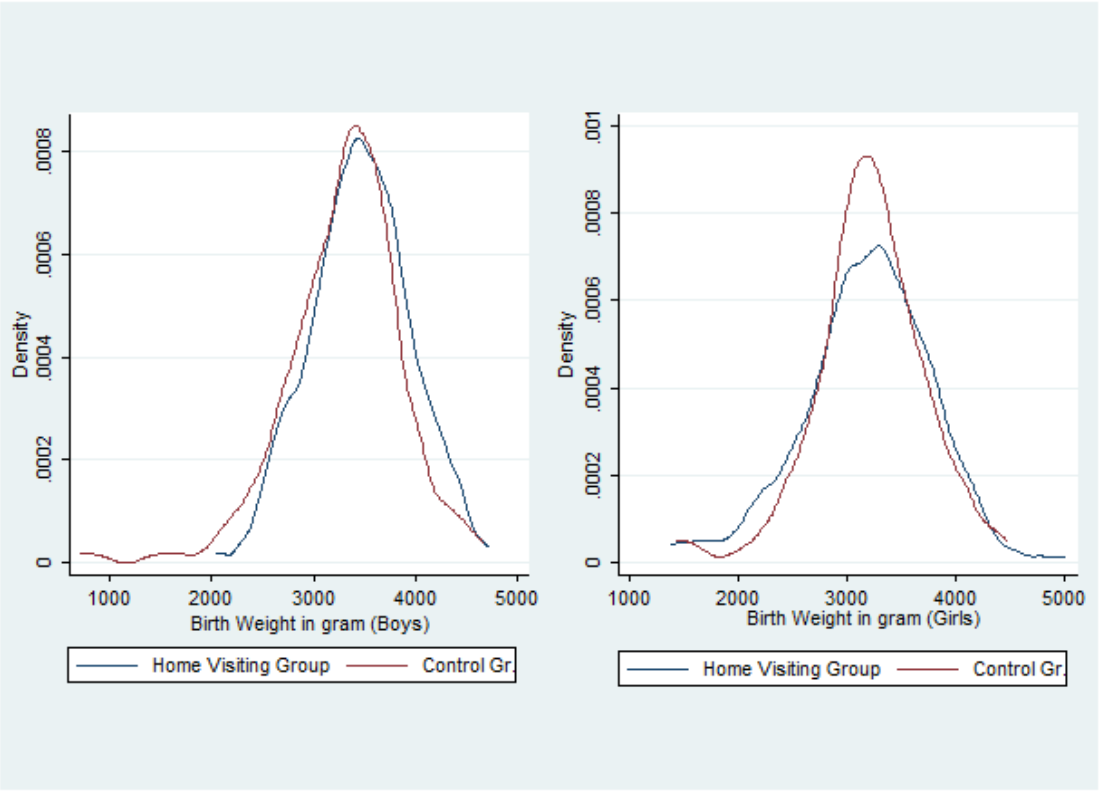
D Density Birth Outcomes for Treatment Group and Control Group

