UNIVERSITY of York

This is a repository copy of *Evaluating the effects of a targeted home visiting program on maternal and child health outcomes*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/127611/

Version: Accepted Version

Article:
Sandner, Malte, Cornelissen, Thomas orcid.org/0000-0001-8259-5105, Jungmann, Tanja et al. (1 more author) (2018) Evaluating the effects of a targeted home visiting program on maternal and child health outcomes. Journal of Health Economics. pp. 269-283. ISSN 0167-6296

https://doi.org/10.1016/j.jhealeco.2018.02.008

Reuse
This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: https://creativecommons.org/licenses/

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

White Rose university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Evaluating the effects of a targeted home visiting program on maternal and child health outcomes

Malte Sandner[§], Thomas Cornelissen*, Tanja Jungmann[†], Peggy Herrmann[‡]

[§]Institute for Employment Research (IAB)

*University of York, CReAM, and IZA

[†]University of Rostock

[‡] Hannover Medical School

November 2017

Abstract

We evaluate the effects of home visiting targeted towards disadvantaged first-time mothers on maternal and child health outcomes. Our analysis exploits a randomized controlled trial and combines rich longitudinal survey data with unique administrative health data. In a context in which the target group has comprehensive health care access, we find no effects of home visiting on most types of health utilization, health behaviors, and physical health measures. However, the intervention has a positive effect on some maternal mental health outcomes, reducing depressions reported in the survey data by eleven percentage points and prescriptions of psycholeptics recorded in the administrative data by seven percentage points.

JEL-Classification: I14

Keywords: Child Health; Health Inequality; Mental Health; Early Childhood Intervention

Contact details: Malte Sandner, Institute for Employment Research (IAB), Regensburger Straße 104, 90478 Nürnberg, Germany (email: malte.sandner@iab.de); Thomas Cornelissen (corresponding author), Department of Economics, University of York, Heslington, York YO10 5DD, United Kingdom (email: thomas.cornelissen@york.ac.uk); Tanja Jungmann, Institute for Special Needs Education and Rehabilitation, University of Rostock, August-Bebel-Str. 28, 18051 Rostock, Germany (email: tanja.jungmann@uni-rostock.de); Peggy Herrmann, Department of Conservative Dentistry, Periodontology and Preventive Dentistry, Hannover Medical School, Carl-Neuberg-Str. 1, 30625 Hannover, Germany (email: Herrmann.Peggy@mh-hannover.de)

Acknowlegments: We are indebted to the AOK Lower Saxony, AOK Bremen and AOK Saxony as well as the Kassenzahnärztliche Vereinigung Lower Saxony and Bremen for providing the data and support with data management. Financial support by the German Federal Ministry for Family, Seniors, Women and Youth (BMFSFJ) (Grant no. IIA6-25080820V6), the State Ministry of Social Affairs and Consumer Protection, Saxony, the TUI Foundation, and the Dürr and Reimann-Dubbers Foundations is gratefully acknowledged. We also would like to thank Gabriella Conti, Christian Dustmann, Tobias König, and Uta Schönberg for helpful comments and all employees of the *Pro Kind* foundation and all members of the *Pro Kind* research group for support, in particular Peter Lutz, Anja Becher and Matthias Schock. This study is registered in the AEA RCT Registry and the unique identifying number is: "AEARCTR-0002135".

1. Introduction

It is well established that children from disadvantaged families suffer more ill-health than children from advantaged backgrounds (e.g. Case et al. 2002; Currie and Stabile 2003; Currie and Lin 2007; Condliffe and Link 2008) and that there is a pronounced link between family income and child health (Currie and Lin 2007; Hoynes et al. 2015; Kuehnle 2014; Milligan and Stabile 2011). The child health gradient in socio-economic status (SES) persists into adulthood (Apouey and Geoffard 2013; Case et al. 2002) and contributes to well-documented income-related health inequalities later in life (e.g., Deaton 2002; van Doorslaer and Koolman 2004). The gradient may indicate under-investment into child health by low-SES parents due to insufficient knowledge about children's health needs (Currie 2000). As child health is related to educational achievement and labor market outcomes, the under-investment is likely to harm economic success later in life (e.g. Currie 2009; Currie et al. 2010; Figlio et al. 2014). Moreover, due to positive externalities of child health (Currie 2000), the costs to society may exceed the private costs. For these reasons, policies that attempt to intervene early in life in order to close the SES gradient in child health have received much attention among policy makers.

