HCEO WORKING PAPER SERIES

Effects of Early Childhood Intervention on Fertility and Maternal Employment: Evidence from a Randomized Controlled Trial

Malte Sandner

Working Paper

2017-074

10/2017

HUMAN CAPITAL AND
ECONOMIC OPPORTUNITY
GLOBAL WORKING GROUP

The University of Chicago
1126 E. 59th Street Box 107
Chicago IL 60637

www.hceconomics.org

Effects of Early Childhood Intervention on Fertility and Maternal Employment: Evidence from a Randomized Controlled Trial

Malte Sandner*

University College London and Institute for Employment Research (IAB)

August 2017

Abstract

This paper presents the results of a randomized study of a home visiting program implemented in Germany for low-income, first-time mothers. A major goal of the program is to improve the participants' economic self-sufficiency and family planning. I use administrative data from the German social security system and detailed telephone surveys to examine the effects of the intervention on maternal employment, welfare benefits, and household composition. The study reveals that the intervention unintentionally decreased maternal employment by 8.7 percentage points and increased subsequent births by 6.6 percentage points, in part through a reduction in abortions.

JEL-Classification: J13, J12, I21, H52
Keywords: Early Childhood Intervention, Randomized Experiment, Fertility

*Financial support from the German Federal Ministry for Family, Seniors, Women and Youth (BMFSFJ) and the Saxony Social Ministry is gratefully acknowledged. I thank the German Record Linkage Center at the Institute for Employment Research and, in particular, Manfred Antoni, Johanna Eberle, and Barbara Hofmann. I also thank all employees of the *Pro Kind* foundation and all members of the *Pro Kind* research group, in particular Anja Becher and Peter Lutz, for support. I also would like to thank Orazio Attanasio, Gabriella Conti, Thomas Cornelissen, Christian Dustmann, James Heckman, Simon Janssen, Melissa Kearney, Uta Schönberg for helpful comments, and participants at the following conferences and seminars are acknowledged: the HCEO Workshop on Human Capital Formation 2016, the Early Childhood Inequality Workshop 2016 in Nürnberg, the Essen Health Conference 2016, the 2016 ASSA conference, the 2015 conference of the German Economic Association, the 2015 EEA conference, and the 2015 SOLE/EALE conference. This study is registered in the AEA RCT Registry and the unique identifying number is: "AEARCTR-0002135".

1 Introduction

Home visiting programs targeted towards disadvantaged families are a type of early childhood intervention that not only aims to improve child outcomes but also to improve maternal outcomes, such as economic self-sufficiency, employment and family planning. To affect these outcomes, family midwives consult mothers at their homes for a longer period after birth to enhance maternal skills (e.g., attachment behavior, interactions, and teaching skills) and to increase the women's personal strengths, including self-efficacy, problem-solving abilities and self-esteem. Home visiting programs are popular in many developed countries. For example, in the United States the federal government initiated the Maternal, Infant, and Early Childhood Home Visiting Program (MIECHV Program) to support state implementation of early childhood home visiting programs. Between 2012-2015, the annual founding for the MIECHV program was $500 million and it served more than 300,000 disadvantaged families and children (U.S. Department of Health and Human Services, 2015). In the U.K. home visiting programs deliver services to 16,000 disadvantaged new families each year (U.K. Department of Health, 2013).

Although home visiting programs aim to improve aspects of the maternal life course, it is arguable whether they will achieve this aim. On the one hand, the intervention could be successful and lead to higher maternal participation in the workforce by improving mothers' awareness of their personal strengths. Due to their higher occupational aspirations, the mothers may decide to delay further births. On the other hand, the intervention could increase women's satisfaction with their maternal role by improving their maternal skills. Greater maternal satisfaction and well-being could increase fertility or the length of time mothers want to stay at home with their child. Both could lead to longer absences from the workforce. It is therefore an open empirical question which of the two effects predominate.

This paper presents the first economic analysis investigating the effects of one such home visiting program on maternal outcomes by exploiting a randomized controlled trial (RCT). The program, named *Pro Kind*, delivers home visits starting during pregnancy until two years after birth for disadvantaged first-time mothers. *Pro Kind*

is the German adaptation of the Nurse Family Partnership (NFP) program which is widely considered the primary MIECHV Program.

The results suggest that the intervention *increased* fertility and *decreased* employment. The effects are sizable, implying that among the intervention group, the probability of a second birth increased by 36 percent (from 18.3 to 24.9 percent) and employment decreased by 24 percent (from 50.7 to 42.0 percent). The effect on fertility is mainly explained by a reduction in abortions in this group of women. I can exclude more favorable family environments, such as more stable partnerships in the treatment group, as a potential mechanism for the fertility effect. However, the intervention positively influenced maternal perceived stress and life satisfaction indicating that mothers in the treatment group are more satisfied with their maternal role. This higher satisfaction may have caused that the mothers choose to stay longer at home and decide against abortions.

My analysis draws on both administrative data from the German social security system that include information on employment, wages, welfare benefits and household composition and survey data from biannual telephone and face-to-face interviews. The administrative data are available for over 90% of the sample over the first three years after the birth of the first child. These data are objectively measured and should not be biased by the treatment and control groups differentially reporting outcomes. The survey data allow for the examination of a much richer set of outcomes, such as conception, abortions, childcare use, and subjective statements about perceived stress and life satisfaction, allowing me to identify channels for the findings derived from the administrative data. To my knowledge, this is the first study combining administrative and survey data to comprehensively evaluate the effects of an early childhood intervention on maternal outcomes.

The results from *Pro Kind* question the findings from previous U.S. medical studies using RCTs that home visiting, in particular NFP, successfully decreased fertility and welfare dependency and increased maternal employment (Olds et al., 2007, 1997; Brooks-Gunn et al., 1994). These results received extensive attention and the fiscal savings due to lower welfare payments represent one reason why governments expand home visiting programs (e.g. Miller, 2015). Because the content, implementation,

and participants were very similar in the *Pro Kind* and U.S. studies, a compelling explanation for the different results may be specific arrangements of the welfare state. For example, the German welfare programs include means-tested welfare payments which do not include work obligations or benefit cuts until the child's third birthday. Additionally, financial incentive programs that are effective to encourage work among low-income families with children, such as the Earned Income Tax Credit (EITC) in the U.S. (e.g. Hoynes and Patel, 2015), do not exist in Germany. This welfare state environment provides few incentives for maternal workforce participation; therefore, the intervention's impact on maternal skills and life satisfaction might dominate over its impact on personal strengths, leading to longer maternity leave periods and subsequent births instead of higher employment. In contrast, in a more strict welfare environment the effect on personal skills may dominate because mothers gain a larger utility increase from early employment. Since welfare arrangements for mothers strongly vary between U.S. states, home visiting may have similar effects on life course as the *Pro Kind* intervention in states with more generous welfare.[1]

The findings on fertility and maternal employment are highly relevant because they contribute to the understanding of how early childhood interventions generate effects on children. For example, shorter spacing between births has negative effects on the test scores of older siblings (Buckles and Munnich, 2012). A literature in economics has shown that children from larger families tend to have lower educational attainment, lower IQ scores, poorer employment outcomes, and a greater likelihood of engaging in risky behavior (Kessler, 1991; Hanushek, 1992; Black et al., 2010). Although the observation period is currently too short to analyze the effects on completed fertility, the results on spacing may reduce the potential for home visiting to improve child development. On the contrary, mothers decided to work less after birth and reported higher well-being, which both can positively affect child development (Bernal and Keane, 2011; Carneiro et al., 2015; Berger and Spiess, 2011). The findings of Sandner and Jungmann (2017) that the *Pro Kind* program has smaller positive effects on infant development compared with studies in the U.S. suggest that,

[1]Welfare payments strongly differ across U.S. states. For example, Stanley et al. (2016) show that maximum Temporary Assistance for Needy Families (TANF) benefits vary between 11% of the of federal poverty level in Mississippi to 47% in New York.

at least in the short term, the development reducing effect predominates.

The findings on maternal welfare dependency and employment decisions are also of high short-term fiscal relevance since disadvantaged mothers receive a substantial amount of total welfare spending in many developed countries.[2] In addition, Dahl et al. (2014) showed that increases in the welfare participation of the current generation can affect the participation behavior of the next generation as well. Following this finding, the effects of home visiting on maternal welfare might also increase the receipt of welfare by their children and have long-term fiscal consequences. However, a cost-benefit perspective that considers the positive effects on life satisfaction and fertility (assuming home visiting increases completed fertility) may conclude that despite higher welfare expenditures, the life course effects may add more benefits than costs.[3]

Finally, the findings of the paper provide new insights into the fertility decision of disadvantaged women. Several studies identified policies and economic circumstances, such as media influence, regional income inequality, regional unemployment rates, and to a lesser extend welfare policies, affecting the fertility behavior of disadvantaged women (Kearney and Levine, 2015a, 2014; Anant et al., 2013; Grogger and Bronars, 2001). However, most studies in this literature only concentrated on birth as fertility outcome and did not identify which role contraceptive use or abortions played in contributing to the changes in birth rates. In contrast, the *Pro Kind* study observes, in addition to births, also contraceptive use, conception, abortions, and miscarriages. The results therefore can provide a much broader insight than previous studies through which channels policies and economic circumstances affect fertility of young low-income women. As the *Pro Kind* results suggest the abortion margin plays an important role when disadvantaged women decide about fertility.

The remainder of the paper is organized as follows: Section 2 reviews the existing literature on the effects of home visiting on maternal life course. Sections 3 and 4 provide descriptions of the *Pro Kind* program, the experimental design, the baseline

[2]For example, in Germany in 2008, families with children younger than three received €4.7 billion in welfare payments.

[3]Several recent articles analyzed the cost-effectiveness of pro-natalist interventions, such as providing public childcare (Bauernschuster et al., 2016), increasing child benefits (González, 2013), and maternity leave benefits (Raute, 2017) in low fertility countries. In comparison to these interventions *Pro Kind* increased fertility at a much lower costs.

sample, and the data used in this study. Section 5 proves the validity of the experimental design, and Section 6 presents the estimation strategy. Section 7 shows the results, and Section 8 compares them with the results of U.S. studies. Section 9 provides concluding remarks.

2 Related Literature

Few studies in the literature have examined the impact of early childhood programs in general, and of home visiting programs in particular, on parents. For example, Heckman et al. (2010, 2013) evaluated 715 outcomes of the famous Perry Preschool Program; although home visits were part of the intervention, none of these outcomes focused on parents. However, the effects on parents might be one undetected link affecting the success of the program. The only program in which effects on parents were systematically evaluated is the NFP in the U.S. This program is conceptually identical to the *Pro Kind* program, and like the *Pro Kind* program, it aims to increase maternal economic self-sufficiency.

The NFP was evaluated in three randomized controlled trials located in Elmira, New York, in 1980; in Memphis, Tennessee, in 1990; and in Denver, Colorado, in 1995. All trials enrolled unemployed and low-income first-time mothers (Olds et al., 1997, 2010, 2004), and both the maternal life course and child outcomes were of prime interest. The availability of follow-up outcome data varies among the trials and ranges from four years to 15 years of data. The NFP literature shows a reduction in the rates of subsequent pregnancies and births and an increase in the intervals between first and second pregnancies and births in all three trials for the first four years after mothers entered the program. A shortcoming of the studies is that they do not present information on whether less sexual activity or more frequent contraception use led to fewer pregnancies, and they do not present information about abortions.

In all three trials, the intervention reduced women's use of welfare, and in two of the three trials, the intervention increased maternal employment. More stable partnerships and the reduction in subsequent births are channels to explain the effects on welfare and employment. Long-term follow-up revealed that the impacts on the

maternal life course did not diminish over the years. The intervention did not affect the mothers' school graduation rates in any of the trials, although higher school attendance was observed in the Elmira trial. Appendix Table A1 summarizes the three trials' results regarding the maternal life course. A recent reanalysis of the NFP Memphis trial using permutation-based inference and correcting for multiple-hypothesis testing (MHT) finds that the significance of the effects on subsequent birth do not survive MHT (Heckman et al., 2017). However, the significance of the effect on welfare dependency persists also in this more rigorous analysis. Only one study (Brooks-Gunn et al., 1994) other than the NFP analyzed the effects of home visiting on the maternal life course using a randomized experiment. In that study, home visiting significantly reduced maternal unemployment.

Cost/benefit analyses of the Elmira and Memphis trials indicate that the NFP reached the fiscal break-even point via its effects on the maternal life course, even before considering effects on the children. In Elmira, the program cost of $3,133 was outweighed by discounted savings of $3,246 (expressed in 1980 U.S.-$) by child age four. The main reason for these savings was increased maternal employment (Olds et al., 1993). In Memphis, the NFP resulted in $12,300 in discounted savings per intervention compared with the program's cost of $11,511 (both expressed in 2006 U.S.-$) by child age twelve. Higher maternal employment and lower government spending on food stamps, Medicaid, AFDC, and TANF generated the savings (Olds et al., 2010). These results show that home visiting programs, and the NFP in particular, have strong effects on the maternal life course and that these effects are fiscally relevant in the U.S. context.

