SOEP The German Socio-Economic Panel study

799 2015

SOEPpapers
on Multidisciplinary Panel Data Research

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

Malte Sandner

DIW SOEP

SOEPpapers on Multidisciplinary Panel Data Research at DIW Berlin

This series presents research findings based either directly on data from the German Socio-Economic Panel study (SOEP) or using SOEP data as part of an internationally comparable data set (e.g. CNEF, ECHP, LIS, LWS, CHER/PACO). SOEP is a truly multidisciplinary household panel study covering a wide range of social and behavioral sciences: economics, sociology, psychology, survey methodology, econometrics and applied statistics, educational science, political science, public health, behavioral genetics, demography, geography, and sport science.

The decision to publish a submission in SOEPpapers is made by a board of editors chosen by the DIW Berlin to represent the wide range of disciplines covered by SOEP. There is no external referee process and papers are either accepted or rejected without revision. Papers appear in this series as works in progress and may also appear elsewhere. They often represent preliminary studies and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be requested from the author directly.

Any opinions expressed in this series are those of the author(s) and not those of DIW Berlin. Research disseminated by DIW Berlin may include views on public policy issues, but the institute itself takes no institutional policy positions.

The SOEPpapers are available at
http://www.diw.de/soeppapers

ISSN: 1864-6689 (online)

German Socio-Economic Panel (SOEP)
DIW Berlin
Mohrenstrasse 58
10117 Berlin, Germany

Contact: Uta Rahmann | soeppapers@diw.de

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

Malte Sandner*

NIW Hannover / Leibniz University Hannover

October, 2015

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 and increased subsequent births. These results contradict those of previous studies from the United States, where home visiting programs successfully increased employment and decreased fertility. Analyzing the reason for the different results, suggests that the program interacts with low employment incentives and generous welfare state arrangements for disadvantaged mothers with young children in Germany.

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

*Financial support by 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 for support and helpful comments.

1 Introduction

Home visiting is an early childhood intervention for disadvantaged mothers that not only aims to improve child outcomes but also to improve maternal outcomes such as employment and family planning. In home visiting programs family midwives affect these outcomes by consulting mothers for a longer period after birth. The family midwives intend to enhance maternal skills (e.g. attachment behavior, interactions, and teaching skills) and to increase women's personal strengths, including self-efficacy, problem-solving abilities and self-esteem. Home visiting programs are popular in many developed countries and particularly in the U.S., where the Obama administration has requested $500 million for fiscal year 2016 and $15 billion over the next 10 years to continue to expand these programs (U.S. Department of Health and Human Services, 2015).

Although home visiting programs aim to improve aspects of the maternal life course, it is arguable whether they will reach 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 higher occupational aspirations, the mothers may decide to delay further birth. 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 positively influence fertility decisions and lead to longer absences from the workforce as a consequence. The only evidence from randomized field experiments regarding which of the two effects predominates comes from the U.S., where home visits successfully decreased fertility and increased maternal employment (Olds et al., 2007, 1997; Brooks-Gunn et al., 1994).

However, it is likely that the fertility increasing effect will dominate in a European welfare system, because in most European countries, and particularly in Germany, social assistance rules for mothers with small children are more generous compared with those in the U.S.[1] The German rules include mean-tested welfare

[1] In 1996, the Temporary Assistance to Needy Families (TANF) program eliminated the legal entitlement to cash welfare by imposing a 60-month lifetime limit on benefits and requiring individuals to leave welfare for work after two years. Furthermore, three of the four stated goals of the TANF program involved reducing non-marital births and encouraging marriage (Blank, 2002). However, even Aid to Families with Dependent Children (AFDC), the program that preceded TANF, was stricter than the welfare system in Germany today. Under AFDC, only single mothers were eligible for cash benefits, which were rather low (the monthly benefits for a single-parent family with

payments which increase with parity and waives work requirements or benefit cuts until the child's third birthday. Additionally, financial incentive programs that encourage work among low-income families with children, such as the Earned Income Tax Credit (EITC) in the U.S., 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 subsequent births instead of higher employment.

This paper analyses the first randomized experiment of one such home visiting program for disadvantaged first-time mothers in Germany, named *Pro Kind*. The results indeed suggest that the intervention increased fertility and decreased employment. The effects are sizable, implying that the probability of a second birth increased among the intervention group by 36 percent and employment decreased by 24 percent relative to the mean of the control group. A lower number of abortions among the women in the treatment group mainly explains the effect on fertility. The effect on abortions is not caused by more favorable family environments (e.g. more stable partnerships) in the treatment group. However, the intervention positively influenced subjective maternal well-being and life satisfaction, which might have influenced abortion and fertility decisions. I can reject the possibility that differences in the content, implementation, and participants of the home visiting programs are able to explain the different results seen in this study compared with U.S. studies. Therefore, the most compelling reason for the different results is the arrangement of the German welfare state.

My analysis draws on administrative data from the German social security system, containing information on employment, wages, welfare benefits and household composition, and on data from biannual telephone interviews. The administrative data are available for over 90% of the sample over the first three years after the birth of the first child. They are objectively measured and should not be biased by the treatment and control groups differentially reporting outcomes. The survey

two children and no income ranged from $120 in Mississippi to $597 in Vermont). Additionally, AFDC primarily used in-kind transfers, such as food stamps, and included significant work obligations (Moffitt, 1998; Gebhardt and Jacobs, 1997).

data allow for the examination of a much richer set of outcomes, such as fertility planning, childcare use, and subjective statements about well-being and life satisfaction, allowing me to identify channels for the findings. To my knowledge, this is the first study combining administrative and survey data to evaluate the effects of an early childhood intervention.