A policy that intervenes particularly early to address health inequalities is home visiting targeted towards disadvantaged families. In such programs, specially trained professionals deliver home visits from pregnancy throughout the first years of children's lives. By providing information, instructions and support to the parents, and in particular to the mothers, home visiting programs try to promote and encourage appropriate acute and preventive health care, to improve home safety to reduce preventable injuries, and to reduce adverse maternal health behaviors, such as smoking. Several countries have recently expanded home visiting programs for disadvantaged families. In the U.S., the Obama administration 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). In the U.K., home visiting programs have been expanded since 2005 delivering services to 16,000 disadvantaged new parents each year (U.K. Department of Health 2013) and in Germany the federal and local governments spent 102 million Euro each year since 2012 to expand home visiting programs (BMFSFJ 2015). Given the substantial cost of these programs, it is of vital interest to know how effective they are in reaching their goals.

Our study helps to answer this question by exploiting a randomized controlled trial (RCT) to examine the health effects of an intensive targeted home visiting program (*Pro Kind*) in Germany, thus contributing to the existing literature in several ways.[1] Previous evaluations of home visiting

[1] Previous literature has investigated effects of the *Pro Kind* program on child development (Sandner and Jungmann 2017), maternal life course outcomes (Sandner 2017) and maternal competencies (Sierau et al. 2016). Results are discussed in section 2. One existing publication in German language (Jungmann et al. 2010) has assessed the effects of

programs have often relied exclusively on survey-based maternal self-reports about health and health utilization. If they had access to administrative data this was mostly restricted to emergency room visits or birth records (Doyle et al., 2016; and the studies surveyed in Avellar and Supplee, 2013; Peacock et al., 2013). In contrast, we have the rare opportunity to combine survey data with more detailed administrative data than previous studies, spanning the period from pregnancy up to two years after birth for over one third of the baseline sample. The survey data include, among others, detailed maternal reports on health behavior, such as breastfeeding and smoking, and on child health, while the administrative data is drawn from insurance records of the German public health insurance system and includes information on all prescription medications, on hospital contacts including diagnoses codes, as well as on midwife contacts and dentist visits. The data allow drilling down to the level of individual health conditions based on diagnoses and prescription codes, providing a more objective and complete measurement of health conditions than is possible based on survey data alone. This wealth of information provides us with the opportunity to analyze more detailed sub-domains of health than any of the previous studies in this area. In addition to child health outcomes, which most studies focus on, we also investigate maternal physical and mental health.

Since we investigate many outcomes, we adjust our results for multiple hypothesis testing (MHT) throughout the analysis and we further check the robustness of our results to different statistical approaches, including inverse probability weighting (IPW) and permutation tests. Overall, we find no clear effects of the intervention on child physical or mental health or on maternal health behavior during the first two years of life. However, we find that the intervention had a positive effect on maternal mental health as evidenced by both, a reduction in prescriptions of psycholeptics recorded in the administrative data by seven percentage points, and reductions in depressions reported in the survey data by eleven percentage points. We further find effects on the utilization of oral health care, explicitly targeted by the intervention, with positive program effects on maternal dental checkups during pregnancy and a reduction of therapeutic dental utilization for children. Both findings are highly relevant, given that maternal depression and stress are related to many adverse child outcomes (Junge et al. 2016; Aizer et al. 2016; Carlson 2015), and oral infections during pregnancy have been linked to preterm delivery (Sanz and Kornmann 2013; Vergnes and Sixou 2007) and childhood caries causes enormous suffering for the affected children and high costs for the society (Casamassimo et al., 2009). Yet, these outcomes have rarely been included in previous evaluations of the health effects of home visiting.[2] While these findings imply that the program has beneficial effects in some limited areas, the overall results suggest that the

Pro Kind on smoking one year after birth and on birth outcomes using preliminary *Pro Kind* survey data only.