Figure 5



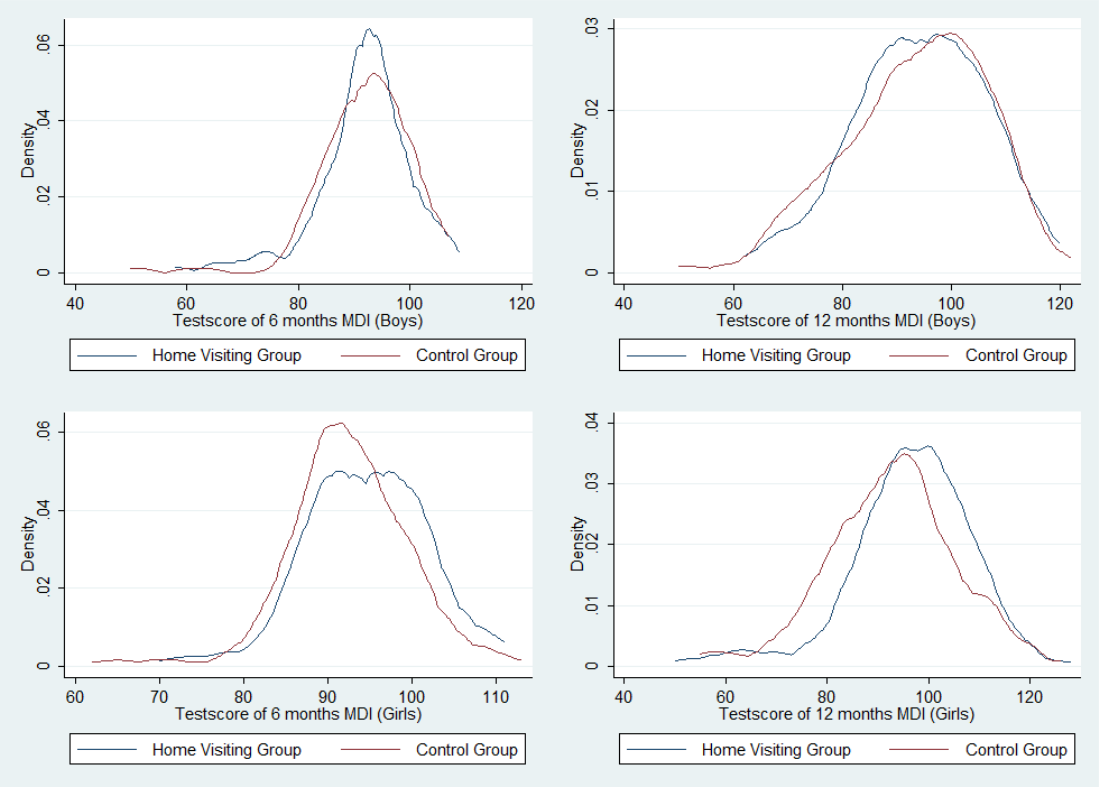
E Density Birth Outcomes for Treatment Group and Control Group (Boys and Girls)

Figure 6



F Density MDI for Treatment Group and Control Group(Boys and Girls)

Figure 7



G Child development including all tests

Table 11: Impact of Home Visiting on Children's Development in SD with all Tests

	6 Months			**12 Months**		
	(1)	(2)	(3)	(4)	(5)	(6)
	A. Mental Developmental Index (MDI)					
Home visiting	0.116	0.111	0.130	0.188*	0.200**	0.225**
	(0.089)	(0.088)	(0.090)	(0.097)	(0.094)	(0.096)
Community fixed effects	No	Yes	Yes	No	Yes	Yes
Household Controls	No	No	Yes	No	No	Yes
Observations	502	502	502	430	430	430
R^2	0.00	0.00	0.01	0.01	0.01	0.02
	B. Psychomotor Developmental Index (PDI)					
Home visiting	0.109	0.118	0.123	0.0589	0.0621	0.0767
	(0.090)	(0.089)	(0.091)	(0.096)	(0.096)	(0.097)
Community fixed effects	No	Yes	Yes	No	Yes	Yes
Household Controls	No	No	Yes	No	No	Yes
Observations	499	499	499	431	431	431
R^2	0.00	0.00	0.03	0.00	0.00	0.04

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

H Child development including all tests

Table 12: Impact of Home Visiting on Children's Development by Gender at Six Months in SD with all Tests