The paper provides new insights into how welfare systems influence fertility. Although the literature presents inconclusive findings regarding whether welfare arrangements (e.g. the amount of welfare or family caps) affect fertility (Moffitt, 1998; Grogger and Bronars, 2001; Kearney, 2004), the effects could be more clear if the welfare system interacts with early childhood interventions. The findings of the *Pro Kind* program suggest that a more generous welfare system can increase fertility when combined with home visiting or other counseling services; additionally, the greater effects on employment in the U.S. imply that these services may increase the effectiveness of workfare reforms for mothers. Attention to these results might be helpful when considering policies from the U.S. that may be implemented in Europe in the future.

The effects of home visiting on fertility can also contribute to the understanding how early childhood interventions generate effects on children. For example, literature in economics has shown that shorter spacing between births has negative effects on the test scores of older siblings (Buckles and Munnich, 2012) and 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). Therefore, the results on fertility might reduce the potential of home visiting to improve child development. The finding of Sandner and Jungmann (2015) that the *Pro Kind* program has smaller effects on child development compared with studies in the U.S. supports this suggestion.

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 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 examine the impact of early childhood programs in general, and of home visiting programs in particular, on parents. For example, Heckman et al. (2010) evaluated 715 outcomes in 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 Nurse Family Partnership Program (NFP). This program is conceptually similar 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. The studies do not present information whether less sexual activity or more frequent contraception use lead to less 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. Only one study (Brooks-Gunn et al., 1994) in addition to the NFP analyzed the effects of home visiting on the maternal life course using a randomized experiment. In that study, home visiting significantly decreased maternal unemployment.

Cost/benefit analyses of the Elmira and Memphis trials indicate that the NFP reaches 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 these 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.

3 The *Pro Kind Program*: A Social Experiment

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 material, 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 conduct the home visits either continuously or in a tandem model with social pedagogues and a pediatric nurse (Brand and Jungmann, 2012). 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. 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 regularly receive feedback, encouragement, reflection, and support from nurse supervisors.

The *Pro Kind* program only registers 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. 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% 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 A1 and Figure A1). 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.

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 a mother lacks childcare arrangements. As a result, there are few incentives

for mothers to participate in the labor market. Furthermore, in addition to the *Pro Kind* program, various services offer help and support, especially for mothers (e.g., the labor agency provides special programs for unemployed people who are younger than 25 years of age and for single mothers).

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 strongest factor that initiates and deepens this relationship is the mother's first experience with a newborn child. Olds et al. (2010) state that through this relationship, nurses can help parents to 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 at their most open-minded are home visiting programs' greatest advantages compared with other interventions.

3.2 Randomization Process and Sample Description

The causal effects of the *Pro Kind* intervention are 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 the woman met the criteria, the supervisor visited the woman at her home. During this visit, participants or, if they were underage, their parents, signed an informed consent form for participating in the study. Thereafter, participants completed a baseline questionnaire to assess demographic and psychological characteristics, as well as risk factors. Up to this point, the mothers had only received 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.[2] After answering the baseline questionnaire, women received the results of

[2]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).

the randomization that assigned them either to the home visit or the control group. The final sample for the *Pro Kind* experiment consists 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.[3] 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.

Table 1 reports the means and the differences in means according to treatment status for the baseline variables.[4] Differences in the average characteristics of the control and treatment groups are small and generally not statistically significant. Migrant status, defined among the mothers as not having German citizenship or not having been born in Germany, is the only demographic characteristic that is significantly different; the control group having a higher proportion of immigrants compared with the treatment group. None of the differences in psychological or physical risk characteristics are statistically significant. Furthermore, I conduct a test of joint significance of all the baseline characteristics. The F-statistic is 1.19; thus, the possibility that the characteristics of the treatment and control groups are the same could not be rejected. Hence, overall, the randomization appears to have successfully created comparable treatment and control groups.

An analysis of the demographic and psychological characteristics of the participants reveals that the women in both groups are young and highly disadvantaged. Most of the mothers were unemployed at the time of the baseline interview and have 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%) have less than eleven years of schooling; many of them have dropped out of school. Furthermore, the average monthly household income is €928.60. Considering the

[3]The monetary incentive was €15 for the interview during pregnancy and at 6 months, €20 for the interview at 12 months, and €25 for the interview at 24 months.

[4]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, and missing values are equally distributed between the control and treatment groups (see Appendix Tables A4 and A5). 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)
Migration Background	0.18	0.12	-0.05* (0.03)
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	0.06 (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)
Not Married			
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: Robust standard errors are reported in parentheses in column 3. Column 3 presents the coefficient on the treatment dummy from a regression model with the treatment dummy plus community dummies. See Appendix Tables A2 and A3 for variable definitions.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

mean household size of 2.49 persons, the participants' average income is 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 the *Pro Kind* Home Visiting

To monitor the quality of the program implementation, the home visitors documented each visit (e.g., duration, covered topics, maternal interest).[5] In total, 12,894 home visits with an average duration of 82 minutes were conducted. The families in

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

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 visit. 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. The home visiting documentation demonstrates that at all developmental stages, home visitors invested 40% of their time with the family to address issues related to the maternal life course and employment (Appendix Table A3). This points out that maternal life course issues and economic self-sufficiency are fundamental topics of the *Pro Kind* program.