[2] Exceptions are the studies by Doyle et al. (2017), Johnston et al. (2004), Mitchell-Herzfeld (2005) Duggan et al. (2004), and Koniak-Griffin et al (2002). While these studies find mixed effects, they rely on survey measures and do not use administrative data for exploring effects on mental health. Dental outcomes are less often included in randomized trials and in no study administrative data is available.

program fails to affect several important domains that are explicitly targeted, such as maternal smoking behavior, breastfeeding, and prevention of accidents.

Our finding of little effects on most types of utilization and on physical health is in line with two recent studies reporting either no effect (Robling et al. 2016, for the UK) or very limited positive effects (Doyle et al. 2015, for Ireland) of similar programs. Our stronger effects on mental health suggest that the mental health needs of deprived mothers in our context are not fully covered by the regular health care provision, but can be addressed to some extent by home visiting. We provide some evidence to support the idea that the beneficial mental health effects are driven by the personal relationship between mother and home visitor that substitutes for a lack of social support of the target group that experiences stressful life events. Finally, our pattern of results sheds some light on the question which types of health outcomes can be influenced by home visiting programs. It suggests that targeted home visiting is most effective for health behaviors in which the SES gradient is steep, to which the program transmits additional information, and which are not too costly to change.

The remainder of the paper is organized as follows: Section 2 presents the setting of the intervention. Section 3 discusses the data and randomization procedure. Section 4 presents the estimation strategy. The main results are shown in Section 5 and discussed alongside some further analysis in Section 6. Section 7 concludes.

2. The *Pro Kind* Program

In the U.S., home visiting, and in particular the Nurse Family Partnership (NFP) program, has been found to be beneficial for maternal and child health. For example, the NFP program reduced days hospitalized for injuries/ingestions and improved maternal health behavior, e.g. less cigarette smoking (for an overview see Olds 2006 and for an economic reanalysis Heckman et al. 2017). In order to investigate how transferable the results from the U.S. are, the German federal government founded the *Pro Kind* pilot project, which represents the first randomized controlled trial to investigate the effectiveness and efficacy of a home visiting program in Germany. The *Pro Kind* program is an adaptation of the NFP program, which provides instructions for home visit frequency, employee selection, teaching material, and guidebooks (see Jungmann et al. 2009; Sierau et al. 2016; Olds 2006) for more information about the *Pro Kind* project and NFP.

The *Pro Kind* intervention started between the 12th and 28th week of pregnancy, extending up to the child's second birthday. The majority of participants (about 60%) were subject to a continuous visitor model in which medical professionals (almost exclusively midwifes, with the exception of one pediatric nurse) conduct the home visits. In a subset of municipalities, covering about 40% of the participants, a tandem model was implemented in which the medical professional

was complemented by a social pedagogue. The medical professional conducted most of the visits up to 6 months after birth, and the social pedagogue conducted most of the visits thereafter (Brand and Jungmann 2012). The frequency of the home visits varied according to the NFP model prescription between weekly, biweekly, and monthly visits, with the highest frequency directly before and after birth. Overall, 52 home visits with an average duration of 90 minutes were scheduled between pregnancy and the child's second birthday. Teaching materials and visit-by-visit guidelines structured the theme and aim of each home visit. Nevertheless, home visitors had the flexibility to adapt the contents to maternal needs and the familial situation. The teaching materials, guidelines, and aims were adapted from NFP in which they were proven successful. All home visitors regularly received feedback, encouragement, reflection, and support from nurse supervisors. Process evaluation data shows that the *Pro Kind* intervention was well implemented. On average families received 32.7 home visits (SD = 19, range: 0-94) and 45.3 home visits when they stayed in the program until the child's second birthday. The median number of home visits is 36, with mothers below the median receiving 17 home visits, and mothers above the median receiving 49 home visits on average.