3 The *Pro Kind Program*

3.1 Background

The home visiting program *Pro Kind* is an adaptation of the previously described NFP program, which provides instructions for home visit frequency, employee selection, teaching materials, and guidebooks (see Jungmann et al., 2009; Olds, 2006, for additional information on the *Pro Kind* program and NFP). The intervention begins

between the 12th and 28th weeks of pregnancy and ends at the child's second birthday. Family midwives and a pediatric nurse conduct the home visits either continuously or in a tandem model with social pedagogues (Brand and Jungmann, 2012).[4] Teaching materials and visit-by-visit guidelines structure the theme and aim of each home visit. Nevertheless, home visitors have the flexibility to adapt the contents to maternal needs and the familial situation. All home visitors got intensive training in the German translations of the NFP guidelines and materials. Additionally, all home visitors regularly receive feedback, encouragement, reflection, and support from nurse supervisors who obtained NFP training in the U.S. The frequency of the home visits varies, according to the NFP model prescription, between weekly, biweekly, and monthly visits, with the highest visit frequency occurring directly before and after birth. Overall, 52 home visits with an average duration of 90 minutes are scheduled between pregnancy and the child's second birthday.

The *Pro Kind* program registers only first-time mothers between their 12th and 28th weeks of gestation. All participants must receive social welfare or unemployment benefits, have an income that qualifies them for social welfare benefits or have excessive debt. Additionally, all participants must have one of the following social risk factors: a low educational level, teenage pregnancy, isolation, health problems, or having been a victim of violence. However, none of these risk factors were binding because all applicants with economic constraints had at least one of them. Project partners, such as gynecologists, job centers, pregnancy information centers, and youth welfare offices, referred approximately 75% of the participants to *Pro Kind*, and approximately 25% of participants self-registered in the program.

The *Pro Kind* program was implemented in three German federal states at 13 implementation sites between 2006 and 2012 (see Appendix Table A2). Although the chosen sites are not fully representative of Germany, the communities cover both rural and urban regions as well as regions in both East and West Germany. This mixture of sites ensures that the program is implemented under varying regional conditions in terms of availability of childcare, healthcare provision, and labor market conditions.

[4]In NFP family nurses conduct the home visits. Since this profession does not exist in Germany, in *Pro Kind* mainly family midwifes, who have a much boarder training than normal midwives, conduct the home visits. As I show later, the profession of the home visitor does not influence the effects of the intervention.

A major goal of the *Pro Kind* program is to improve families' economic self-sufficiency by helping parents develop a perspective for their future and make appropriate decisions about planning future pregnancies, finishing their education, and finding employment. The question arises why home visiting in general, and *Pro Kind* in particular, would produce effects in these domains. This question is especially crucial because the German welfare state offers generous benefits to the mothers of infants and toddlers. For example, there are no work obligations or welfare cuts as long as the child is under three years old, even when childcare is available (German Federal Employment Agency, 2014). As a result, there are few incentives for mothers to participate in the labor market.

The main answer why the *Pro Kind* program could have additional effects on maternal life course and employment is given by the relationships that the home visitors develop with the mothers during their pregnancies and their children's early years. The mother's first time experience of giving birth is the strongest factor that initiates and deepens this relationship. Olds et al. (2010) state that through this relationship, nurses can help parents gradually gain a sense of mastery for overcoming challenges and position themselves to create the kind of life they want. Furthermore, mothers with newborns are often open-minded to guidance during this fundamental life transition, during which they make important choices that shape the trajectories of their lives and those of their children. Thus, the home visitors' ability to build relationships and meet clients in an open-minded life situation are home visiting programs' greatest advantages over other interventions.

Besides increasing economic self-sufficiency, the intervention also aims to improve maternal parenting, child development and maternal and child health. Previous evaluations on the other goals of the *Pro Kind* program found positive effects on child cognitive development, which were lower than in NFP and concentrated on girls, on certain maternal skills, and on maternal depression (Sandner and Jungmann, 2017; Sandner et al., 2017; Sierau et al., 2015).

3.2 Randomization Process and Sample Description

The causal effects of the *Pro Kind* intervention were evaluated using a randomized controlled trial. At the beginning of the randomization process, all women answered a brief screening questionnaire, typically by telephone, to assess whether they fulfilled the affiliation criteria. If a woman met the criteria, the supervisor visited the woman at her home. During this visit, the participant (or, if she was underage, her parents) signed an informed consent form for participation in the study. Thereafter, participants completed a baseline questionnaire to assess demographic and psychological characteristics, as well as risk factors. Until this point, the mothers had received only information on the research study and as little information as possible on the home visits to minimize the "John Henry" effect for mothers in the control group.[5] After answering the baseline questionnaire, women received the results of the randomization that assigned them either to the home visit or the control group. The final sample for the *Pro Kind* experiment consisted of 755 mothers, of whom 394 were assigned to the treatment group and 361 to the control group.

After randomization, mothers in both research groups had access to the regular German welfare state services. They received an address list with support services in their communities and monetary incentives for participating in the study.[6] Therefore, families in the control group also received more support than the average first-time low-income family in Germany. However, only women in the treatment group received the *Pro Kind* home visits, and no other comparable home visiting services were available in any of the communities.

Table 1 reports the means and the differences in means according to treatment status for the baseline variables.[7] Differences in the average characteristics of the control and treatment groups were small and not statistically significant. Hence, overall, the randomization appears to have successfully created comparable treatment and control groups.

[5]The "John Henry" effect explains an unexpected outcome of an experiment caused by the control group's knowledge of its role in the experiment. This knowledge encourages the group to perform differently and often better than they would have otherwise, eliminating the effect of the experimental manipulation (Salkind, 2010).

[6]The monetary incentive was €15 for the each telephone interview.

[7]I use sample means or values from a multivariate imputation procedure in the case of missing values for baseline variables. However, complete data are available for most variables (see Appendix Tables A3 and A4). The present missing values are equally distributed between the control and treatment groups and results hardly change when non-imputed data are used.

Table 1: Descriptive Statistics

	Control Group Means (1)	**Treatment Group Means** (2)	**Treatment vs. Control** (3)
Demographic Characteristics			
Age in Years	21.53	21.27	-0.27 (0.31)
Week in Pregnancy	20.30	19.76	-0.53 (0.42)
Teenage	0.44	0.47	0.03 (0.04)
Foreign Nationality	0.08	0.07	-0.01 (0.02)
HH-Income per Month (€)	916.62	937.28	17.54 (40.60)
Debt Over €3,000	0.17	0.19	0.02 (0.03)
No Graduation	0.75	0.78	0.03 (0.04)
Low Income	0.81	0.82	0.01 (0.03)
No Employment	0.86	0.82	-0.04 (0.03)
No Partner	0.28	0.29	0.01 (0.03)
Unmarried	0.93	0.90	-0.03 (0.02)
Living with Parents	0.27	0.28	0.01 (0.03)
Persons in HH	2.45	2.55	0.09 (0.12)
Psychological and Physical Characteristics			
Unwanted Pregnancy	0.17	0.18	0.01 (0.03)
Daily Smoking	0.34	0.34	-0.01 (0.03)
Social Isolation	0.08	0.06	-0.02 (0.02)
Foster Care Experience	0.19	0.23	0.04 (0.03)
Experience of Neglect	0.39	0.38	-0.01 (0.04)
Experience of Loss	0.54	0.49	-0.05 (0.04)
Experience of Violence, ever	0.09	0.08	-0.01 (0.04)
Depression	0.13	0.10	-0.03 (0.02)
Anxiety	0.18	0.17	-0.01 (0.03)
Stress	0.29	0.31	0.03 (0.03)
Aggression	0.19	0.14	-0.04 (0.03)
Med. Indicated Risk Preg.	0.11	0.11	-0.01 (0.02)
Body Mass Index (BMI)	25.31	25.22	0.16 (0.39)
Sum Risk Factors	5.86	5.73	0.04 (0.03)
Observations	361	394	755

Notes: Column 3 contains estimates of the average difference in characteristics between mothers in the control and treatment group. Dependent variables shown in the first column. The treatment indicator has the value one if the mother is in the treatment group. Robust standard errors are reported in parentheses in Column 3. See Appendix Tables A3 and A4 for variable definitions.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

An analysis of the demographic and psychological characteristics of the participants reveals that the women in both groups were young and highly disadvantaged. Most of the mothers were unemployed at the time of the baseline interview and had never been regularly employed. The low employment levels seem to be a consequence of the fact that a high percentage of the mothers (approximately 75%) had less than eleven years of schooling; many of them dropped out of school. Furthermore, the average monthly household income was €928.60. Considering the mean household size of 2.49 persons, the participants' average income was below the poverty line in Germany. These figures indicate that *Pro Kind* was successful in recruiting families on welfare and those with low education levels, who were the target population of

the intervention.

3.3 Utilization of *Pro Kind* Home Visiting

To monitor the quality of the program implementation, the home visitors documented each visit (e.g., duration, covered topics, and maternal interest).[8] In total, 12,894 home visits with an average duration of 82 minutes were conducted. The families in the treatment group received 32.7 home visits on average (SD = 19, range: 0-94). Only 9 of the 394 mothers in the treatment group received no home visits. Because participation in the *Pro Kind* program is voluntary, 166 (42.2%) mothers decided to leave the program before the child's second birthday (main reasons: no further interest [n = 68], not reachable [n = 37], and moving away from a *Pro Kind* community [n = 28]). Considering only families who received the *Pro Kind* home visits until the child's second birthday increases the average number of home visits to 45.3 (SD = 10.7, range: 11-94) showing that the intervention was well implemented for families who stayed until the end of the program.

In NFP and subsequently *Pro Kind* each home visit is devoted to a specific topic which is defined in the home visit guideline books. Of the 42 scheduled home visits after birth, eight visits focus on contraception and family planing, six on life goals, and five on financial issues. The materials of the home visits with the topic contraception and family planing mainly focus on using safe contraception methods. The visits with the topic life goals focus on how the mother can combine motherhood and apprenticeship or employment and on exercises to increase self-efficacy. The materials for financial issues explain, for example, how to keep a book of household accounts. Although the home visitors had the flexibility to adapt the contents to maternal needs and the familial situation, the regular supervision monitored that all topics were covered and all materials were used. The home visitors' documentation about how much time in the families was spent on certain topics demonstrates that the declared rates are close to the NFP guidelines (Appendix Table A5). Additionally, the participants stated in interviews, which measured the quality of the program implementation, that all materials were regularly used during the home visits (Brand

[8] See Brand and Jungmann (2014) for further description of program design and implementation.

and Jungmann, 2011).

4 Data

4.1 Administrative Data

The German Record Linkage Center (GRLC) of the Institute for Employment Research obtained individual-level labour market biographies from the German social security system and matched them to the treatment indicator and date of affiliation based on the participants' full name, full address, and date of birth.[9] The data contain information on maternal outcomes, such as employment, type of employment, wage, welfare benefit use, job search, age, community of residence and household composition. Studies that have also used these German social security data in other settings include, for example, Schmieder et al. (2012), Card et al. (2013) and Dustmann et al. (2009). From the submitted information of 740 participants, the GRLC was able to track 703 participants to their labor market biographies.[10] For all the tracked participants complete labor market biographies are available from their first time labor market entry until 36 months after the birth of the treatment child.[11] My primary outcomes of employment and welfare use thus have an effective postrandomization "attrition rate" of 7%. Only household composition, which I use as measure of fertility, has a slightly higher "attrition rate" of 11% because the information was available only if the mother was either engaged in a job search or received welfare benefits.[12]

4.2 Survey Data

In addition to the administrative data, I use extremely detailed survey data from biannual telephone interviews and from a face-to-face interview at 12 months after birth. The telephone interviews begin during pregnancy and continue at six-month intervals

[9]Staff of the GRLC (www.record-linkage.de) linked the data. Questions concerning the linkage can be directed to the GRLC. The GRLC receives funding from the German Research Foundation (grant number: BE3172/1-2).

[10]Of the 755 participants in the baseline sample, 15 refused to provide informed consent and were not used for the merging process.

[11]Throughout this article, the treatment child indicates the the mother's first child, who was in focus of the intervention.

[12]Information on community of residence is only available if the mother was employed, engaged in job search, or received welfare benefits. The information is not available if the mothers simply "stayed at home" without being employed, looking for a job or receiving welfare benefits.

until the child's third birthday. The interviews are computer-assisted and contain all of the questions that are recommended when using the German Socioeconomic Panel (SOEP) as a reference data set, including questions on the participants' household, income, employment, childcare use, family planning and partnership. Furthermore, the interviews contain the SOEP activity calendar to record the participants' employment status on a monthly basis, questions about use of contraceptives, and the SOEP mother-child questionnaire to record maternal attitudes towards each newborn child of the mother (Anger et al., 2009). The face-to-face interview contains besides psychological child development scales questions about maternal life satisfaction and perceived stress.

The telephone interviewers attempted to contact all mothers at each point in time, except in cases of miscarriage or infant death. To guarantee a high participation rate, the interviewer attempted to contact the participant four times within two months near the interview date. If no contact could be made during this time span, the interviewer attempted to contact the mother for the next scheduled interview four months later. If contact was made for this interview, a combined interview regarding the time span for the two interviews was conducted. However, no interview covered a period longer than 12 months to avoid recall bias. Therefore, some participants missed one or two telephone interviews and continued to participate in subsequent telephone interviews. The main reasons for missed interviews were switching telephone numbers or refusing to participate. Overall, nearly 80% (n = 602) of the mothers were interviewed at least once after pregnancy, and for 71% (n = 539) of the mothers, data are available for at least 12 months after birth. Moreover, 39% (n = 296) participated in all interviews without missing data for any months after birth. Data from the face-to-face interview is available for 56% (n = 422) of the mothers.