4 Data

4.1 Administrative Data

The German Record Linkage Center (GRLC) of the Institute for Employment Research obtained individual-level labor 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.[6] 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 are, 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.[7] For all tracked participants data are available from affiliation into the project until 36 months after the birth of the treatment child. My primary outcomes of employment and welfare

[6] 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).

[7] 15 participants of the 755 participants in the baseline sample refused participation in the informed consent and were not used for the merging process.

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 is only available if the mother was either engaged in a job search, or received welfare benefits.[8]

4.2 Telephone Survey Data

In addition to the administrative data, I use data from biannual telephone interviews with the mothers. The telephone interviews begin during pregnancy and continue at six-month intervals 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, maternal well-being, and life satisfaction (Siedler et al., 2009). 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 toward each newborn child of the mother (Anger et al., 2009).

The telephone interviewers attempted to contact all of the mothers at each time point, 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 could be 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%

[8]Information on age and 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.

(n=539) of the mothers, data are available for at least 12 months after birth. 39% (n=296) participated in all interviews without missing data for any months after birth.

5 Validity of the Experimental Design

Table 2: Sample Composition Telephone Interviews

	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 Survey 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]
Observations	*755*	*394*

Notes: Robust standard errors in parentheses and p-values in brackets. Administrative data in Panel A is available for 36 months after birth of treatment child. TG = Treatment Group; CG = Control Group.

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

Table 3 presents the differences in the baseline demographic characteristics be-

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)
Migration	-0.0594**	-0.0592**	-0.0546*	-0.0548	-0.0701
	(0.0259)	(0.0298)	(0.0317)	(0.0355)	(0.0462)
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)
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: Robust standard errors are shown in parentheses. Dependent variables shown in the first column. The treatment indicator has the value one if the mother is in the treatment group. Column (1) contains estimates of the average difference in characteristics between mothers in the control and treatment group including community fixed effects for the participants merged with the administrative data. Column (2)-(5) contain these estimates for the survey data. See Appendix Tables A4 and A5 for variable definitions. TG = Treatment Group; CG = Control Group.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

tween the treatment and control groups for the administrative data (Column 1) and the survey data grouped by the 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. Alone the difference in the proportion of mothers with migrant backgrounds, which

is already significant at baseline, remains significant for almost all of the interviews.[9]

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 nonlinear models for the binary outcomes instead. In a first step, I estimate models without h_{ic} and α_c and than I include h_{ic} and α_c as a robustness check. 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 can not 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 to the intervention is not merged to the administrative data. However, the effect of treatment on the treated would be marginally different to the present results because the implementation research showed that 97.7% of the treatment group participants received at least some treatment.

[9] Appendix Table A7 shows that some characteristics and risk factors differ between those who dropped out and those who participated in the follow-up interviews. Generally, the participating mothers are older and have fewer cumulative risk factors. The only difference between the participants who are merged and those who are not merged with the administrative data is migration status. This likely because migrants participate less frequently in the labor market and are less frequently eligible for welfare benefits.

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. The next three rows separate employment into part/full time employment, apprenticeship or marginal employment. Column 4 shows the mean total number of months in one of the occupations.

Among mothers in the control group, 52 percent participated in the labor market in the first 36 months after birth. They were employed for 6.39 months on average during this period, indicating a high amount of job fluctuation and short employment periods in the sample. Participants are most frequently employed in marginal employment, but apprenticeship also plays a large role, in particular when total months employed are considered. The prevalence of apprenticeship demonstrates that many participants have not completed their vocational training before giving birth and that they are oriented towards completing it after the birth.[10]

Analyzing the treatment impact on employment reveals that home visiting reduces 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 are employed for at least one month by 8.8 percentage points to a rate of 43.4 percentage points; the average number of months employed is reduced by 1.6 months to 4.87 months, which is a 23.8 percent decrease relative to the mean time worked by the mothers in the control group. When analyzing the different types of employment, the effect is strongest for part time/full time employment for which the treatment reduced the extensive margin by 27.2 percent and the intensive margin by 39.3 percent relative to the mean of the control group.

The fourth row, "Welfare", indicates whether and for how many months on av-

[10] In Germany an apprenticeship includes on the job training in a company and attendance of a vocational school. Completing an apprenticeship is strongly correlated with labor market success in Germany.

Table 4: Maternal Life Course Outcomes 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)
Any Employment	0.521	-0.088**	0.019	6.392	-1.550**	0.018
	[0.479]	(0.038)		[9.086]	(0.652)	
Parttime/Fulltime Employed	0.191	-0.052*	0.061	1.642	-0.645**	0.043
	[0.393]	(0.028)		[4.826]	(0.319)	
Apprenticeship	0.202	-0.012	0.696	2.369	-0.223	0.620
	[0.402]	(0.030)		[5.999]	(0.449)	
Marginal Employment	0.305	-0.054	0.114	2.299	-0.664*	0.055
	[0.461]	(0.034)		[5.071]	(0.345)	
Welfare	0.964	0.030**	0.023	31.92	1.840**	0.042
	[0.295]	(0.013)		[12.71]	(0.904)	
Observations	*341*	*703*		*341*	*703*	
Second Child in HH	0.183	0.066**	0.037			
	[0.363]	(0.032)				
Observations	*323*	*677*				

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

erage a mother lived in a household that received public assistance. The figure in column 1 shows, corresponding with the affiliation criteria, that 96.4 percent of the mothers in the control group received public assistance for at least one months during the first 36 months after birth. Moreover, the total number of months (31.92) indicates that the participants' households received welfare in 88.6 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 fertility. "Second Child in HH" is a binary variable that takes value 1 if two or more children are living in the household and 0 with one child or no children living in the household.[11] Because the data record 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 live in a household with two or

[11]There can be no children in the household in the event of a miscarriage of the first pregnancy or the adoption of the treatment child.