One of the major goals of *Pro Kind* is the improvement of maternal and child health. In order to achieve these goals the project aims at

(I) increasing utilization of immunizations and preventive care (prenatal checkups, dental checkups and oral health care, post-natal checkup for the child),

(II) improving maternal health behavior, e.g. healthy diet, breastfeeding, reduction of maternal smoking, child passive smoking, alcohol and substance use,

(III) increasing maternal self-efficacy to improve her mental health and well-being,

(IV) ensuring that mothers interact safely with the child and provide a safe environment.

The last topic was the most dominant in the *Pro Kind* guidelines with 18 home visits scheduled to cover safety aspects. In 11 home visits the focus was on maternal health behavior with five home visits focusing on smoking reduction and five visits focusing on a healthy diet. Materials of four visits included utilization of immunization and other preventive services, and another four home visits were devoted to oral health care. This topic was particularly highlighted by the fact that participants received a box with oral health care products, such as tooth brushes, and home visitors received an extra training session on oral health from a dentist (see Meyer et al. 2014 for a description of the oral health module). Health and health-related topics dominated the implementation of the *Pro Kind* program. In self-ratings of the devoted time for each home visit, the home visitors stated that they invested 50 percent of their time with the family during pregnancy to the domains Maternal Health, Environmental Health and Social and Health Services. This share slightly decreased to 38 percent in the first year of life and to 34 percent in the second year of life

(Brand et al. 2013).

In order to recruit families who are most likely to benefit from the *Pro Kind* intervention, the program was restricted to financially and socially disadvantaged first-time mothers with at least a basic command of the German language. Financial disadvantage is mainly defined as receipt of social welfare benefits while social risk factors included, for example, low education, teenage pregnancy, and health problems. Project partners, like gynecologists, job centers, pregnancy information centers, and youth welfare offices, referred about 75 percent of the participants to *Pro Kind*. About 25 percent self-registered into the program, which was advertised, for example, in gynecologist's offices and pregnancy information centers.

The *Pro Kind* project was implemented between 2006 and 2012 in 13 municipalities of three German federal states covering rural and urban regions as well as regions in Eastern and Western Germany. This mixture of sites ensures that the program was implemented under varying regional conditions in terms of the availability of child care, health care provision, and labor market conditions.

Besides health, the intervention also aimed to improve maternal parenting, child development and the development of the maternal life course. Previous literature about the outcomes of the *Pro Kind* program examined the effects of the intervention on child cognitive development, maternal skills, and maternal life course (Sandner and Jungmann 2017; Sandner 2017; Sierau et al. 2016). These articles found positive effects on child cognitive development concentrated on girls, higher maternal skills in some domains, more subsequent births, and lower maternal employment for mothers in the treatment group. A systematic examination of the effects of the intervention on child and maternal health is absent so far.

3. Data and Randomization

The *Pro Kind* program was implemented as a randomized controlled trial. 755 expectant mothers were enrolled into the trial and randomly assigned via a computer routine based on Efron´s biased coin approach (Efron, 1971) either to the treatment group (N=394) or to the control group (N=361). The random treatment assignment was stratified by municipality, maternal age ($<$ 18 vs. $\geq$ 18 years), and maternal nationality (German vs. non-German). 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 and to obtain the strata information. If a woman met the criteria, the supervisor visited the woman at her home. During this visit, the participant (or, if she was underage, her parents) signed an informed consent form for participation in the study, which included consent to allow access to their health insurance data for research purposes. Thereafter, participants completed a baseline questionnaire to assess

demographic and psychological characteristics. After answering the baseline questionnaire, women received the results of the randomization that assigned them either to the home visit or the control group. Women in both groups received an address list with support services in their communities, a small monetary incentive for participating in the study, feedback about the children's developmental status, and had access to the regular German health care services. Only expectant mothers in the treatment group received the described home visits.