5 Validity of the Experimental Design

Differences in attrition or in the pre-randomization characteristics of the treatment and control analysis samples would raise concerns regarding the validity of the experiment for identifying causal inference. Therefore, Table 2 summarizes the sample

Table 2: Sample Composition Administrative Data and Telephone Survey Data

	Control Mean (std. dev.) for Full Sample (1)	Difference Between TG and CG (2)
Panel A: Administrative Data		
Consent to Merging	0.986	-0.012
	(0.117)	(0.010)
		[0.257]
Merged	0.945	-0.026
	(0.229)	(0.018)
		[0.162]
Panel B: Telephone Interview Data		
At Least One Interview After Birth	0.784	0.026
	(0.412)	(0.029)
		[0.381]
Data Available for 12 Months After Birth	0.698	0.030
	(0.460)	(0.033)
		[0.357]
Data Available for 24 Months After Birth	0.557	0.045
	(0.497)	(0.036)
		[0.214]
Complete Data from Birth Until Third Birthday	0.380	0.024
	(0.486)	(0.036)
		[0.500]
Panel C: Face-to-Face Interview Data		
Face-to-Face Interview 12 Months after Birth	0.548	0.020
	(0.498)	(0.036)
		[0.580]
Observations	*361*	*755*

Notes: Robust standard errors in parentheses and p-values in brackets. TG = Treatment Group; CG = Control Group.

composition from the administrative (Panel A, Column 1) and survey data (Panel B and C, Column 1) and analyzes the treatment-control balance (Column 2). The results in Column 2 indicate no significant differences between the treatment and control groups in the response rate for either the merged administrative data or the survey data.

Table 3 presents the differences in the baseline demographic characteristics between the treatment and control groups for the administrative data (Column 1) and the telephone interview data grouped by data availability (Columns 2-5). Appendix Table A6 shows the differences in psychological characteristics. The results reveal that the attrition only slightly reduced the equal distribution of the baseline characteristics.[13] Sandner and Jungmann (2017) show that no selective attrition occurs for

[13]Appendix Table A7 shows that some characteristics and risk factors differed between those who dropped out and those who participated in the follow-up interviews. Generally, the participating mothers were older and had fewer cumulative risk factors. The only difference between the participants who were merged and those who were not merged with the administrative data was non German citizenship. This difference is likely due to those women who

Table 3: Selective Attrition between TG and CG Demographic Characteristics - Administrative and Survey Data

		Difference TG/CG for:			
	Merged	At Least One Interview After Birth	Data Available for 12 Months After Birth	Data Available for 24 Months After Birth	Complete data from Birth Until Third Birthday
	(1)	(2)	(3)	(4)	(5)
Demographic Characteristics					
Age in Years	-0.314	-0.0637	0.0411	0.0872	0.313
	(0.329)	(0.364)	(0.393)	(0.445)	(0.578)
Week in Pregnancy	-0.423	-0.623	-0.429	-0.164	0.0986
	(0.433)	(0.466)	(0.495)	(0.548)	(0.665)
Foreign Nationality	-0.0153	-0.0127	-0.0141	-0.0086	-0.0016
	(0.0187)	(0.0219)	(0.0234)	(0.0261)	(0.0353)
Teenage	0.0358	0.0223	0.0173	0.000	0.0185
	(0.0376)	(0.0404)	(0.0425)	(0.0467)	(0.0550)
Mon. HH-Inc. in €	18.24	33.60	5.046	-3.292	31.79
	(43.69)	(48.27)	(48.63)	(54.22)	(67.26)
Debt over 3000 €	0.0259	0.0275	0.0230	0.0319	0.0565
	(0.0294)	(0.0319)	(0.0342)	(0.0381)	(0.0478)
Education Risk	0.0310	0.0213	0.0214	0.0223	0.0505
	(0.0319)	(0.0359)	(0.0387)	(0.0441)	(0.0552)
Income Risk	0.0193	0.00392	0.0117	0.0229	0.0102
	(0.0291)	(0.0327)	(0.0349)	(0.0399)	(0.0506)
Employment Risk	-0.0272	-0.0353	-0.0429	-0.0495	-0.0734
	(0.0279)	(0.0312)	(0.0336)	(0.0384)	(0.0495)
No Partner	0.0163	0.0324	0.0422	0.0351	0.0268
	(0.0346)	(0.0369)	(0.0386)	(0.0435)	(0.0546)
Unmarried	-0.0296	-0.0394	-0.0436	-0.0343	-0.0571
	(0.0221)	(0.0231)	(0.0252)	(0.0285)	(0.0376)
Living with Parents	0.00674	0.0104	-0.00503	-0.0155	-0.0311
	(0.0336)	(0.0365)	(0.0383)	(0.0422)	(0.0508)
Persons in HH	0.0508	0.148	0.0897	0.0316	-0.0784
	(0.126)	(0.136)	(0.136)	(0.148)	(0.181)
Lower Saxony	0.0319	0.0189	0.0346	0.0238	0.00308
	(0.0365)	(0.0395)	(0.0416)	(0.0460)	(0.0570)
Bremen	-0.0234	-0.00335	-0.0178	-0.00195	0.0247
	(0.0345)	(0.0377)	(0.0399)	(0.0447)	(0.0552)
Saxony	-0.00851	-0.0155	-0.0167	-0.0219	-0.0278
	(0.0356)	(0.0383)	(0.0406)	(0.0451)	(0.0523)
	703	602	539	438	296

Notes: Column (1) contains estimates of the average difference in characteristics between mothers in the control and treatment groups for the participants merged with the administrative data. Dependent variables shown in the first column. Column (2)-(5) contain these estimates for the survey data. The treatment indicator has the value one if the mother is in the treatment group. Robust standard errors are shown in parentheses. See Appendix Tables A3 and A4 for variable definitions. TG = Treatment Group; CG = Control Group.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

the face-to-face interview.

recently migrated having fewer years of employment and welfare history, which reduces the probability of identifying them in the social security data.

6 Estimation Strategy

To analyze the effects of the intervention on maternal employment, fertility, childcare use, and partnership stability, I estimate the intent-to-treat (ITT) effects of the *Pro Kind* intervention using the multivariate model in Equation 1:

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

where Y_{ic} denotes an outcome variable for mother i from community c. HV_{ic} is a dummy variable that takes a value of one if the mother belongs to the treatment group. h_{ic} is a vector of demographic and psychological family characteristics at baseline; α_c are community dummies; and ϵ_{ic} is the error term. β_1 measures the difference in outcome Y between the treatment and control groups.

I estimate the extensive and intensive margin of employment and welfare benefits with linear models. The results are not sensitive for estimating non-linear models for the binary outcomes instead and for including or excluding baseline variables. In the estimations with the administrative data, the only available baseline characteristics are maternal age and community of residence at baseline, whereas in the survey data several baseline characteristics can be included to give more precision to the estimates. I cannot estimate the effect of treatment on the treated using the randomly assigned treatment intended as an instrumental variable for treatment received because the data on compliance with the intervention are not merged with the administrative data. However, the effect of treatment on the treated would be marginally different from the present results because the implementation research showed that 97.7% of the treatment group participants received at least some treatment.

7 Results

7.1 Administrative Data

Table 4 examines the effects of *Pro Kind* on employment, public assistance and household composition using administrative data from the German social security system. In the first row, Column 1 presents the percentage of mothers who were employed for

at least one month in the first three years after the birth of the treatment child.[14] Any employment summarizes the employment types, part-/full-time employment, apprenticeship, and marginal employment.[15] Among the mothers in the control group, 51 percent participated in the labor market in the first 36 months after birth. They were employed for 6.1 months on average during this period, indicating a high amount of job fluctuation and short employment periods in the sample. Analyzing the treatment impact on employment reveals that home visiting reduced the percentage of mothers with any employment (extensive margin) and the number of months employed (intensive margin). These effects are large and significant. The treatment reduced the rate of mothers who were employed for at least one month by 8.7 percentage points, to a rate of 42 percentage points; the average number of months employed was reduced by 1.51 months to 4.59 months, which is a 24.7 percent decrease relative to the mean time worked by the mothers in the control group. Appendix A8 demonstrates the results for different types of employment with the effect being strongest for part-time/full-time employment, for which the treatment reduced the extensive margin by 26.4 percent and the intensive margin by 39.1 percent relative to the mean of the control group.

The second row, "Welfare", indicates whether and for how many months on average a mother lived in a household that received public assistance. The figure in Column 1 shows, corresponding to the affiliation criteria, that 94.1 percent of the mothers in the control group received public assistance for at least one month during the first 36 months after birth. Moreover, the total number of months (28.39) indicates that the participants' households received welfare in 78.8 percent of the months during this period. In line with the reduction in employment, the treatment significantly increased the share of participant households on welfare and the number of months on welfare.

Next, I turn to the outcome of fertility. "Second Child in HH" is a binary variable that takes a value of 1 if two or more children are living in the household and 0 if

[14] If randomization (on average three months before birth) is chosen as starting point of the analysis, the employment and welfare levels change slightly, but the effect sizes of the results do not change.

[15] In Germany, an apprenticeship includes on-the-job training in a company and attendance of a vocational school. Completing an apprenticeship, which usually takes three years, is strongly correlated with labor market success in Germany. Marginal employment is, according to German social security law, an employment relationship with a low absolute level of earnings (a wage of less than 450 Euros per month) or an employment relationship of short duration.

Table 4: Maternal Life Course Outcomes before Randomization and 36 Months after the Birth of the Treatment Child - Administrative Data

	Extensive Margin			Intensive Margin (in Months)		
	Control Mean	Diff. TG/CG	p-values	Control Mean	Diff. TG/CG	p-values
	(1)	(2)	(3)	(4)	(5)	(6)
0 - 36 Months after Birth						
Any Employment	0.507	-0.087**	0.020	6.102	-1.509**	0.017
	[0.501]	(0.038)		[8.825]	(0.630)	
Welfare	0.941	0.037**	0.014	28.39	1.634**	0.046
	[0.235]	(0.015)		[11.51]	(0.817)	
Observations	*341*	*703*		*341*	*703*	
Second Child in HH	0.183	0.066**	0.037			
	[0.363]	(0.032)				
Observations	*323*	*677*				
0 - 36 Months before Randomization						
Any Employment	0.592	0.015	0.678	9.384	0.201	0.816
	[0.492]	(0.037)		[11.60]	(0.864)	
Welfare	0.788	-0.007	0.819	14.531	-0.371	0.706
	[0.408]	(0.031)		[13.10]	(0.981)	
Observations	*341*	*703*		*341*	*703*	

Notes: Standard deviations in square brackets; robust standard errors in parentheses. Columns (2) and (5) report the coefficient and standard error on Home Visiting (HV) from estimating equation (1) by OLS. Data is available on a monthly base. TG = Treatment Group; CG = Control Group; HH = Household.
* p < 0.1, ** p < 0.05, *** p < 0.01

one child or no children are living in the household.[16] Because the data recorded household composition only for households that receive welfare benefits or were engaged in job seeking, the number of observations is slightly reduced. The results show that while 18.3 percent of control group participants lived in a household with two or more children within the 36 months after the birth of the treatment child, this rate is 6.6 percentage points higher in the treatment group, leading to 24.8 percent of households with more than one child, which is an increase of 36 percent relative to the mean of the control group. As a robustness check, I assigned no second child to mothers with missing fertility data, which slightly increases the treatment coefficient on fertility. All effects on employment, welfare and fertility also hold and become slightly larger if the models include the community and age of the mother as controls (Appendix A8).

[16]There could be no children in the household in the event of miscarriage of the first pregnancy or when the treatment child did not live with the mother, most likely because the child was given to foster parents or, lived with the grand parents or the father. This was the case in 11 households. As a robustness check, I excluded these women from the fertility and employment and welfare analyses without any changes showing that these women were equally balanced between treatment and control groups.

The last rows in Table 4 show the differences between treatment and control groups for employment and welfare receipt within 36 months before randomization. Differences in both outcomes are small and far away from being statistical significant which is also true if shorter time periods before randomization are chosen. These findings further enforce that the differences in outcomes after randomization can be solely attributed to the intervention.

Figure 1: Employment, Welfare, and Fertility by Age of the Treatment Child - Administrative Data

Note: The figure displays the average rate of second birth, employment and welfare receipt for participants in treatment group (TG) and control group (CG) from randomization until 36 months after birth. The solid vertical line shows the month of randomization. The dashed vertical line shows the month of birth of the treatment child.

To examine the dynamics in the results, Figure 1 displays the development of employment, welfare and fertility over time for treatment and control groups within 36 months after birth (39 months after randomization). While the effect on fertility is small until 18 months after birth, mothers in the treatment group have already a lower employment and a higher welfare rate than their counterparts in the control group. Starting with 18 months, the treatment effect on births emerges. The fertility increase affects the difference in employment only slightly whereas the gap in welfare receipt increases after 18 months. These findings suggest that the higher fertility only partly

explains the decreasing treatment effect on employment because employment was already lower in the treatment group before fertility increased. Only in the third year increased fertility seems to have a slight employment reducing and welfare increasing effect. Appendix Figure 1 presents dynamics in employment and welfare also before randomization, confirming the estimation results of Table 4 with no differences before randomization and strong divergences in the outcomes after randomization.

Appendix Table A9 presents heterogeneous effects for teenage vs. non-teenage mothers and municipalities where only family midwives conducted the home visits vs. municipalities where also social pedagogues were involved in the home visits. Additionally, I analyze whether the effect on fertility differs between mothers who were employed and non-employed after birth. The differences in treatment effects between teenagers and older participants are only small and not significant. Furthermore, home visitors' profession does not have an significant effect on outcomes, although the effect on second births is slightly smaller in municipalities were only family midwifes conducted the home visits. However, the analysis regarding employment and fertility shows that the fertility effect is concentrated in mothers who had not worked in the first 24 months after birth which confirms that the lower employment in the treatment group in the first 24 months is not explained by mother stopping employment after a subsequent birth.