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 by 36% relative to the mean of the the control group.

Overall, the results from the administrative data indicate that the intervention has unintended effects which are in contrast to the results of studies from the U.S. Instead of the intended higher level 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 are likely caused by the increased maternal fertility, which might reduce the available time and flexibility of the mothers for labor market participation. All effects from the administrative data also hold and become slightly larger if the models include the community and age of the mother as controls (Appendix A8). Selective attrition or reporting bias could not cause the results because employment data and public assistance are available for all mothers for the 36 months after the birth of the treatment child. Only household composition is not available for all participants. However, bias is unlikely because most households with a second child in the *Pro Kind* sample will receive welfare benefits, and therefore, they are included in the administrative data. 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 results of the telephone survey including the 296 mothers who participated in all interviews until the third birthday of the treatment child.[12] The first six 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 asks for second births and not only for household composition.

In the survey data, the rate of employment in the control group is quite similar to the rate in the administrative data (Table 5, Column 1). Only in the different types of employment the extensive and intensive margin in the survey data is higher than

[12] 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.

in the administrative data. The differences in employment between the treatment and control groups are smaller in the survey than in the administrative data and are not statistically significant. However, the signs of the coefficients are in the same direction in the two datasets with the sole exception being months of apprenticeship. The lower number of risk characteristics in the survey sample relative to the administrative data sample might explain the differences in the level of employment and the size of the treatment effects (see Appendix A7). Estimation models including predicted probabilities of participating in all interviews calculated by logit models support this suggestion. This inverse probability weighting leads to very similar results as the ones obtained from the administrative data (see Appendix A9).

In line with the reported higher employment, fewer mothers than indicated by the administrative data state to receive welfare in the control group. However, 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 10.2 percentage points, which is even higher than in the administrative data.

The last five rows in Table 5 contain information which is only measured through the telephone surveys including the occurrence of a second pregnancy, inconsistent use of contraceptives, constant partnership, childcare use, and school attendance. These five outcomes can help to identify channels why the intervention has the unintended effects on employment, welfare use and fertility which we observe in the administrative data.

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 become 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 do not affected second pregnancies, they also do not

Table 5: Maternal Life Course Outcomes 36 Months after the 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	0.555	-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)				
Childcare Utilization	0.584	0.083	0.144	7.175	1.894*	0.071
	[0.495]	(0.056)		[8.571]	(1.046)	
School	0.102	-0.014	0.681	0.934	0.072	0.879
	[0.304]	(0.025)		[3.877]	(0.331)	
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

affected inconsistent use of contraceptives.[13]

The next three rows examine partner stability, childcare use and school attendance. Partner stability presents the percentage of women who stayed with the same partner from pregnancy until the third birthday of the treatment child. The treatment does not change the rate of mothers in a stable partnership 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. Home visiting might influence childcare use because the home visitor supports the treated mothers in the childcare application process. This improved childcare access might positively influence a mother's fertility decision if she perceives external childcare as a relief of

[13]The question for use of contraceptives is asked in three interviews at 15, 27 and 36 months. A mother uses contraceptives inconsistently if she states in one interview that she does not always use a contraceptive method. Mothers who are sexual inactive, pregnant or trying to get pregnant are excluded from the sample.

strain. The data shows that the intervention slightly increased the average months in childcare.[14] Nevertheless, the effect does not appear strong enough to explain the intervention's impact on total fertility. Finally, school attendance is an indicator that could explain the lower employment rate in the treatment group in addition to the higher birth rate. 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.[15]

Overall, the results of the survey data confirm the findings from the administrative data that the intervention increases second births and welfare dependency. The results on employment are not as strong as in the administrative data, which might be explained by the survey participants having fewer risk characteristics. Investigating the channels for the results indicates that a change in second pregnancy outcomes most likely explains the increase in second births, while partner stability, school attendance and childcare use are unlikely to be the explanation. This finding is again in contrast to the results from the U.S. where the intervention in all three trials reduced not only second births, but also second pregnancies.

The analyses only included mothers who participated in all interviews. Although there are no differences between treatment and control group baseline characteristics in this sample, the results require careful interpretation because the survey sample does not have the same characteristics as the baseline sample. Therefore, in the next section, I 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 only includes mothers who participated in all interviews. Presumably, the rate is lower because some mothers only participate in one interview after birth, which is most likely before a further

[14] Childcare utilization is a broad measure of whether a child attends childcare. It does not include hours or quality of childcare. Childcare for welfare receiving mothers is usually completely financed by the state in Germany.
[15] Enrollment in higher education is of negligible relevance in the treatment and control groups.

pregnancy occurred. The rate of second pregnancies in the treatment group is 5.5 percentage points higher, but the difference is statically not significant at the ten percent level, thereby confirming the results from the analyses with the mothers who participated in all interviews. Altogether, 175 second pregnancies occurred among the mothers who participated at least in one interview after birth.[16]

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 Outcome			
	Control Mean	Treatment Mean	Overall Mean
Life Birth	0.527	0.634	0.589
Abortion	0.243	0.149	0.189
Misscarriage	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. In Panel B all pregnancies from Panel A. Panel C is a multinomial logit estimation with Life Birth as baseline category. TG = Treatment Group; CG = Control Group;
* p < 0.1, ** p < 0.05, *** p < 0.01

Panel B investigates the outcomes of these 175 second pregnancies, which can be live birth, abortion, miscarriage and unobserved pregnancy outcome. Along with the results of the previous sections, the results of Panel B reveal that the percentage of pregnancies that led to a live birth is 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 is only slightly higher in the treatment group.