We use administrative health insurance and survey data to measure the health and the health service utilization of mother and child and the health behavior of the mother. In Germany, health insurance is compulsory and covers all costs for preventive utilization (including dental examinations), prescriptions, hospitalizations, GP visits, and other outpatient services such as routine midwife visits before and after birth.[3] For the *Pro Kind* sample, which consists in large parts of welfare recipients, health insurance contributions are covered by the welfare office. Individuals can choose their health insurance company but competition between the health insurance companies is low because of almost identical charges and similar services. Additional private health care expenditure is very rare, in particular for this target group. Therefore, health service utilization recorded in the public health insurance data is a sensible proxy for health and preventive utilization in Germany.

Our main administrative dataset consists of health insurance records from the AOK, the largest public health insurance company that covers around 30 to 40 percent of the German population depending on the federal state.[4] We have access to this data for all three German federal states in which the *Pro Kind* trial was implemented (Lower Saxony, Bremen and Saxony). The AOK data include hospital utilization (date, days in hospital, DRG, ICD, costs), all medical prescriptions (date, ATC-code, costs) and utilization of midwifes before and after birth. We use further administrative data from the Association of Statutory Health Insurance Dentists (Kassenzahnärztliche Vereinigung, KZV) for Lower Saxony and Bremen.[5] This data includes dates and cost codes of dentist visits. Cost codes enable us to derive the diagnosis and to classify the services into diagnosis and counseling, prophylactic and therapeutic utilization. All visits within the borders of the federal state are covered. Finally, we use data from four face-to-face interviews (during pregnancy, as well as 6, 12 and 24 months after birth). These include questions on maternal health behavior, health

[3] In Germany, the public health insurance pays for routine midwife visits but pregnant women are responsible to contact a midwife and arrange appointments. Midwife visits can start during pregnancy and continue until eight weeks after birth. Visits can continue until one year after birth if medically indicated. The appointments mainly include health checks of the child and the mother and some counselling.

[4] Low income households are overrepresented in the AOK because 15 percent of the population, who are above a certain income threshold, or who are civil servants, can choose to be insured with a private health insurer and some public health insurance companies are more focused on specialized employees with often higher incomes.

[5] The regional KZVs are public bodies with compulsory membership for all dentists who provide services financed through the public health insurance. Dentists receive remuneration for publicly financed services not from the public health insurances directly, but indirectly through the KZVs.

care utilization for mother and child, and questions on maternal and child physical health and maternal mental health. All interviewers were blind to the treatment status of the participant.

Figure 1 gives an overview of the numbers of observations in the baseline sample and the different administrative and follow-up survey samples used in our analysis. The baseline sample covers all expectant mothers who originally registered for the study and participated in the initial interview. As Figure1 shows, the AOK insurance data comprises somewhat above 30 percent of the observations in the baseline sample. This data availability results from the combination of two factors. Although about 55 percent of the original participants are insured with the AOK, for one third of them a successful match to their health insurance record was impossible because their self-reported health insurance number was missing or incorrectly reported. For individuals included in the sample, data is available over our full period of analysis (up to 24 months after birth). As we show below, presence in the AOK sample is not related to treatment status.[6] Overall, the AOK sample still has a convenient size of 244 mothers and 220 children, which is larger than in other early childhood intervention trials. For example, the Perry Preschool Project had 123 participants (Heckman et al. 2010) and in Doyle et al. (2015), who analyze a home visiting program in Ireland, 173 of the initial 233 participants remained in the study at the 6-month interview.

As Figure 1 also shows, data from the dental health administrative sample is available for about 70 percent of the baseline sample in the regions of Lower Saxony and Bremen. The main reason for missing observations in the dental health sample is similar to that for the public health insurance data. However, the regional associations of public health insurance dentists cover all individuals in a given region, which explains the higher data availability for dental services.