Overall, the results from the administrative data indicate that the intervention had unintended effects, which are in contrast to the results of studies from the U.S. Instead of the intended higher levels of maternal employment and economic self-sufficiency and a lower rate of second births, we observe the opposite. The reduction in employment and the increase in welfare dependency were likely observed because the mother decided to stay at home longer directly after birth, and then these mothers who stayed at home decided to have another child. Only in the third year after birth, parity seems to increase welfare receipt and to slightly reduce employment in the treatment group. In the next section, I use survey data to examine which channels most likely explain the identified results.

7.2 Survey Data

Table 5 presents the results of the telephone survey for the first three years after birth including the 296 mothers who participated in all interviews until the third birthday of the treatment child.[17] The first three rows of Table 5 include the same outcomes as Table 4. The only difference is that the variable "Second Child in Household" is labeled "Second Birth" because the survey directly asked for second births and not only for household composition. To increase the comparison between the survey sample and the administrative sample and to account for the potential bias that non-response may introduce, I weighted the models with the predicted probabilities of participating in all interviews. The weights are calculated by a logit model using the baseline characteristics presented in Appendix 3 and 4.

In the survey data, the extensive and intensive rate of employment are slightly higher than in the administrative data. However, the differences in employment between the treatment and control groups are similar in size without being statistically significant because of larger standard errors. In line with the reported higher employment, fewer mothers in the control group reported receiving welfare than indicated by the administrative data. However, also in this category, the treatment effect corresponds in size and significance to that in the administrative data. Analyzing fertility in the survey data shows that the rate of second births in the control group is comparable to the respective figure in the administrative data. The difference between the treatment and control groups in the survey data is 8.6 percentage points, which is even higher than in the administrative data.

The last six rows in Table 5 contain information which was measured only through the telephone surveys, including the occurrence of a second pregnancy, inconsistent use of contraceptives, constant partnership, change in marriage status, school attendance, and childcare use. These six outcomes can help to identify channels why the intervention had unintended effects on employment, welfare use and fertility, which we observed in the administrative data. School attendance and childcare use were recorded on a monthly basis while the status at the time of the interview was recorded

[17] I include only mothers who participated in all interviews to ensure that the outcomes can be interpreted in the same way as the outcomes from the administrative data.

Table 5: Maternal Life Course Outcomes 36 Months after Birth of the Treatment Child - Survey Data

	Extensive Margin			Intensive Margin (in Months)		
	Control Mean	Diff. TG/CG	p-values	Control Mean	Diff. TG/CG	p-values
	(1)	(2)	(3)	(4)	(5)	(6)
Any Employment	0.600	-0.061	0.347	8.312	-1.415	0.246
	[0.491]	(0.065)		[9.322]	(1.217)	
Welfare	0.919	0.052*	0.071	26.093	2.217*	0.089
	[0.273]	(0.029)		[10.852]	(1.301)	
Second Birth	0.203	0.086	0.147			
	[0.403]	(0.059)				
Second Pregnancy	0.360	0.011	0.863			
	[0.481]	(0.065)				
Inconsistent Use of Contraceptives	0.205	0.047	0.405			
	[0.404]	(0.056)				
Constant Partnership	0.401	-0.005	0.927			
	[0.491]	(0.057)				
Change in Marriage Status	0.163	-0.012	0.794			
	[0.327]	(0.045)				
School	0.120	-0.016	0.714	1.081	0.468	0.537
	[0.326]	(0.044)		[4.087]	(0.758)	
Childcare Utilization	0.589	0.061	0.349	7.268	1.953	0.125
	[0.493]	(0.065)		[8.869]	(1.269)	
Observations	137	296		137	296	

Notes: Robust standard errors in parentheses; Standard deviations in square brackets. Columns (2) and (5) report the coefficient and standard error on home visiting (HV) from estimating equation (1) by OLS. The estimations are weighted by the inverse probability to participate in all interviews. TG = Treatment Group; CG = Control Group; HH=Household.

for the other outcomes.[18]

Analyzing the rate of second pregnancies reveals that, in contrast to the rate of second births, it does not differ between the treatment and control groups. In both groups, approximately one-third of the mothers became pregnant a second time within 36 months after the birth of the treatment child. This finding indicates that a difference in pregnancy outcomes must be present, at least to some extent. As expected, since the home visits did not affect second pregnancies, they also did not affect inconsistent use of contraceptives.[19]

The next four rows examine partner stability, marriage status, school attendance and childcare use. Partner stability represents the percentage of women who stayed with the same partner from pregnancy until the third birthday of the treatment

[18]Childcare utilization is a broad measure of whether a child attends childcare. It does not include hours or quality of childcare.

[19]The question regarding the use of contraceptives was asked in three interviews at 15, 27 and 36 months. A mother was considered to use contraceptives inconsistently if she stated in one interview that she did not always use a contraceptive method. Mothers who were sexually inactive, pregnant or trying to become pregnant were excluded from the sample.

child. Change in marriage status indicates whether a women became married during the observation period. The treatment did not change the rate of mothers in a stable partnership or the marriage rate, indicating that it is not a more stable family situation which lead to more births or that a higher family income from a partner decreases maternal employment probability. School attendance is an indicator that could explain the lower employment rate in the treatment group due to mothers returning to or starting school after birth. Increased school attendance would be in line with the goals of the intervention. However, the survey data reveal no increase in school attendance for the mothers in the treatment group.[20]

The home visitors may recommend less or later childcare use to the mothers which may explain the lower rate of employment in the treatment group. However, following the *Pro Kind* guidelines, the home visitors supported the mothers in the childcare application process and in finding adequate childcare if the mothers wanted to use childcare. Additionally, the home visitors might gave advice that childcare is subsidized or even completely free for mothers on welfare in Germany. The results show that the intervention slightly increased the average months of childcare use, suggesting that home visitors were successful at supporting mothers in locating care. An analysis of the timing of childcare use reveals that the two groups used childcare similarly in the first 12 months; only afterwards the usage increased stronger for the treatment group. Therefore, lower childcare use does not explain the lower employment after birth in the treatment group. Instead, it seems that some mothers, although not working, used institutional childcare. If these mothers perceived external childcare as a relief of strain, the better provision might be one reason why they decided to have a second child.

Overall, the results of the survey data confirm the findings from the administrative data that the intervention increased second births and welfare dependency and decreased employment. Appendix A10 shows the unweighted results, which appear similar in size with the exception that the coefficient for second birth becomes significant. Investigating the channels for the results indicates that a change in second pregnancy outcomes most likely explains the increase in second births, while partner

[20]Enrollment in higher education was of negligible relevance in the treatment and control groups.

stability and school attendance are unlikely to be explanations. This finding is again in contrast to results from the U.S., where the intervention reduced not only second births, but also second pregnancies in all three trials.

The analyses included only mothers who participated in all the interviews with data available for the complete 36 month after birth. To use all available data, in the next section I remove the sample restriction and include all mothers who participated in at least one interview after birth to examine how pregnancy outcomes, as the main driver of the fertility effect, differ between the treatment and control groups.

7.2.1 Pregnancy Outcomes

Table 6, Panel A shows that the rate of second pregnancies in this sample of control group mothers is slightly lower than in the sample that includes only mothers who participated in all the interviews. Presumably, the rate is lower because some mothers participated in only one interview after birth, which was most likely before a further pregnancy occurred. The rate of second pregnancies in the treatment group is 5.5 percentage points higher, but the difference is not statically significant at the ten percent level, thereby confirming the results from the analyses of the mothers who participated in all the interviews. Altogether, 175 second pregnancies occurred among the mothers who participated at least in one interview after birth.[21]

Panel B presents the outcomes of these 175 second pregnancies, which could be live birth, abortion, miscarriage or unobserved pregnancy outcome. Along with the results of the previous sections, Panel B reveals that the percentage of pregnancies that led to a live birth was higher in the treatment group (63%) than in the control group (53%), resulting in 103 observed second births. Additionally, the table demonstrates that abortions (24% vs. 15%) and miscarriages (14% vs. 9%) were more common in the control group than in the treatment group. In contrast, the percentage of pregnant women with unobserved pregnancy outcomes was only slightly higher in the treatment group.

Panel C uses a multinomial logit function to examine the differences in pregnancy outcomes in greater detail. I am interested in whether the treatment influenced the

[21]The 175 pregnancies only include the first pregnancy of each participant after the birth of the treatment child.

Table 6: Second Pregnancy Outcomes in Treatment and Control Groups

Panel A: Second Pregnancy Occurred			
	Control Mean	Diff. TG/ CG	p-value
Pregnancy after First Birth	0.261	0.055	0.136
	[0.440]	(0.037)	
Obs.	*283*	*602*	
Panel B: Second Pregnancy Outcomes (Descriptives)			
	Control Mean	Treatment Mean	Overall Mean
Live Birth	0.527	0.634	0.589
Abortion	0.243	0.149	0.189
Miscarriage	0.135	0.089	0.109
Unobserved	0.095	0.129	0.114
Obs.	*74*	*101*	*175*
Panel C: Multinomial Logit			
	Birth vs. Abortion	Birth vs. Miscarriage	Birth vs. Unobserved
Home Visiting	-0.677*	-0.600	0.123
	(0.405)	(0.503)	(0.512)
Obs.	*175*	*175*	*175*

Notes: Standard errors in parentheses; Standard deviations in square brackets. The table includes all mothers with at least one interview after birth. Panel B includes all pregnancies from Panel A. Panel C is a multinomial logit estimation with Live Birth as baseline category. TG = Treatment Group; CG = Control Group;
* p < 0.1, ** p < 0.05, *** p < 0.01

probability of a live birth relative to the other three outcomes. The analysis reveals that the probability of a pregnancy ending in an abortion instead of a live birth was significantly lower in the treatment group than in the control group. For miscarriage, the coefficient is in the same direction and of approximately the same size but is not significant. Finally, the probability of not observing the outcome of the pregnancy relative to that of a live birth was only slightly higher in the treatment group. These findings confirm that the differences in fertility between the two groups were not caused by selective attrition; rather, they were the result of a reduced number of abortions and miscarriages in the treatment group.

One concern about the validity of the abortion result is the possibility of misreporting. The validity is violated if mothers in the treatment group reported abortions or pregnancies differently than mothers in the control group. This might be the case if mothers in the treatment group did not want the home visitor to know about an unwanted pregnancy; therefore, being in the treatment group might increase a mother's likelihood to underreport pregnancies and, consequently, abortions and/or miscarriages. However, the interviewers guaranteed the participants that their answers would not be relayed to the home visitors. Additionally, it is very unlikely

that the home visitors would not recognize a client's pregnancy, possibly making the reporting bias smaller in the treatment group than in the control group. Another concern is that abortions might be reported as miscarriages as a socially more acceptable termination. However, this misreporting would imply that the intervention's effects on abortions are only a lower bound, since miscarriages were higher in the control group.

Placing the rate of abortions in the *Pro Kind* program in relation to the abortion rates in the overall population helps interpret the abortion results. From 2008 to 2011, there were approximately 16 abortions per 100 live births in the overall German population.[22] Ratios for at-risk mothers who are comparable to the *Pro Kind* sample are not available. However, data for unmarried women, who might be more similar to the *Pro Kind* sample than the overall population, indicate 27 abortions per 100 live births (Statistisches Bundesamt, 2014). The control group of the *Pro Kind* sample had a ratio of 46 abortions per 100 live births, whereas the ratio in the treatment group was 23 to 100, which is close to the population average and lower than the average for unmarried mothers. This might indicate that mothers in the treatment group were as confident in their ability to raise a second child as average mothers.

Despite the finding that a lower percentage of pregnancies ended in an abortion in the treatment group than in the control group, it remains unclear whether this is the result of appropriate family planning decisions, which is a goal of the *Pro Kind* program. In this context, appropriate decisions mean that only mothers who plan a second birth and who are able to meet the challenges of another child give birth to a second child. As a first insight, the analysis of the survey data indicated that the treatment did not affect partner stability, which might be related to appropriate family planning. To investigate in greater detail the question of whether appropriate family planning increased, I analyzed the life situations of the *Pro Kind* second-time mothers in treatment and control groups and compared them with SOEP second-time mothers.

Table 7 includes data from the interview after the birth of the second child for 97 of the 103 second children. The first two rows present responses to questions

[22]German official statistics only report the rate of abortions in comparison to live births and not the rate of pregnancies that end in abortion.

Table 7: Life Situations of Mothers who Gave Birth to a Second Child

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	CG		TG		p-value	SOEP	
	n	Mean	n	Mean	Diff. C-T	n	Mean
After Birth of sec. Child							
Unplanned Preg.	35	0.57	62	0.61	0.689	799	0.19
Father Does not Live in HH	35	0.29	60	0.40	0.262	803	0.06
No Other Care Apart From Mother	35	0.31	62	0.48	0.104	804	0.08
Mother has no Partner	33	0.06	58	0.17	0.130	803	0.01
Age of the Sec. Child in Mo.	32	8.41	62	6.49	0.352	802	6.96
Age of the Moth. at Births in Years	35	23.4	62	23.9	0.594	766	32.08

Notes: P-values are based on z-statistic of a two-group test of proportions. The presented data contains all second children for whom data are available. *Age of the Sec. Child in Mo.* gives the age of the second child at the time of the interview. CG = Control Group; TG = Treatment Group.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

concerning whether the child was unplanned and whether the mother had a partner. If the mothers had made appropriate family planning decisions, one would expect that unplanned pregnancies and pregnancies among women without partners would be uncommon among second-time mothers. However, 61% of the mothers in the treatment group stated that their second child was unplanned. In the control group, this rate was 57%. Furthermore, other characteristics, such as "no partner" or "father does not live in the household", occurred more often in the treatment group. Although none of these differences are statistically significant, the results may indicate that mothers in the treatment group with fewer resources had a second child and that these mothers were less accurate with respect to family planning than the mothers in the control group. The findings that only 39% of the mothers planned their second pregnancy and 48% had no care support from another person indicates little appropriate family planning in the treatment group. These figures are even more illustrative when compared to population representative SOEP mothers (Columns 6 and 7): 81% of the SOEP mothers stated that the pregnancy that led to a second child was planned, and only 8% stated that they were alone responsible for the child.