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

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

the 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 is 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 not significant. Finally, the probability of not observing the outcome of the pregnancy relative to that of a live birth is 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.

Placing the rate of abortions in the *Pro Kind* program in relation to the abortion rates in average populations helps to interpret the abortion results. From 2008 to 2011, there were approximately 16 abortions per 100 live births in the overall German population.[17] 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 data on the population average indicate 27 abortions per 100 live births (Statistisches Bundesamt, 2014). The control group of the *Pro Kind* sample has a ratio of 46 abortions per 100 live births, whereas in the treatment group the ratio is 23 to 100, which is close to the population average and lower than average for unmarried mothers. This might indicate that mothers in the treatment group are as confident in their ability to raise a second child as average mothers.

Despite the finding that a lower percentage of pregnancies in the treatment group ended in an abortion, 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 planned a second birth and who are able to meet the challenges of another child give birth to a second child. An 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 whether appropriate family planning increased, I analyze the life situations of the *Pro Kind* second time mothers in treatment and control

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

groups and compare them with SOEP second time mothers.

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

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Control		Treatment		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 base on z-statistic of a two-group test of proportions. The presented data contains all second children for who data is available. *Age of the Sec. Child in Mo.* gives the age of the second child at the time of the interview. C=Control Group; T=Treatment Group.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7 includes data from the interview after the birth of the second child on 97 of the 103 second children. The first two rows present responses to questions concerning whether the child was unplanned or 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, the other characteristics, such as "no partner" or "father does not live in the household", occurred more often in the treatment group. Although none of the differences is statistically significant, the results may indicate that mothers with fewer resources had a second child in the treatment group and that these mothers were less responsible with respect to family planning compared with the mothers in the control group. The fact that only 39% of the mothers planned their second pregnancy and 48% have 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 state that the pregnancy that led to a second child was planned, and only 8% state that they are alone responsible for the child.

Now the question is why the mothers in the treatment group decide to have another child despite not having planned a second birth and seeming to be unable to

meet the challenges of having another child. As an explanation, the home visitor might directly influence the decision of the pregnant mother. There are no recommendations concerning abortions in the *Pro Kind* or the 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-depths interviews that abortion was almost no topic in the nurse supervision, they also stated that a nurse or midwife would hardly give the advice to abort a pregnancy to a client. However, mothers in treatment and control groups also get the encouragement to keep the baby from other sources because German law permits abortions only if the woman received consultation from a family counseling office. Therefore, in addition to the direct advice of the home visitors and the family counseling office, it is likely that also others channels are important for the decision of the mother.

In the literature on further parity progression, life satisfaction and well-being play important roles. As Margolis and Myrskylä (2015) show, a decline in life satisfaction during transition to parenthood reduces the probability of subsequent births. Therefore, a potential reason why the *Pro Kind* program increased fertility is increased satisfaction with their lives and their maternal role. This higher satisfaction might have resulted from more positive experiences and greater attachment with the first child due to enhanced maternal skills in the treatment group. To test this hypothesis, the next section investigates whether the *Pro Kind* intervention influences reports of maternal life satisfaction and well-being.

7.2.2 Life Satisfaction and Well-being

I begin the analysis with a descriptive overview of the treatment and control groups' outcomes and the SOEP data for first-time mothers. These outcomes were obtained at the interview 27 months after the birth of the treatment child. Appendix Table A10 shows that on 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. The mothers in the treatment group reported feeling sad, angry, or worried less often and happy more often. Compared with the first-time mothers from the SOEP sample, the

mothers in the *Pro Kind* treatment group also have higher well-being and are more satisfied in most categories. Table 9 shows that the differences between the control and treatment groups are significant at the 10% level for the well-being index, which captures satisfaction with life in a variety of specific areas and in general.[18] The standardized effect sizes are meaningful, with values near 0.15 SD.

Table 8: Well-Being and Satisfaction with Life

	(1)	(2)	(3)	(4)	(5)	(6)
	Index of Well-Being		Index of Life Satisfaction 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*
R^2	*0.02*	*0.18*	*0.01*	*0.26*	*0.01*	*0.18*

Notes: Standard errors (in parentheses). Well-Being is an index of less often sad, angry, 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, child care availability and family life. All dependent variables are standardized with mean of zero. Controls include extended baseline variables, community fixed effects and age of the treatment child. Measurement is in average at 28 months after birth of the treatment child. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

After showing that the *Pro Kind* program increased maternal life satisfaction and well-being, the investigation sought to determine whether these subjective measures are related to fertility decisions. Mothers who abort their pregnancy in the *Pro Kind* sample have a general satisfaction value of 5.74 which is lower than the general satisfaction of mothers with and without a second birth. Although it is unclear whether low life satisfaction caused the abortions or the abortions led to low life satisfaction, this association provides a first indication that low life satisfaction is correlated with abortions. Further evidence that the greater life satisfaction in the treatment group is related to fertility comes from comparing the mothers who gave birth to a second child in the treatment group with those in the control group. Their life satisfaction levels differed significantly, with a value of 7.61 in the treatment group and 6.42 in the control group (T=-3.06; nTG=60; nCG=33). It is possible that the birth of the second child caused this increase in happiness. However, it is likely that greater life satisfaction was also influenced by better experiences with the first child and that, as a result, the mothers were already happier before they

[18] Well-being is based on an index indicating how often one is happy versus sad, angry, or worried. Life satisfaction in different areas is based on an index of eight questions related to satisfaction with health, housework, household income, personal income, place of dwelling, free time, childcare availability and family life.

became pregnant a second time. If this is the case, this higher level of happiness could be an explanation for the lower rate of abortions in the treatment group.