	6 Months					
	Boys			Girls		
	(1)	(2)	(3)	(4)	(5)	(6)
	A. Mental Developmental Index (MDI)					
Home visiting	0.0268	0.0183	0.0150	0.196*	0.167	0.167
	(0.138)	(0.133)	(0.139)	(0.116)	(0.116)	(0.114)
Community fixed effects	No	Yes	Yes	No	Yes	Yes
Household Controls	No	No	Yes	No	No	Yes
Observations	235	235	235	267	267	267
R^2	0.00	0.00	0.04	0.01	0.01	0.03
	B. Psychomotor Developmental Index (PDI)					
Home visiting	0.0459	0.0211	0.000340	0.165	0.191	0.200
	(0.137)	(0.134)	(0.138)	(0.118)	(0.120)	(0.122)
Community fixed effects	No	Yes	Yes	No	Yes	Yes
Household Controls	No	No	Yes	No	No	Yes
Observations	233	233	233	266	266	266
R^2	0.00	0.00	0.05	0.01	0.01	0.04

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

I Child development including all tests

Table 13: Impact of Home Visiting on Children's Development by Gender at twelve Months in SD with all Tests

	12 Months					
	Boys			**Girls**		
	(1)	(2)	(3)	(4)	(5)	(6)
	A. Mental Developmental Index (MDI)					
Home visiting	0.111	0.147	0.201	0.256**	0.277**	0.288**
	(0.147)	(0.145)	(0.147)	(0.127)	(0.126)	(0.125)
Household Controls	No	No	Yes	No	No	Yes
N	199	199	199	231	231	231
R^2	0.00	0.01	0.05	0.02	0.02	0.03
	B. Psychomotor Developmental Index (PDI)					
Home visiting	0.0274	0.0262	0.00733	0.0850	0.0372	0.0721
	(0.143)	(0.141)	(0.143)	(0.131)	(0.130)	(0.131)
Household Controls	No	No	Yes	No	No	Yes
N	199	199	199	232	232	232
R^2	0.00	0.00	0.07	0.00	0.00	0.03

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

J Distribution of cognitive treatment effects

Figure 8: Distribution of treatment effects for girls mdia

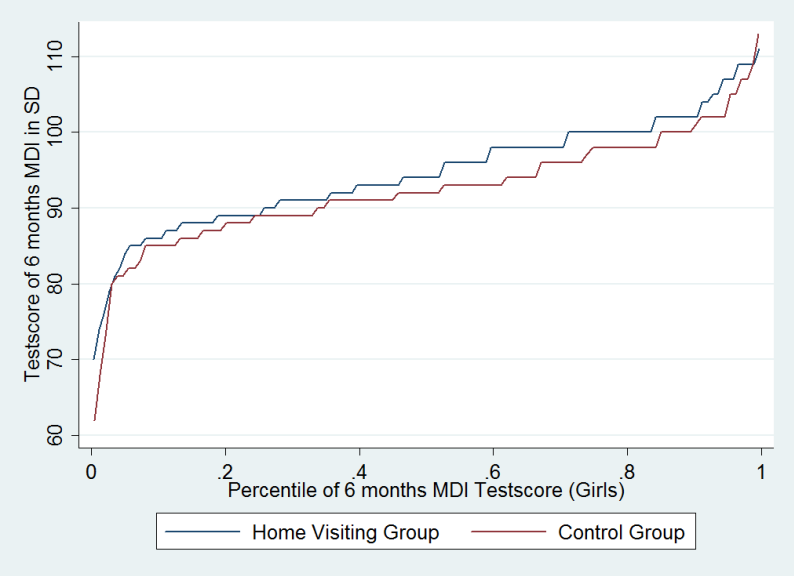
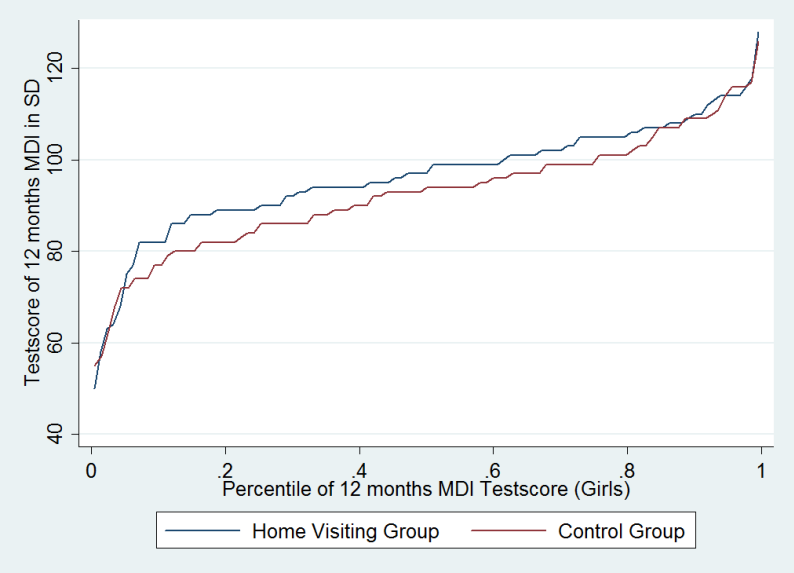