The next question is why the mothers in the treatment group decided to have another child despite not having planned a second birth, being unemployed and seeming to be unable to meet the challenges of having another child. As an explanation, the home visitor might have directly influenced the decision of the pregnant mother. There are no recommendations concerning abortions in the *Pro Kind* or NFP guidelines, and I do not have information about the behavior of the home visitors in this

situation. Although, the nurse supervisors stated in in-depth interviews that abortion was essentially not a topic in the nurse supervision, they also stated that a nurse or midwife would hardly advise a client to abort a pregnancy. However, two reasons argue against direct advice of the home visitor as the main explanation for less abortions. First, mothers in the treatment and control groups also received encouragement to keep the baby from other sources because German law permits abortions only if the woman has received consultation from a family counseling office. Second, many of the mothers decided against an abortion after the second birthday of the treatment child, when the intervention had already ended. Therefore, in addition to the direct advice of the home visitors and the family counseling office, it is likely that other channels were also important for the decision of the mothers in the treatment group.

7.2.2 Life Satisfaction and Well-being

One potential channel why the *Pro Kind* program increased both fertility and the duration that mothers stayed at home directly after birth is increased satisfaction with their lives and their maternal role. As Sierau et al. (2015) show in previous work on *Pro Kind*, the intervention significantly increased parental self-efficacy and maternal feelings of attachment. These more positive experiences to the first child might have lead to higher life satisfaction in the treatment group. To test more directly whether the intervention increased maternal life satisfaction, the section investigates whether the *Pro Kind* intervention influenced reports of maternal well-being, life satisfaction, and perceived stress. These outcomes were obtained at a face-to-face interview 12 months after the birth of the treatment child and in a telephone interview on average 28 month after birth of the treatment child.

Appendix Table A11 shows the life satisfaction, well-being and stress questions and their means for treatment and control groups. At 12 months, mothers in the treatment group give lower ratings in all eleven stress items and higher ratings in eight of ten satisfaction items. At 28 months, in eight of the nine satisfaction dimensions, the mothers in the treatment group reported being more satisfied than the mothers in the control group. The results are similar for the four questions regarding well-being.

Table 8: Well-Being, Stress and Satisfaction with Life

	(1)	(2)	(3)	(4)	(5)	(6)
	Face-to-Face Interview at 12 Months					
	Stress Index		Life Satisfaction Index			
Home Visiting	-0.117**	-0.125***	0.108**	0.124**		
	(0.048)	(0.045)	(0.054)	(0.054)		
Household Controls	No	Yes	No	Yes		
Observations	*422*	*422*	*422*	*422*		
	Telephone Interview at 28 Months					
	Index of Well-Being		Life Satisfaction Index in Different Areas		Satisfaction with Life in General	
Home Visiting	0.189***	0.167***	0.118*	0.106*	0.155*	0.147**
	(0.069)	(0.043)	(0.061)	(0.051)	(0.097)	(0.062)
Household Controls	No	Yes	No	Yes	No	Yes
Observations	*434*	*429*	*430*	*425*	*432*	*427*

Notes: Standard errors in parentheses. *Stress* is an index of ten questions asking for perceived stress in different areas. *Life Satisfaction* is an index of ten questions concerning satisfaction in different domains and general life satisfaction. *Well-Being* is an index of being less often sad, angry, or worried and more often happy. *Life Satisfaction in Different Areas* is an index of eight questions concerning satisfaction with health, housework, household income, personal income, place of dwelling, free time, childcare availability and family life. The indices are constructed following Kling et al. (2007). Controls include demographic and psychological baseline variables (see Appendix Tables A3 and A4), community fixed effects and age of the treatment child. Measurement occurred in average at 12 months and 28 months after the birth of the treatment child. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The mothers in the treatment group reported feeling sad, angry, or worried less often and happy more often.

Table 8 shows that at both points of time the life satisfaction indecies are significantly higher in the treatment group. The index of stress is significantly lower at 12 months and the well-being index is significantly higher at 28 months in the treatment group.[23] The standardized effect sizes are meaningful, with values higher than 0.10 SD. The effect on life satisfaction and stress occurs already at 12 months after birth, which is before the fertility effect emerges. Therefore, the increased fertility in the treatment group did not cause the higher satisfaction and lower perceived stress in this group supporting that higher life-satisfaction induced the effect on fertility and the longer absence from the workforce. On the other hand, the effect on life satisfaction remains stable at 28 months when many mother have already given births to a second child. The finding of Sandner et al. (2017) who show that mothers in the treatment group used less anti-depressants and reported less symptoms of clinical depression further support that the *Pro Kind* intervention increased maternal well-being. Overall, the findings indicate that mothers who received guidance and encouragement by

[23]The *life satisfaction index* is based on ten questions related to satisfaction in different domains and general life satisfaction. The *stress index* is based on eleven questions asking for perceived stress in different areas. The *well-being index* is based on four questions and the *life satisfaction index* in different areas is based on eight questions. Following Kling et al. (2007), indices are defined to be the equally weighted average of z-scores of its components. The z-scores are calculated by subtracting the control group mean and dividing by the control group standard deviation.

home visitors enjoyed their maternal life more which may influenced their abortion and employment decisions.

8 Comparison to U.S. Results

In this section, I first discuss whether program implementation and participants' characteristics can explain why the effects on maternal outcomes differ so substantially from the results of the U.S. studies, mainly the NFP. Then I discuss different explanations for the reversed effects, such as welfare state arrangements for mothers with small children and contraception use.

Table 9: Program and Participant Characteristics of Pro Kind and NFP

	Nurse Family Partnership			Pro Kind
Characteristics of the Intervention				
Location	Elmira, NY	Memphis, TE	Denver, CO	Germany
Year	1980	1990	1995	2007
Evaluation Design	RCT	RCT	RCT	RCT
Randomized Participants	264	1139	490	755
Home Visits (Mean)	32	33	27.5	33
Materials	NFP Guidebooks			German Adaptation of NFP Guidebooks
Home Visitors (Qualification)	Family Nurses			Family Midwifes, Social Pedagogues
Home Visitors (Training)	NFP Guidelines			NFP Guidelines
Participants Characteristics				
Parity	First			First
Date of Randomization	Pregnancy			Pregnancy
Socioeconomic Status	Low			Low
Age	18.9	18.1	19.9	21.4
Unmarried in %	100	98	85	92
Years of Education	10.7	10.2	11.2	10.7
Results Maternal Life Course				
Employment	+	+	+	-
Second Birth	-	-	-	+

Notes: + indicates a positive treatment effect, - indicates a negative treatment effect.

Table 9 summarizes the *Pro Kind* implementation and participants' characteristics. The table shows that, compared with the NFP home visitors, the *Pro Kind* home visitors received the same training and used the same materials and guidebooks (translated into German), during their home visits. The only difference occurs in the profession of the home visitors. One could be concerned that family midwifes are more pro-natalist than family nurses. However, in Germany family midwifes obtain training far beyond delivery and postpartum care, resulting in a comparable quali-

fication as family nurses in the U.S. Additionally, the results on fertility show that the effects are not larger in municipalities where only family midwifes conducted the home visits (Appendix Table A9). Another similarity between the *Pro Kind* program and in the NFP trials is the average number of conducted home visits. Consequently, program costs are very similar in the two programs. As discussed above, the average intervention cost in the NFP Memphis trial was $11,511 (expressed in 2006 U.S. dollars). The average cost of the *Pro Kind* intervention was €8,705 (expressed in 2008 Euros), or approximately $11,752 assuming an exchange rate of 1.35 €/$ (Maier-Pfeiffer et al., 2013). Finally, the implementation data show that the home visitors spent a similar amount of time on the various program topics in *Pro Kind* and NFP (Appendix Table A5). The next rows in Table 10 compare sample characteristics between *Pro Kind* and NFP. With respect to marriage status and years of education, the populations in the NFP randomized trials show similar characteristics. Only the average age of the participants appears slightly younger. However, it is important to note that in both *Pro Kind* and NFP, all participants were disadvantaged, pregnant, first-time mothers. These criteria alone should result in highly comparable populations in the U.S. and German studies. Taking all these aspects together, it is unlikely that differences in implementation or the participants' characteristics alone can explain the dramatic difference between the results from the U.S. studies and the German study.

One alternative explanation might be the different welfare state arrangements in Germany and the U.S. for families with children under three. In Germany, social assistance is means-tested and increases with parity. There are no work obligations or benefit cuts until the child's third birthday. If a welfare-dependent mother decides to work, the benefits are withdrawn at a rate of almost 100%. As an example of these low incentives, Blundell et al. (2009) showed that the budget line for a low-wage single mother with two children was hardly affected by her working hours.[24] In contrast, in the U.S., welfare programs include work obligations, in-kind transfers and

[24] A single mother with one child receives approximately €1,370 in welfare payments per month (which is $1850, assuming an exchange rate of 1.35 €/$). If she earns the German hourly minimum wage of €8.5, she earns €1,200 with full-time employment after deductions for health insurance (childcare is generally free for low-income mothers in Germany). Thus, the single mother must work full-time and must earn an hourly wage of €9.5 to meet the reservation wage.

family caps, limiting either partially or completely any additional benefit for having a subsequent child while receiving welfare benefits. In addition, Meyer and Rosenbaum (2001) or Hoynes and Patel (2015) demonstrated that since 1984, changes in tax and transfer programs, such as the EITC, sharply increased the incentive for low-income mothers and single mothers to work. These welfare arrangements can rationalize the differences between the *Pro Kind* and the NFP findings, assuming that home visiting in the U.S. and Germany increases home production skills and labour market skills of the mothers. In Germany, employment does only provide little additional utility for the mother because welfare benefits are high. Therefore, the intervention effect on maternal skills may have a higher marginal benefit and dominate over the effect on labour market skills. In contrast, in the U.S., employment leads to a great utility increase for the mothers. Therefore, the increase in labour market skills may dominate over the effect on maternal skills and lead to higher employment.

Another reason for the different fertility effects apart from welfare, might be different levels of knowledge concerning contraception among young women in the U.S. and Germany. Although institutional settings for abortions are comparable in both countries and all contraceptives are generally available for purchase in Germany and the U.S.,[25] it has been documented that teen pregnancies are higher in the U.S. than in Germany (Kearney and Levine, 2012). One explanation for the higher rate in the U.S. is less contraceptive use and knowledge of contraceptives among young women (Darrach et al., 2001). If this is the case, home visiting in the U.S. may have more space to achieve an impact, e.g., due to recommendation of safer contraception methods, whereas it is more difficult to reach additional benefits in this topic in Germany. In line with this argument, NFP reduces not only births but also pregnancies. While it is difficult to reject this explanation, two findings challenge it. First, assuming that knowledge of contraception is lower among teenagers, I would expect to find a significant smaller effect of the *Pro Kind* intervention on teenage fertility, which is not the case. Second, Kearney and Levine (2015b) showed that mandatory sex education

[25] German law permits abortions up to the 12th week of a pregnancy if the woman received consultation and passed a subsequent waiting period of three days. After the 12th week of the pregnancy, abortions are possible without time limits if there is a risk to the life and health of the mother (medical indication) or if the pregnancy is the result of a crime (criminal indication). The expenses for abortions based on the two indications are typically borne by health insurances, whereas abortions following a consultation are paid privately. Although abortion laws are more lenient in the U.S. relative to Germany, abortion is legal in both countries; therefore, a comparable situation persists (Levine, 2004; Cygan-Rehm and Riphahn, 2014).

has only limited effects on teen births and that lack of knowledge seems not to be the main driver of teen fertility. Therefore, it remains an open question to what extent education in contraception by home visitors or increased maternal personal strengths reduced further pregnancies in the U.S. studies.

9 Conclusion

Home visiting programs are a popular type of early childhood intervention for supporting disadvantaged families. While many studies have investigated how these programs affect child outcomes, this study used a randomized experiment to answer the much less thoroughly investigated question of how home visiting programs affect the maternal life course. The few previous studies that investigated this topic found that home visiting programs had positive effects on maternal employment and reductions in fertility. In contrast, this analysis of the *Pro Kind* program reveals that the intervention had negative effects on employment and positive effects on fertility. The effects on fertility were mainly driven by the lower number of abortions in the treatment group. Furthermore, the *Pro Kind* program increased the life satisfaction and well-being of the participating mothers. Although it is not clear how the observed increase in fertility and decrease in employment will affect child development and government spending, it is encouraging that the relationship between home visitors and mothers seem sufficiently strong to affect two very fundamental life decisions of disadvantaged mothers. This finding shows how promising the concept of home visiting is in general.