8 Comparison to U.S. Results

In this section, I investigate why the effects on maternal outcomes differ so substantially from the results in the U.S. studies, mainly from the NFP. First, I discuss whether program implementation, participants' characteristics, or contraceptive and abortion availability can explain the different outcomes. Second, I examine how the welfare state arrangements for mothers with small children in the two countries influence the program outcomes.

The concern that different program implementation leads to the different outcomes is causeless for several reasons. First, the *Pro Kind* home visitors used the same materials and guidebooks, translated into German, during their home visits as the NFP home visitors. Second, 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 A3). Third, the average number of conducted home visits in the *Pro Kind* program ($\bar{x} = 33$) is close to the average number of home visits in the NFP trials (Elmira: $\bar{x} = 32$, Memphis: $\bar{x} = 33$, Denver: $\bar{x} = 27.5$). Finally, the similar average number of home visits leads to very similar program costs. As discussed above, the average intervention costs in the NFP Memphis trial amounted to \$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). All of these aspects, demonstrate that the *Pro Kind* program was comparable in material, content and intensity to the NFP. Therefore, it is unlikely that a difference in the implementation of the *Pro Kind* program can explain the different effects from the U.S. studies.

Analyzing the *Pro Kind* sample characteristics reveals that it has similar characteristics with respect to marriage status, years of education and poverty level in comparison to the populations in the NFP randomized trials. Only the average age

of the participants appears slightly younger. At affiliation the participants in the NFP trials were, on average, between 18 and 19 years old in comparison with 21 years in the *Pro Kind* study (Olds et al., 1997; Kitzman et al., 1997; Olds et al., 2002). 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.

Contraception availability and abortion rules in the two countries might also explain the different outcomes. However, the institutional settings are comparable in both countries. 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 bared by health insurances, whereas abortions following a consultation are to be paid privately. Although abortion laws are relatively 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). All contraceptives are generally available for purchase in Germany and the U.S. and knowledge about contraception and awareness of different forms of contraception is almost complete in both countries (Johnson and Pion, 2013).

After showing that the participants' characteristics, program implementation or contraceptive and abortion availability are comparable in the two countries, it is most compelling that the differences in the arrangement of the welfare systems explain the opposite results. In Germany, social assistance is means-tested and increases with parity. The benefits are withdrawn at a rate of almost 100%, so Blundell et al. (2009) show as an example that the budget line for a low-wage lone mother with two children is hardly affected by her working hours.[19] Additionally, there are no work obligations or benefit cuts until the child's third birthday. In

[19] 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 with full-time employment €1,200 after deduction of health insurance (childcare is generally free for low-income mothers in Germany). Thus, the single mother has to work full-time and has to earn an hourly wage of €9.5 to meet the reservation wage.

contrast, in the U.S., welfare programs include work obligations, in-kind transfers and family caps, limiting either in part or completely any additional benefit for having a subsequent child while receiving welfare benefits. In addition, (Meyer and Rosenbaum, 2001) demonstrate that between 1984 and 1996, changes in tax and transfer programs, such as the EITC, sharply increased the incentive for low income mothers and single mothers to work.

In a series of papers, Kearney and Levine (2012, 2014) examined reasons for teen childbearing across developed countries and across U.S. states. Their approaches concentrate on teens, but they are also helpful to answer why home visiting has different effects in Germany and the U.S. because approximately 50% of the participants in all home visiting studies are teenagers. In addition the findings for teens can be more generally applied to first-time mothers of low socioeconomic status (SES). Although welfare regulations are more generous in Germany, Kearney and Levine (2012) showed that U.S. teens are four times as likely to give birth compared to teens in Germany. Kearney and Levine (2014) explain a sizable share of the geographic variation in teen childbearing with variation in income inequality across U.S. states and developed countries. In regions with high income inequality, girls on the bottom of the distribution choose immediate utility enhancing motherhood at a young age instead of investing in their own economic progress because they perceive that they have little chance of advancement. This setting can also help to explain the decision for a second child in combination with home visiting.