Figure 9: Distribution of treatment effects for girls mdib



K Self productivity and direct complementaries

Table 14: Self productivity and direct complementaries between 6 and 12 months

	(1) z2mdib	(2) z2mdib	(3) z2mdib	(4) z2mdib	(5) z2mdib
grup_kod	0.174*	0.177*	0.187*	0.191*	0.240**
	(0.094)	(0.095)	(0.097)	(0.098)	(0.098)
z2mdia	0.277***	0.269***	0.280***	0.268***	
	(0.058)	(0.061)	(0.062)	(0.066)	
z2pdia	0.0763	0.0800	0.0676	0.0650	
	(0.061)	(0.064)	(0.063)	(0.066)	
geschlecht		-0.0177		-0.0337	-0.0721
		(0.098)		(0.098)	(0.102)
altinjahr		0.00883		0.00371	-0.00323
		(0.016)		(0.017)	(0.015)
sswa		-0.00851		-0.00793	-0.00568
		(0.008)		(0.008)	(0.009)
alt_risk		0.00604		0.0764	0.136
		(0.154)		(0.194)	(0.194)
immi		0.0722		0.0720	0.0821
		(0.148)		(0.159)	(0.158)
edu_risk		-0.0954		-0.00815	-0.0452
		(0.146)		(0.177)	(0.195)
inc_risk_em		-0.146		-0.0805	-0.156
		(0.125)		(0.179)	(0.184)
beruf_risk		0.0251		0.109	0.0435
		(0.122)		(0.163)	(0.170)
part_risk		0.0299		0.153	0.160
		(0.116)		(0.167)	(0.162)
wohn_eltern		0.117		0.0899	-0.00904
		(0.135)		(0.151)	(0.162)
wunsch_risk_em			-0.0255	-0.0292	0.0210
			(0.153)	(0.176)	(0.172)
soz_risk_em			0.133	0.142	-0.148
			(0.205)	(0.239)	(0.261)
miss_risk_em			-0.103	-0.00497	-0.113
			(0.133)	(0.195)	(0.191)
lost_risk_em			0.167	0.215	0.304**
			(0.113)	(0.157)	(0.148)
gewalt1_risk_em			-0.122	-0.111	-0.00168
			(0.177)	(0.192)	(0.189)
gewalt2_risk_em			0.0733	0.130	0.0615
			(0.127)	(0.173)	(0.179)
psych_risk_em			0.141	0.175	0.236
			(0.141)	(0.195)	(0.188)
dass_da_risk_em			-0.0364	-0.0409	-0.140
			(0.182)	(0.216)	(0.236)
dass_aa_risk_em			0.149	0.187	-0.0310
			(0.159)	(0.192)	(0.198)
dass_sa_risk_em			0.133	0.173	0.239
			(0.145)	(0.180)	(0.174)
faf_risk_em			-0.0180	0.0631	-0.0310
			(0.158)	(0.191)	(0.194)
risk_ges_em			-0.0366	-0.0781	-0.0544
			(0.045)	(0.114)	(0.114)
N	342	342	342	342	393
R^2	0.11	0.13	0.13	0.14	0.07

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$