A randomized experiment was used to evaluate the effects of *Pro Kind* on the maternal life course. Therefore, the effects can be causally linked to the intervention. For the main analysis, I used administrative data that are not subject to the risk of missing data or reporting error. For the analysis of the channels that lead to the unintended outcomes, I relied on survey data that suffered from survey non-response. Nevertheless, a comparison of the baseline characteristics of the treatment and control groups indicates that this attrition was not selective. Therefore, it is unlikely that the sample attrition reduced the validity of the results.

Previous studies that examined the effect of home visiting on the maternal life course were conducted in the U.S., whereas the *Pro Kind* program is located in Germany. The content and implementation of the program and the program participants are very similar in *Pro Kind* and the U.S. studies. Therefore, the differences in the two countries' welfare systems might explain much of the variation in outcomes between the previous studies and the *Pro Kind* study. Studies of other early childhood programs, particularly when they are implemented in settings other than the U.S., should consider the results of the *Pro Kind* program. One example in which an evaluation of a home visiting program did not consider maternal employment, welfare receipt or fertility as outcomes is a recent large RCT in the U.K., which also evaluated an adaptation of NFP (Robling et al., 2016). That study considered second pregnancies as the only maternal life course outcome. Consistent with the *Pro Kind* results, the findings showed no differences between the treatment and control groups in second pregnancies which suggests that the results for other life course outcomes might be also similar to the *Pro Kind* findings.

References

Anant, E., A. Gassman-Pines, and C. Gibson-Davis (2013). Community-Wide Job Loss and Teenage Fertility: Evidence From North Carolina. *Demography 50*(6), 2151–2171.

Anger, S., J. R. Frick, J. Goebel, M. M. Grabka, O. Groh-Samberg, H. Haas, E. Holst, P. Krause, M. Kroh, H. Lohmann, J. Schupp, I. Sieber, T. Siedler, C. Schmitt, C. K. Spiess, I. Tucci, and G. G. Wagner (2009). Developing SOEPsurvey and SOEPservice - The (Near) Future of the German Socio-Economic Panel Study (SOEP). *SOEPpapers 155*, 1–35.

Bauernschuster, S., T. Hener, and H. Rainer (2016). Children of a (Policy) Revolution: The Introduction of Universal Child Careand its Effect on Fertility. *Journal of the European Economic Association 14*(4), 975–1005.

Berger, E. M. and C. K. Spiess (2011). Maternal Life Satisfaction and Child Outcomes: Are They Related? *Journal of Economic Psychology 32*(1), 142–158.

Bernal, R. and M. P. Keane (2011). Child Care Choices and Children's Cognitive Achievement: The Case of Single Mothers. *Journal of Labor Economics 29*(3), 459–512.

Black, S., P. J. Devereux, and K. G. Salvanes (2010). Small Family, Smart Family? Family Size and the IQ Scores of Young Men. *Journal of Human Resources 45*(1), 33–58.

Blundell, R., M. Brewer, P. Haan, and A. Shephard (2009). Optimal Income Taxation of Lone Mothers: An Empirical Comparison of the UK and Germany. *The Economic Journal 119*(535), F101–F121.

Brand, T. and T. Jungmann (2011). *Final report about the implementation of the pilot project Pro Kind*. Hannover.

Brand, T. and T. Jungmann (2012). Implementation Differences of Two Staffing Models in the German Home Visiting Program Pro Kind. *Journal of Community Psychology 40*(8), 891–905.

Brand, T. and T. Jungmann (2014). Participant Characteristics and Process Variables Predict Attrition from a Home-Based Early Intervention Program. *Early Childhood Research Quarterly 29*(2), 155–167.

Brooks-Gunn, J., M. McCormick, S. Shapiro, A. A. Benasich, and G. Black (1994). The Effects of Early Education Intervention on Maternal Employment, Public Assistance, and Health Insurance: The Infant Health and Development Program. *American Journal of Public Health* (84), 924–931.

Buckles, K. and E. Munnich (2012). Birth Spacing and Sibling Outcomes. *Journal of Human Resources 47*(3), 613–643.

Card, D., J. Heining, and P. Kline (2013). Workplace Heterogeneity and the Rise of West German Wage Inequality. *Quarterly Journal of Economics 128*(3), 967–1015.

Carneiro, P., K. V. Løken, and K. G. Salvanes (2015). A Flying Start? Maternity Leave Benefits and Long-Run Outcomes of Children. *Journal of Political Economy 123*(2).

Cygan-Rehm, K. and R. T. Riphahn (2014). Teenage pregnancies and births in Germany: patterns and developments. *Applied Economics 46*(28), 3503–3522.

Dahl, G., A. R. Kostol, and M. Mogstad (2014). Family Welfare Cultures. *Quarterly Journal of Economics 129*(4), 1711–1752.

Darrach, J., S. Singh, and J. Frost (2001). Differences in Teenage Pregnancy Rates Among Five Developed Countries: The Roles of Sexual Activity and Contraceptive Use. *Family Planning Perspectives 6*(33), 244–250.

Dustmann, C., J. Ludsteck, and U. Schönberg (2009). Revisiting the German Wage Structure. *Quarterly Journal of Economics 124*(2), 843–881.

German Federal Employment Agency (2014). Zweites Buch Sozialgesetzbuch – SGB II Fachliche Hinweise, § 10 SGB II.

González, L. (2013). The Effect of a Universal Child Benefit on Conceptions, Abortions, and Early Maternal Labor Supply. *American Economic Journal: Economic Policy 5*(3), 160–188.

Grogger, J. and S. G. Bronars (2001). The Effect of Welfare Payments on the Marriage and Fertility Behavior of Unwed Mothers: Results from a Twins Experiment. *Journal of Political Economy 109*(3), 529–545.

Hanushek, E. (1992). The Trade-off Between Quantity and Quality. *Journal of Political Economy 100*(1), 84–117.

Heckman, J., M. Holland, K. Makino, R. Pinto, and M. Rosales-Rueda (2017). An Analysis of the Memphis Nurse-Family Partnership Program. *NBER Working Paper 23610*.

Heckman, J., S. H. Moon, R. Pinto, P. Savelyev, and A. Yavitz (2010). Analyzing Social Experiments as Implemented: A Reexamination of the Evidence From the HighScope Perry Preschool Program. *Quantitative Economics 1*(1), 1–46.

Heckman, J., R. Pinto, and P. Savelyev (2013). Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes. *American Economic Review 103*(6), 2052–2086.

Hoynes, H. and A. Patel (2015). Effective Policy for Reducing Inequality? The Earned Income Tax Credit and the Distribution of Income. *NBER Working Paper 21340*.

Jungmann, T., Y. Ziert, V. Kurtz, and T. Brand (2009). Preventing Adverse Developmental Outcomes and Early Onset Conduct Problems through Prenatal and Infancy Home Visitation: The German Pilot Project Pro Kind. *European Journal of Developmental Science 3*(3), 292–298.

Kearney, M. S. and P. B. Levine (2012). Why is the Teen Birth Rate in the United States So High and Why Does It Matter? *Journal of Economic Perspectives 26*(2), 141–166.

Kearney, M. S. and P. B. Levine (2014). Income Inequality and Early Nonmarital Childbearing. *Journal of Human Resources 49*(1), 1–31.

Kearney, M. S. and P. B. Levine (2015a). Investigating recent trends in the U.S. teen birth rate. *Journal of Health Economics 41*, 15–29.

Kearney, M. S. and P. B. Levine (2015b). Media Influences on Social Outcomes: The Impact of MTV's 16 and Pregnant on Teen Childbearing. *American Economic Review 105*(12), 3597–3632.

Kessler, D. (1991). Birth Order, Family Size, and Achievement: Family Structure and Wage Determination. *Journal of Labor Economics 9*(4), 413–426.

Kitzman, H., D. Olds, C. Henderson, C. Hanks, R. Cole, R. Tantelbaum, and et al (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.

Kling, J. R., J. B. Liebman, and L. F. Katz (2007). Experimental Analysis of Neighborhood Effects. *Econometrica 75*(1), 83–119.

Levine, P. B. (2004). *Sex and Consequences: Abortion, Public Policy, and the Economics of Fertility*. Princeton, NJ: Princeton University Press.

Maier-Pfeiffer, A., K. Kutz, and M. Sandner (2013). Kosten, Nutzen und Finanzierung in den Frühen Hilfen. In T. Brand and T. Jungmann (Eds.), *Kinder schützen, Familien stärken*, pp. 111–124. Weinheim: Beltz Juventa.

Meyer, B. and D. Rosenbaum (2001). Welfare, the Earned Income Tax Credit, and the Labor Supply of Single Mothers. *Quarterly Journal of Economics, 116*(3), 1063–1114.

Miller, T. R. (2015). Projected Outcomes of Nurse-Family Partnership Home Visitation During 1996-2013, USA. *Prevention Science 16*(6), 765–777.

Olds, D., J. Eckenrode, C. Henderson, H. Kitzman, J. Powers, R. Cole, and et al. (1997). Long-Term Effects of Home Visitation on Maternal Life Course and Child Abuse and Neglect: 15-Year Follow-Up of a Randomized Trial. *The Journal of the American Medical Association 278*(8), 637–643.

Olds, D., C. Henderson, C. Phelps, and et. al (1993). Effects of Prenatal and Infancy Nurse Home Visitation on Government Spending. *Med Care* (31), 155–174.

Olds, D., C. Henderson, R. Tatelbaum, and R. Chamberlin (1988). Improving the Life-Course Development of Socially Disadvantaged Mothers: A Randomized Trial of Nurse Home Visitation. *American Journal of Public Health 78*(11), 1436–1445.

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. *Archives of Pediatrics and Adolescent Medicine 164*(5), 419–424.

Olds, D., H. Kitzman, R. Cole, J. Robinson, K. Sidora, D. Luckey, and et al (2004). Effects of Nurse Home-Visiting on Maternal Life Course and Child Development: Age 6 Follow-Up Results of a Randomized Trial. *Pediatrics 114*(6), 1550–1559.

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.

Olds, D. L. (2006). The Nurse–Family Partnership: An Evidence-Based Preventive Intervention. *Infant Mental Health Journal 27*(1), 5–25.

Olds, D. L., H. Kitzman, C. Hanks, R. Cole, E. Anson, K. Sidora-Arcoleo, D. W. Luckey, C. R. Henderson, J. Holmberg, R. A. Tutt, A. J. Stevenson, and J. Bondy (2007). Effects of Nurse Home Visiting on Maternal and Child Functioning: Age-9 Follow-up of a Randomized Trial. *Pediatrics 120*(4), e832–e845.

Raute, A. (2017). Can Financial Incentives Reduce the Baby Gap? Evidence from a Reform in Maternity Leave Benefits. *NBER Working Paper 23793*.

Robling, M., M.-J. Bekkers, K. Bell, C. C. Butler, R. Cannings-John, S. Channon, B. C. Martin, J. W. Gregory, K. Hood, A. Kemp, J. Kenkre, A. A. Montgomery, G. Moody, E. Owen-Jones, and et al. (2016). Effectiveness of a Nurse-Led Intensive Home-Visitation Programme for First-Time Teenage Mothers (Building Blocks): A Pragmatic Randomised Controlled Trial. *The Lancet 387*(10014), 146–155.

Salkind, N. J. (2010). *Encyclopedia of Research Design*. Thousand Oaks and CA: Sage.

Sandner, M., T. Cornelissen, T. Jungmann, and P. Herrmann (2017). Evaluating the effects of a targeted home visiting program on maternal and child outcomes. *HCEO Workingpaper 2017_31.*

Sandner, M. and T. Jungmann (2017). Gender-Specific Effects of Early Childhood Intervention: Evidence from a Randomized Controlled Trial. *Labour Economics 45*, 59–78.

Schmieder, J. F., T. von Wachter, and S. Bender (2012). The Effects of Extended Unemployment Insurance Over the Business Cycle: Evidence from Regression Discontinuity Estimates Over 20 Years. *Quarterly Journal of Economics 127*(2), 701–752.

Sierau, S., V. Dähne, T. Brand, V. Kurtz, K. von Klitzing, and T. Jungmann (2015). Effects of Home Visitation on Maternal Competencies and Family Environment, and Child Development: a Randomized Controlled Trial. *Prevention Science 16*(7), 1–12.

Stanley, M., I. Floyed, and M. Hill (2016). *TANF Cash Benefits Have Fallen by More Than 20 Percent in Most States and Continue to Erode.* Washington DC.

Statistisches Bundesamt (2014). Schwangerschaftsabbrüche - Fachserie 12 Reihe 3 - 2014.