Kearney and Levine (2014) formalize their idea in a framework, which I adapt to the decision of a low SES mother to have a second child. An individual chooses to have a subsequent baby in the current period if the following condition is met:

$$u_0^b + V^{low} > u_0^d + qV^{high} + (1-q)V^{low} \tag{2}$$

where u_0^b is the current period utility if she has a subsequent baby and u_0^d is the current period utility if she delays additional childbearing. I assume that for young women of low SES having another baby is utility-enhancing in the current period, such that $u^b > u^d$.[20] V^{low} presents the discounted value of the young mother's future

[20] If $u^b < u^d$, it is never optimal to have a second baby in the current period and the model trivially predicts

utility stream with a second baby. The future utility is determined on V^{low} because having a second baby reduces her flexibility in the labor market and thereby the probability to leave welfare, decreasing the future stream of her own earnings. If the young woman delays further childbearing, there is some positive probability that she will achieve the "high" utility position V^{high} in future periods. The young woman does not perfectly observe this probability. Instead, she bases her decision on her subjective perception of the probability q. Rearranging Expression 2, we see that a young woman will choose to delay additional childbearing in the current period if and only if:

$$q(V^{high} - V^{low}) > u_0^b - u_0^d \tag{3}$$

A young woman will choice to delay second childbearing if the interaction between her perceived chance of economic success q and the size of economic improvement $(V^{high} - V^{low})$ due to delay of further child bearing is greater than her immediate benefit from additional childbearing $(u_0^b - u_0^d)$.

Now, I add home visiting to the framework and assume that home visiting increases u_0^b because home-visited mothers are more satisfied with the maternal role and enjoy more positive experiences with their first child. The effects of the *Pro Kind* program on maternal life satisfaction support this assumption. Second, I assume that home visiting increases q because home visiting aims to improve maternal personal strengths, such as locus of control, self-efficacy and problem-solving abilities. Therefore, home visiting gives the mother the sense of mastering, which leads to a higher perception that she can reach V^{high}. The positive NFP effects on maternal sense of mastery (Olds et al., 2004) and the effects on employment support this assumption. I further assume that home visiting increases u_0^b and q to the same effect in Germany and the U.S.

We observe that home visiting increases the right term of Expression 3 in the U.S. and in Germany. In contrast, in the U.S., an increase in q has a greater impact on fertility decision than in Germany because the gap between V^{high} and V^{low} is greater

"delay" to the optimizing choice. It seems reasonable to expect that for the majority of high-SES young women, $u^b < u^d$, and the results of the empirical analysis are consistent with that supposition.

in the U.S. than in Germany. The gap is larger because income from welfare V^{low} is much higher in Germany and the net benefit of labor market participation V^{high} is lower in Germany due to the absence of in-work benefit policies. In the U.S., the effect on the perceived higher change in advancement q has a strong effect on birth decision, whereas this effect is rather small in Germany. Therefore, the increase of immediate utility from a further child predominates and leads to a fertility increase and employment decrease effect of home visiting in Germany, whereas the opposite can be observed in the U.S.

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 uses 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 are 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.

A randomized experiment is 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 use 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 rely on survey data that suffered from survey non-response. Nevertheless, a comparison of the baseline characteristics for the treatment and control groups indicates that this attrition was not selective. Therefore, it is unlikely that the sample attrition resulted in problems with the validity of the results.

Previous studies that examined the effect of home visiting on the maternal life

course were performed 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. The findings on welfare dependency and employment decisions are of high fiscal relevance since disadvantaged mothers show strong welfare persistence and receive a substantial amount of total welfare spending in many countries.[21]

Furthermore, the results of this study can increase our understanding of the mechanisms through which early childhood interventions operate. On the one hand, it is likely that improved maternal life satisfaction and well-being can contribute to positive effects on child outcomes.[22] On the other hand, especially in disadvantaged populations, shorter birth spacing can have a negative effect on child development that counteracts the positive effects (Buckles and Munnich, 2012). Which of the two effects will predominate requires further investigation. For the *Pro Kind* program, it seems that the positive effects of the intervention on child development are diminished by the increased fertility. Sandner and Jungmann (2015) show that while the treatment increases cognitive development at ages 6 and 12 months for girls, these effects vanish at age 24 months.

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

[22] For example, Berger and Spiess (2011) and Petterson and Albers (2001) show that higher maternal life satisfaction and fewer depression syndromes are related to improved child cognitive development.

References

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 SOEP-survey and SOEPservice - The (Near) Future of the German Socio-Economic Panel Study (SOEP). *SOEPpapers 155*, 1–35.

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.

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.

Blank, R. M. (2002). Evaluating Welfare Reform in the United States. *Journal of Economic Literature 40*(4), 1105–1166.

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 (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.

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

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

Gebhardt, T. and H. Jacobs (1997). American Conditions? - Social Assistance in the USA and Germany : a Comparison From a Historical, Institutional and Legal Perspective. *Zeitschrift für Sozialreform 43*(8), 597–634.

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., 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.

Johnson, S. and C. Pion (2013). Current Methods and Attitudes of Women Towards Contraception in Europe and America. *Reproductive Health 10*(7), 1–9.

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. (2004). Is There an Effect of Incremental Welfare Benefits on Fertility Behavior? A Look at the Family Cap. *Journal of Human Resources 39*(2), 295–325.

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.

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.

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.

Margolis, R. and M. Myrskylä (2015). Parental Well-being Surrounding First Birth as a Determinant of Further Parity Progression. *Demography 52*(4), 1147–1166.

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.

Moffitt, R. (1998). The Effect of Welfare on Marriage and Fertility: What Do we Know and What Do We Need to Know? In R. Moffitt (Ed.), *Welfare, the Family and Reproductive Behaviour: Research Perspectives*. Washington D.C.: National Academy Press.

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.

Petterson, S. and A. B. Albers (2001). Effects of Poverty and Maternal Depression on Early Child Development. *Child Development 6*(72), 1794–1813.

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

Sandner, M. and T. Jungmann (2015). Gender-Specific Effects of Early Childhood Intervention: Evidence from a Randomized Controlled Trial. *mimeo*.

Schmieder, J. F., T. v. 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.

Siedler, T., J. Schupp, C. K. Spieß, and G. G. Wagner (2009). The German Socio-Economic Panel as Reference Data Set. *Schmollers Jahrbuch 129*(2), 367–374.

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

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

U.S. Department of Health and Human Services (February 19, 2015). HHS Awards $386 Million to Support Families Through the Home Visiting Program.

Appendix A

See Tables A1-A10 and Figures A1-A2.

Table A1: NFP Results Elmira, Memphis and Denver

NFP Results Elmira

Outcome	Observation Period		
	6 Months	**4 Years**	**15 Years**
School:	More School Enrollment 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: 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 A4: 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
Migration	Binary	1 if Participant is not Born in Germany or 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
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 A5: 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	755

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)
	703	602	539	438	296

Notes: Robust standard errors are shown in parentheses. Dependent variables shown in column (1). The treatment indicator has the value one if the mother is in the treatment group. Column (2) contains estimates of the average difference in characteristics between mothers in the control and treatment group including community fixed effects for the participants merged with the administrative data. Column (3)-(6) contain these estimates for the survey data. See Appendix Tables A4 and A5 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)
Migration	-0.190***	0.0484	0.0410	0.0350	0.0761**
	(0.0509)	(0.0323)	(0.0287)	(0.0263)	(0.0265)
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)
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)
	755	755	755	755	755

Notes: Robust standard errors are shown in parentheses. Dependent variables shown in column (1). The treatment indicator has the value one if the mother is merged or participated in the intervviews. Column (2) contains estimates of the average difference in characteristics between attriting and non-attriting mothers including community fixed effects for the participants merged with the administrative data. Column (3)-(6) contain these estimates for the survey data. See Appendix Tables A4 and A5 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)
Any Employment	0.528	-0.092***	0.008	6.526	-1.598***	0.007
	[0.499]	(0.027)		[9.180]	(0.476)	
Parttime/Fulltime Employed	0.194	-0.058**	0.033	1.689	-0.636*	0.099
	[0.396]	(0.023)		[4.902]	(0.351)	
Apprenticeship	0.207	-0.016	0.606	2.426	-0.294	0.555
	[0.406]	(0.030)		[6.078]	(0.482)	
Marginal employment	0.304	-0.054*	0.093	2.328	-0.651**	0.048
	[0.461]	(0.029)		[5.142]	(0.289)	
Welfare	0.964	0.013	0.235	32.78	1.223	0.238
	[0.295]	(0.010)		[11.91]	(0.975)	
Observations	*329*	*684*		329	684	
Second Child in HH	0.187	0.065**	0.032			
	[0.363]	(0.026)				
Observations	*316*	*663*				

Notes: Robust standard errors 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 from affiliation to 36 months after birth. Estimations include community fixed effects and controls for age and being underaged. TG = Treatment Group; CG = Control Group; HH=Household.

Table A9: 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)	
Parttime/Fulltime Employed	0.303	-0.028	0.628	2.486	-0.643	0.314
	[0.460]	(0.058)		[5.301]	(0.638)	
Apprenticeship	0.304	-0.076	0.211	3.256	0.029	0.975
	[0.461]	(0.060)		[6.360]	(0.941)	
Marginal employment	0.262	-0.045	0.419	2.570	-0.801	0.210
	[0.440]	(0.055)		[5.625]	(0.638)	
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)				
School	0.120	-0.016	0.714	1.081	0.468	0.537
	[0.326]	(0.044)		[4.087]	(0.758)	
Childcare	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. Estimations include community fixed effects and baseline controls. TG = Treatment Group; CG = Control Group; HH=Household.

Table A10: Descriptive Statistics for Well-Being and Life-Satisfaction

	Control Group			Treatment Group			SOEP		
	Mean	sd	n	Mean	sd	n	Mean	sd	n
How Often or Seldom Have You Experienced this Feeling in the Last Four Weeks?									
Angry	3.05	1.00	195	2.91	1.09	239	3.11	0.92	498
Worried	2.09	1.04	194	1.77	0.94	238	1.98	0.93	498
Happy	3.66	0.90	195	3.76	0.88	237	3.73	0.86	498
Sad	2.71	1.07	195	2.49	1.03	237	2.62	1.04	498
How Satisfied are you Today with the Following Areas of Your Life?									
Health	6.55	2.97	194	6.83	2.88	235	7.34	1.89	726
Housework	6.92	2.33	193	7.37	2.32	231	6.62	1.78	671
Household Income	4.92	2.70	193	5.58	2.89	235	6.16	2.28	726
Personal Income	4.14	2.90	191	4.57	3.05	233	4.84	2.86	686
Place of Dwelling	6.56	3.16	194	6.63	3.12	235	7.60	2.13	726
Free Time	5.67	2.91	195	6.23	2.87	234	6.22	2.25	727
Child Care Availability	6.73	3.01	192	6.68	3.33	228	6.75	2.62	630
Family Life	7.46	2.35	195	7.63	2.52	234	7.86	1.76	602
Life in General	7.13	2.10	195	7.44	1.91	237	7.33	1.54	727

Notes: For the outcomes in the first four rows the scale is: 1=Very Rarely, 2=Rarely, 3=Occasionally, 4=Often, 5=Very Often. For the other outcomes the scale is: 0=totally unhappy to 10=totally happy. SOEP includes mothers whose first child has an age between two and three years. The average age of the first child in the *Pro Kind* sample is 30.06 months. sd=standard deviation.