The National Center for Children Families and Communities (2005). *Sample NFP: Sample County Health Department -Inition (August 1999) through December 31, 2003.*

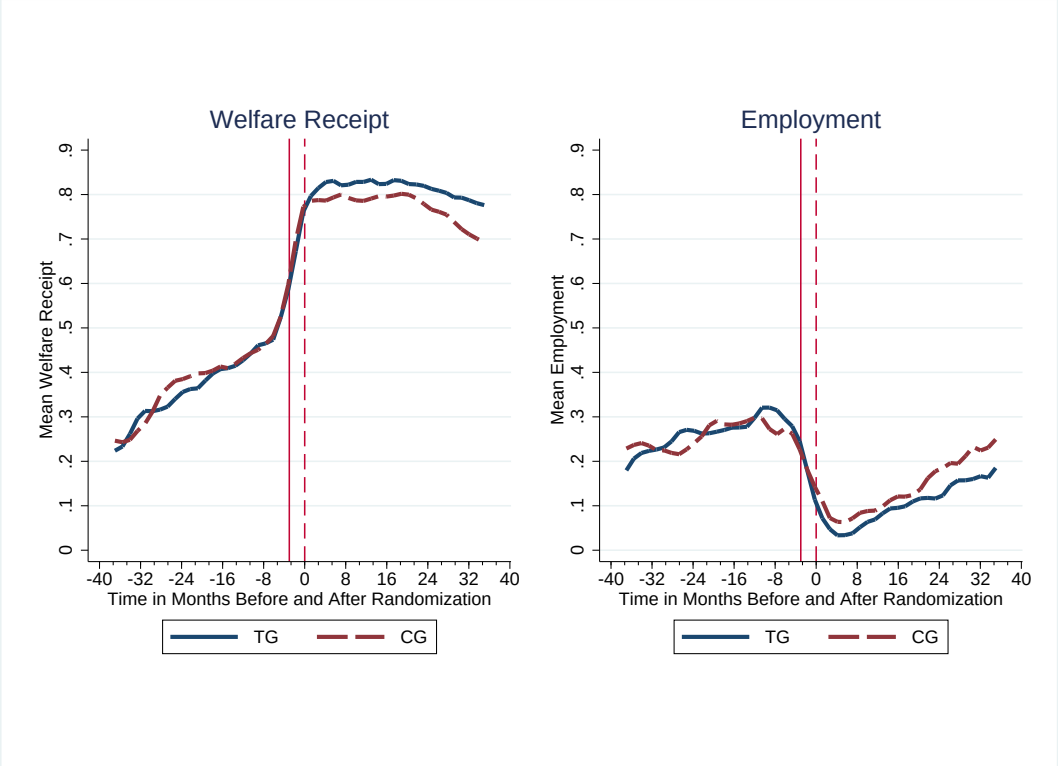
U.K. Department of Health (2013). Family Nurse Partnership programme to be extended: https://goo.gl/OZZbnJ.

U.S. Department of Health and Human Services (2015). HHS Awards $386 Million to Support Families Through the Home Visiting Program: http://www.hhs.gov/news/press/2015pres/02/20150219a.html.

Appendix A

See Figure A1 and Tables A1-A10.

Figure A1: Employment and Welfare, Before and After Randomization - Administrative Data



Note: The figure displays the average rate of second birth, employment and welfare receipt for participants in treatment group (TG) and control group (CG) from randomization until 36 months after birth. The solid vertical line shows the month of randomization. The dashed vertical line shows the month of birth of the treatment child.

Table A1: NFP Results Elmira, Memphis and Denver

NFP Results Elmira

Outcome	Observation Period		
	6 Months	**4 Years**	**15 Years**
School:	More School Enrolment of School Dropouts		
Employ.:		More Employment (15.54 Mon. vs. 8.64 Mon.)	More Employment (95 Mon. vs. 80 Mon.) ($p<0.1$)
Fertility:		Fewer Subsequent Pregnancies (0.58 vs. 1.02)	Fewer Subsequent Pregnancies (1.5 vs. 2.2) Fewer Subsequent Births (1.1 vs. 1.6) Longer Interval Between First and Subsequent Birth (65 Mon. vs. 37 Mon.)
Welfare:			Less Mon. Eligible to Welfare (60 Mon. vs. 90 Mon.)

NFP Results Memphis

Outcome	Observation Period			
	2 Years	**6 Years**	**9 Years**	**12 Years**
Employ.:		More Employment ($p<0.1$)	More Employment ($p<0.1$)	More Employment ($p<0.1$)
Fertility:	Fewer Subsequent Pregnancies (0.36 vs. 0.47)	Fewer Subsequent Pregnancies (1.16 vs. 1.38) Fewer Subsequent Births per Year (1.08 vs. 1.28)	Fewer Cumulative Subsequent Births per Year (0.81 vs. 0.93)	
Welfare:		Less Mon. Eligible to Transfer per Year (7.21 Mon. vs. 8.96 Mon.)	Less Mon. Eligible to Transfer per Year (5.21 Mon. vs. 5.92 Mon.)	

NFP Results Denver

Outcome	Observation Period	
	2 Years	**4 Years**
Employ.:	More Employment (6.83 Mon. vs. 5.65 Mon.)	More Employment (15.13 Mon. vs. 13.38 Mon.)
Fertility:	Fewer Subsequent Pregnancies (0.29 vs. 0.41) Fewer Subsequent Births (0.12 vs. 0.19)	Longer Interval Between First and Subsequent Birth (24.51 Mon. vs. 20.39 Mon.)

Notes: If not indicated differently, all treatment effects are significant with $p<0.05$. Employ. = Employment. Mon. = Months Source: NFP Results Elmira (Olds et al., 1988, 1997), Memphis (Kitzman et al., 1997; Olds et al., 2004, 2007, 2010), Denver (Olds et al., 2002, 2004)

Table A2: Randomization Outcomes per Municipality

Federal State	Community	CG	TG	Enrollment Period
Lower Saxony	Braunschweig	26	32	1.11.2006 - 30.4.2009
	Celle	15	25	
	Garbsen	10	12	
	Göttingen	12	13	
	Laatzen	4	4	
	Wolfsburg	11	15	
	Hannover	54	52	
Bremen	Bremen	77	83	15.4.2007 - 15.3.2009
	Bremerhaven	31	29	
Saxony	Leipzig	36	44	1.1.2008 - 31.12.2009
	Plauen	13	18	
	Muldentalkreis	16	12	
	Dresden	46	43	
	Vogtlandkreis	10	12	
$\sum$		**361**	**394**	

Table A3: Baseline Variable Definitions - Demographic Characteristics

Variable	Type	Description	n
Age in Years	Metric	Participants' Age in Years at Baseline	755
Week in Pregnancy	Metric	Week in Pregnancy at Randomization	755
Teenage	Binary	1 if Participant is Younger than 20 Years	755
Foreign Nationality	Binary	1 if Participant has no German Nationality	755
Monthly HH-Income in €	Metric	Monthly Net-Income in Participants' Household	647
Debt over € 3000	Binary	1 if Debt is over € 3000 in Participants' Household	728
Education Risk	Binary	1 if Participant has less than 11 Years of Schooling	755
Income Risk	Binary	1 if Net-Income is below € 1250 in Participants' Household	647
Employment Risk	Binary	1 if Participant has no Regular Employment	755
No Partner	Binary	1 if Participant is in a Partnership	755
Unmarried	Binary	1 if Participant is not married or living in divorce	755
Living with Parents	Binary	1 if Participant Lives in her Parents Household	751
Persons in HH	Metric	Number of Persons in Participants' Household at Baseline	737

Table A4: Baseline Variable Definitions - Psychological and Physical Characteristics

Variable	Type	Description	n
Unwanted Pregnancy	Binary	1 if Participant States that Pregnancy was Unwanted	747
Daily Smoking	Binary	1 if Participant Smokes Daily	755
Isolation	Binary	1 if Participant has Infrequently Contact to Friends or Relatives	747
Foster Care Experience	Binary	1 if Participant Lived at Least Once in a Foster Family or Foster Care	735
Neglect Experience	Binary	1 if Indication of Neglect Experience during Childhood	730
Lost Experience	Binary	1 if Participant Lost an Attachment Figure due to Death or Divorce	736
Violence Experience	Binary	1 if Participant ever Experienced Violence in her Life	751
Depression	Binary	1 if Value higher 20 for Depression on the Depression Anxiety Stress Scale (DASS)	749
Anxiety	Binary	1 if Value higher 15 on Anxiety on the DASS	744
Stress	Binary	1 if Value higher 25 on Stress on the DASS	749
Aggression	Binary	1 if Value higher 10 on the *Fragebogen zur Erfassung von Aggressivitätsfaktoren (FAF)*	743
Medically Indicated Risk Preg.	Binary	1 if participant has physical problems or if participant is older than 35	724
Body-Mass-Index	Metric	Participants' $Weight/Height^2$ (Weight Before Pregnancy)	750
Sum Risk Factors	Metric	Sum of Risk Factors (Risk factors include: Being Underage, No Graduation, Low Income, No Employment, Unwanted Pregnancy, Alcohol Consumption Once a Week, Regular Drug Use, No Partner, Social Isolation, Foster Care Experience, Experience of Neglect, Experience of Loss, Experience of Violence (During Pregnancy), Experience of Violence (Ever), Mental Illness, Depression, Anxiety, Stress and Aggression.	755

Table A5: Topical Focus of the Home Visits in NFP and *Pro Kind*

	Pro Kind Average	NFP-Average	Recommended Average by NFP
During Pregnancy			
Maternal Health	28%	37%	35-40%
Maternal and Parental Role	19%	23%	23-25%
Environmental Health	10%	11%	5-7%
Life Course Development	16%	13%	10-15%
Family and Friends	15%	16%	10-15%
Social and Health Services	12%	-	-
During Infancy			
Maternal Health	16%	20%	14-20%
Maternal and Parental Role	30%	36%	45-50%
Environmental Health	11%	14%	7-10%
Life Course Development	17%	15%	10-15%
Family and Friends	14%	15%	10-15%
Social and Health Services	11%	-	-
During Toddlerhood			
Maternal Health	13%	17%	10-15%
Maternal Role	30%	37%	40-45%
Environmental Health	10%	14%	7-10%
Life Course Development	22%	17%	18-20%
Family and Friends	14%	15%	10-15%
Social and Health Services	11%	-	-

Notes: The percentage rates give the share of the total time in the family, which the home visitors spent for a certain topic. The data is collected by a documentation system, in which the home visitors note the duration and the covered topic for each home visit. Source: Jungmann et al. (2009); The National Center for Children Families and Communities (2005).

Table A6: Selective Attrition between TG and CG Psychological Characteristics - Administrative and Survey Data

	Difference TG/CG for:				
	Merged	At Least One Interview After Birth	Data Available for 12 Months After Birth	Data Available for 24 Months After Birth	Complete data from Birth Until Third Birthday
	(1)	(2)	(3)	(4)	(5)
Unwanted Pregnancy	0.0122	0.0224	0.0318	0.0183	-0.00863
	(0.0288)	(0.0310)	(0.0313)	(0.0333)	(0.0416)
Daily Smoking	0.00186	0.000532	-0.0133	-0.00888	-0.0256
	(0.0360)	(0.0384)	(0.0407)	(0.0442)	(0.0540)
Isolation	-0.00685	-0.0146	-0.00474	-0.00712	0.0151
	(0.0189)	(0.0204)	(0.0213)	(0.0246)	(0.0319)
Foster Care Exper.	0.0409	0.0471	0.0424	0.0548	0.0573
	(0.0313)	(0.0321)	(0.0338)	(0.0359)	(0.0430)
Neglect Experience	0.00810	-0.00346	-0.0136	-0.00800	0.0396
	(0.0368)	(0.0393)	(0.0416)	(0.0460)	(0.0565)
Lost Experience	-0.0474	-0.0679*	-0.0667	-0.0485	0.000505
	(0.0377)	(0.0408)	(0.0431)	(0.0480)	(0.0585)
Violence Ever	-0.00510	-0.00210	-0.0127	-0.0247	-0.0393
	(0.0211)	(0.0213)	(0.0219)	(0.0239)	(0.0318)
Depression	-0.0154	-0.00256	0.00532	0.0110	0.0173
	(0.0241)	(0.0250)	(0.0262)	(0.0289)	(0.0368)
Anxiety	-0.00761	0.00400	0.00552	0.00189	0.00193
	(0.0287)	(0.0301)	(0.0315)	(0.0348)	(0.0438)
Stress	0.0329	0.0277	0.0214	0.0202	0.00161
	(0.0348)	(0.0374)	(0.0394)	(0.0438)	(0.0540)
Aggression	-0.0328	-0.0450	-0.0462	-0.0652*	-0.0819**
	(0.0282)	(0.0294)	(0.0312)	(0.0336)	(0.0401)
Body-Mass-Index	-0.0154	-0.265	-0.114	-0.170	0.391
	(0.401)	(0.445)	(0.477)	(0.519)	(0.652)
Medic. Indic. Risk Preg.	0.00459	0.0135	0.0113	-0.0132	-0.00358
	(0.0240)	(0.0255)	(0.0274)	(0.0297)	(0.0373)
Sum Risk Factors	-0.0336	-0.120	-0.140	-0.121	-0.0928
	(0.184)	(0.192)	(0.200)	(0.217)	(0.271)
Observations	*703*	*602*	*539*	*438*	*296*

Notes: Column (1) contains estimates of the average difference in characteristics between mothers in the control and treatment group for the participants merged with the administrative data. Column (2)-(5) contain these estimates for the survey data. Dependent variables are shown in the first column. The treatment indicator has the value one if the mother is in the treatment group. Robust standard errors are shown in parentheses. See Appendix Tables A3 and A4 for variable definitions.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A7: Selective Attrition between "Attritors" and "Non-Attritors"

	Difference "Attritors" / "Non-Attritors" for:				
	Merged	At Least One Interview After Birth	Data Available for 12 Months After Birth	Data Available for 24 Months After Birth	Complete data from Birth Until Third Birthday
	(1)	(2)	(3)	(4)	(5)
Age in Years	0.801	1.261**	1.679***	1.858***	2.180***
	(0.623)	(0.390)	(0.344)	(0.313)	(0.313)
Week in Pregnancy	-0.480	1.404**	1.162*	0.808	1.060*
	(0.829)	(0.520)	(0.463)	(0.424)	(0.428)
Foreign Nationality	-0.204***	-0.0068	0.0011	0.0010	0.036*
	(0.0382)	(0.0245)	(0.0218)	(0.0200)	(0.0201)
Teenage	-0.0505	-0.137**	-0.150***	-0.165***	-0.201***
	(0.0716)	(0.0449)	(0.0398)	(0.0363)	(0.0364)
Mon. HH-Inc. in €	-61.91	194.9***	111.0*	135.3**	158.7***
	(85.35)	(53.59)	(47.53)	(42.64)	(42.55)
Debt over 3000 €	0.0902	0.0374	0.0513	0.0386	0.0538
	(0.0552)	(0.0348)	(0.0309)	(0.0283)	(0.0286)
Education Risk	0.0167	-0.130***	-0.153***	-0.159***	-0.170***
	(0.0610)	(0.0381)	(0.0337)	(0.0307)	(0.0310)
Income Risk	0.0693	-0.0686	-0.0652*	-0.0858**	-0.106***
	(0.0559)	(0.0351)	(0.0312)	(0.0285)	(0.0288)
Employment Risk	-0.00974	-0.0732*	-0.0790**	-0.0905***	-0.121***
	(0.0531)	(0.0334)	(0.0296)	(0.0271)	(0.0272)
No Partner	0.164*	-0.00840	-0.0384	0.000605	0.0552
	(0.0648)	(0.0410)	(0.0365)	(0.0334)	(0.0337)
Unmarried	0.0963*	-0.0227	-0.0391*	-0.0351*	-0.0572**
	(0.0396)	(0.0251)	(0.0223)	(0.0204)	(0.0205)
Living with Parents	-0.0840	-0.00294	-0.0267	-0.0346	-0.0352
	(0.0648)	(0.0410)	(0.0363)	(0.0331)	(0.0334)
Persons in HH	-0.312	-0.0562	-0.195	-0.194	-0.163
	(0.234)	(0.151)	(0.133)	(0.122)	(0.124)
Unwanted Pregnancy	0.0418	0.00448	-0.0617*	-0.0816**	-0.0409
	(0.0545)	(0.0343)	(0.0305)	(0.0278)	(0.0282)
Daily Smoking	0.158*	-0.0502	-0.0309	-0.0844*	-0.0520
	(0.0679)	(0.0429)	(0.0382)	(0.0348)	(0.0353)
Isolation	-0.0485	-0.0185	-0.0184	0.00138	0.0179
	(0.0367)	(0.0232)	(0.0206)	(0.0189)	(0.0191)
Foster Care Exper.	0.0859	-0.116**	-0.0885**	-0.109***	-0.0862**
	(0.0590)	(0.0370)	(0.0329)	(0.0301)	(0.0305)
Neglect Experience	0.119	-0.0889*	-0.0641	-0.0625	-0.0140
	(0.0697)	(0.0439)	(0.0391)	(0.0358)	(0.0362)
Lost Experience	0.0587	0.00802	0.00973	-0.0509	-0.0322
	(0.0718)	(0.0453)	(0.0403)	(0.0368)	(0.0373)
Violence Ever	0.00843	-0.0576*	-0.0564*	-0.0442*	-0.00606
	(0.0401)	(0.0252)	(0.0224)	(0.0205)	(0.0208)
Depression	-0.0194	-0.0587*	-0.0507*	-0.0383	-0.00834
	(0.0462)	(0.0290)	(0.0258)	(0.0237)	(0.0240)
Anxiety	0.0211	-0.0611	-0.0553	-0.0435	-0.00755
	(0.0545)	(0.0343)	(0.0305)	(0.0279)	(0.0283)
Stress	0.0765	-0.0229	-0.0309	-0.0178	0.00896
	(0.0660)	(0.0416)	(0.0370)	(0.0339)	(0.0343)
Aggression	0.0525	-0.0563	-0.0358	-0.0486	-0.0423
	(0.0533)	(0.0335)	(0.0298)	(0.0273)	(0.0276)
Body-Mass-Index	0.200	0.433	1.015*	0.908*	0.882*
	(0.766)	(0.483)	(0.428)	(0.392)	(0.396)
Medic. Indic. Risk Preg.	-0.00159	-0.0211	-0.00257	-0.0157	0.00158
	(0.0457)	(0.0288)	(0.0256)	(0.0235)	(0.0237)
Sum Risk Factors	0.752*	-0.772***	-0.837***	-0.879***	-0.587**
	(0.349)	(0.219)	(0.194)	(0.177)	(0.180)
Lower Saxony	-0.110	-0.0413	-0.0530	-0.0539	0.0160
	(0.0697)	(0.0440)	(0.0391)	(0.0358)	(0.0362)
Bremen	0.0843	0.0769	0.0626	0.0650	0.0730*
	(0.0652)	(0.0410)	(0.0365)	(0.0334)	(0.0338)
Saxony	0.0252	-0.0356	-0.00958	-0.0111	-0.0890*
	(0.0677)	(0.0426)	(0.0379)	(0.0347)	(0.0350)
Observations	*755*	*755*	*755*	*755*	*755*

Notes: Column (1) contains estimates of the average difference in characteristics between mothers merged and mothers who are not merged with the administrative data. Column (2)-(5) contain these estimates for mothers who participated and non participated in the survey. Dependent variables are shown in the first column. The treatment indicator has the value one if the mother is merged or participated in the survey. Robust standard errors are shown in parentheses. See Appendix Tables A3 and A4 for variable definitions.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A8: Maternal Life Course Outcomes 36 Months after Birth of the Treatment Child - Administrative Data

	Extensive Margin			Intensive Margin (in Months)		
	Control Mean	Diff. TG/CG	p-values	Control Mean	Diff. TG/CG	p-values
	(1)	(2)	(3)	(4)	(5)	(6)
			0 - 36 Months after Birth			
Any Employment	0.511	-0.092***	0.007	6.224	-1.557***	0.008
	[0.501]	(0.027)		[8.919]	(0.468)	
Parttime/Fulltime Employed	0.188	-0.056**	0.045	1.592	-0.601	0.104
	[0.392]	(0.026)		[4.744]	(0.336)	
Apprenticeship	0.206	-0.039	0.297	2.298	-0.295	0.504
	[0.406]	(0.036)		[5.852]	(0.442)	
Marginal employment	0.291	-0.054*	0.099	2.252	-0.642**	0.048
	[0.455]	(0.030)		[5.062]	(0.284)	
Welfare	0.951	0.017	0.194	29.17	1.085	0.251
	[0.235]	(0.013)		[10.81]	(0.892)	
Observations	*329*	*684*		*329*	*684*	
Second Child in HH	0.187	0.065**	0.032			
	[0.363]	(0.026)				
Observations	*316*	*663*				

Notes: Standard errors clustered on community level in square brackets; Standard deviations in parentheses. Columns (2) and (5) report the coefficient and standard error on Home Visiting (HV) from estimating equation (1) by OLS. Data is available on a monthly base. Estimations include community fixed effects and controls for age and being underaged. TG = Treatment Group; CG = Control Group; HH = Household.
* p < 0.1, ** p < 0.05, *** p < 0.01

Table A9: Heterogeneous Effects - Administrative Data

	Second Birth		**Welfare (Intensive Margin)**		**Employment (Intensive Margin)**	
	Effect Size	p-values	Effect Size	p-values	Effect Size	p-values
	(1)	(2)	(3)	(4)	(5)	(6)
Heterogeneous Effects for Teenager						
TG * Teenager	-0.052	0.429	1.199	0.466	0.999	0.429
	(0.065)		(1.645)		(1.263)	
Teenager	0.024	0.593	-0.933	0.466	-1.812*	0.053
	(0.046)		(1.279)		(0.935)	
TG	0.085**	0.033	1.200	0.264	-1.840**	0.025
	(0.040)		(1.075)		(0.819)	
Observations	*677*		*703*		*703*	
Heterogeneous Effects by Profession of the Home Visitor						
TG * Family Midwife	-0.044	0.502	-1.372	0.408	-0.978	0.456
	(0.066)		(1.657)		(1.311)	
Family Midwife	0.002	0.961	-1.064	0.390	-0.118	0.909
	(0.045)		(1.237)		(1.029)	
TG	0.086**	0.043	1.846*	0.077	-1.294	0.112
	(0.042)		(1.042)		(0.813)	
Observations	652		673		673	
Heterogeneous Effects for Non-Employed						
TG * Any Employment (0-24 M.)	-0.087	0.127				
	(0.059)					
Any Employment (0-24 M.)	-0.100**	0.016				
	(0.042)					
TG	0.085**	0.042				
	(0.042)					
Observations	*677*					

Notes: Robust standard errors in parentheses. Columns (1), (3) and (5) report the coefficient and standard error on Home Visiting (HV) from estimating equation (1) extended by the interaction between treatment group and subgroup indicator. Family Midwife is one for mothers living in municipalities (Bremen, Bremerhaven, Leipzig) where only family midwifes conducted the home visits. Data is available on a monthly base. TG = Treatment Group; M = Months.
* p < 0.1, ** p < 0.05, *** p < 0.01

Table A10: Maternal Life Course Outcomes 36 Months after the Birth of the Treatment Child - Unweighed Survey Data

	Extensive Margin			Intensive Margin (in Months)		
	Control Mean	Diff. TG/CG	p-values	Control Mean	Diff. TG/CG	p-values
	(1)	(2)	(3)	(4)	(5)	(6)
Any Employment	0.555	-0.008	0.896	7.569	-0.752	0.481
	[0.499]	(0.058)		[9.231]	(1.066)	
Parttime/Fulltime Employed	0.299	-0.010	0.852	2.365	-0.522	0.339
	[0.460]	(0.053)		[5.087]	(0.544)	
Apprenticeship	0.255	-0.035	0.479	2.672	0.442	0.554
	[0.438]	(0.049)		[5.810]	(0.744)	
Marginal employment	0.248	-0.015	0.757	2.533	-0.671	0.272
	[0.434]	(0.050)		[5.705]	(0.610)	
Welfare	0.912	0.050*	0.084	26.511	1.274	0.301
	[0.284]	(0.028)		[11.017]	(1.230)	
Second Birth	0.175	0.102**	0.036			
	[0.382]	(0.048)				
Second Pregnancy	0.321	0.031	0.574			
	[0.469]	(0.055)				
Inconsistent Use of Contraceptives	0.226	0.019	0.702			
	[0.419]	(0.049)				
Constant Partnership	0.401	-0.005	0.927			
	[0.491]	(0.057)				
Change in Marriage Status	0.183	-0.025	0.566			
	[0.387]	(0.044)				
School	0.102	-0.014	0.681	0.934	0.072	0.879
	[0.304]	(0.025)		[3.877]	(0.331)	
Childcare Utilization	0.584	0.083	0.144	7.175	1.894*	0.071
	[0.495]	(0.056)		[8.571]	(1.046)	
Observations	*137*	*296*		*137*	*296*	

Notes: Standard deviations in square brackets; robust standard errors in parentheses. Columns (2) and (5) report the coefficient and standard error on Home Visiting (HV) from estimating equation (1) by OLS. Data is available on a monthly base from affiliation to 36 months after birth. TG = Treatment Group; CG = Control Group; HH = Household.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A11: Descriptive Statistics for Life-Satisfaction and Stress

	Control Group		Treatment Group	
	Mean	sd	Mean	sd
	Face-to-Face Interview at 12 Months			
How stressful do you perceive the following areas? (1=Not stressful at all; 4= Very stressful)				
School/Job	2.14	1.04	1.70	0.91
Child/Children	1.68	0.62	1.67	0.61
Financial Situation	2.36	0.92	2.26	0.93
Housing Situation	1.97	1.06	1.83	1.01
Relation to Child's Father	1.98	1.08	1.94	1.13
Detachment from Parents	1.43	0.75	1.32	0.66
Homework	1.69	0.78	1.67	0.73
Leisure	1.56	0.78	1.49	0.75
Social Life	1.47	0.70	1.40	0.68
Daily Issues	1.88	0.74	1.82	0.68
Overall Life Situation	1.94	0.76	1.85	0.65
Please rate how satisfied are you overall with your life: (1= Not at all; 4= Completely true)				
Life Could Hardly Be Happier	2.97	0.73	3.11	0.66
Many Things Will Come True	3.10	0.61	3.17	0.61
I am Satisfied With My Life	3.17	0.67	3.16	0.68
Time Will Bring Interesting and Pleasant Things	3.61	0.57	3.60	0.55
I am Satisfied With My Life Situation	2.97	0.73	3.01	0.69
How satisfied are you with your... (1= very unsatisfied; 4= very satisfied)				
Housing Situation	2.78	1.07	2.85	1.05
Financial Situation	2.34	0.88	2.50	0.90
Level of Education	2.22	0.98	2.44	1.05
Occupational Situation	2.24	0.95	2.39	0.97
Support by the Child's Father	2.66	1.19	2.80	1.18
Observations	*198*		*224*	
	Telephone Interview at 28 Months			
How often or seldom have you experienced this feeling in the last four weeks? (1= very often; 4= very seldom)				
Angry	3.05	3.05	3.05	1.09
Worried	2.09	1.04	1.77	0.94
Happy	3.66	0.90	3.76	0.87
Sad	2.71	1.07	2.49	1.03
How satisfied are you today with the following areas of your life? (0= very unsatisfied; 10= very satisfied)				
Health	6.55	2.96	6.83	2.88
Housework	6.93	2.32	7.37	2.30
Household Income	4.92	2.68	5.58	2.89
Personal Income	4.14	2.87	4.57	3.04
Place of Dwelling	6.56	3.15	6.63	3.12
Free Time	5.67	2.91	6.23	2.86
Child Care Availability	6.73	2.99	6.68	3.28
Family Life	7.46	2.35	7.63	2.52
Life in General	7.13	2.10	7.44	1.91
Observations	*195*		*237*	