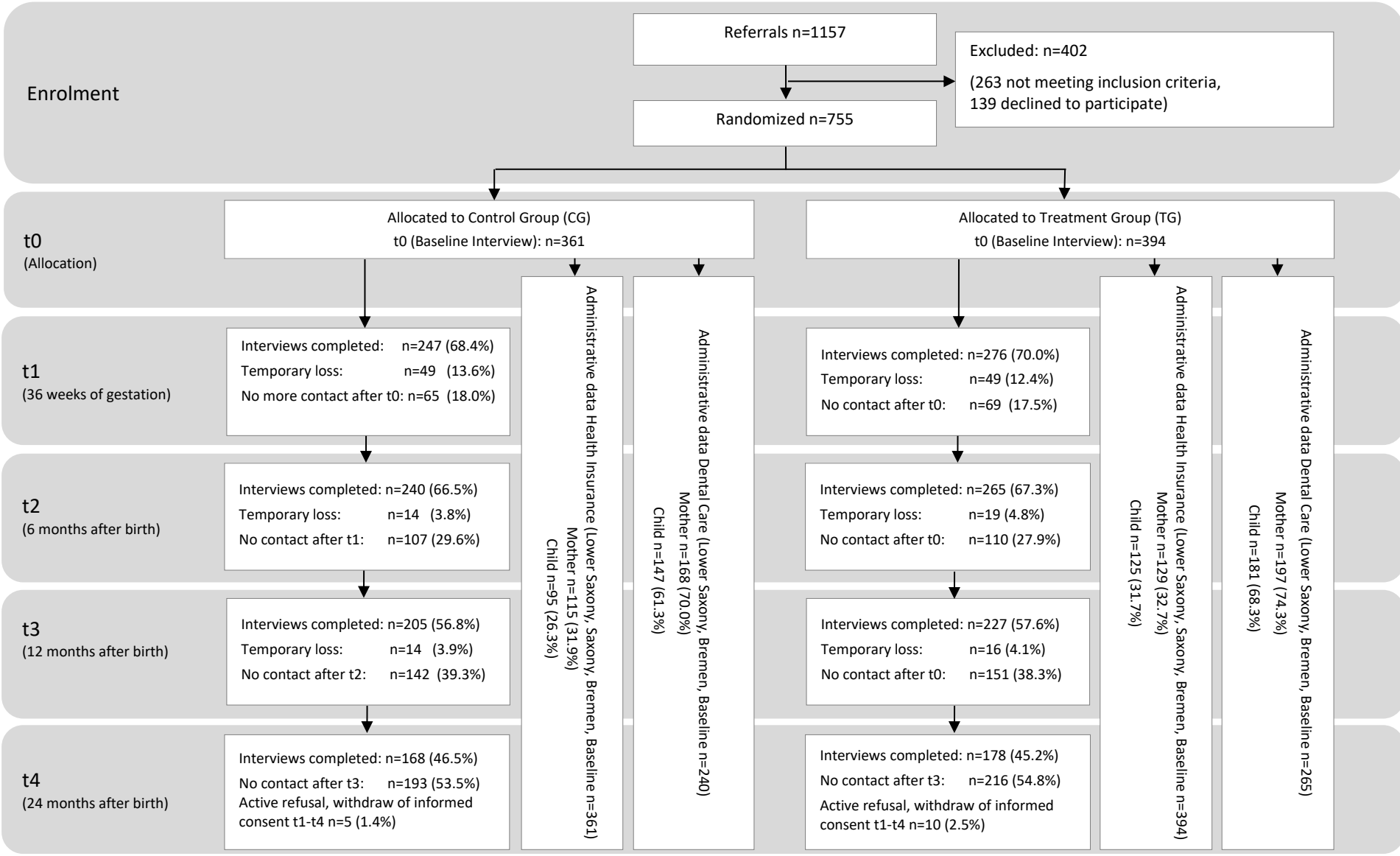
In the survey data there is some attrition over time, with about 70 percent of mothers in the baseline sample answering to the survey at pregnancy, and the response rate gradually declining to just below 50 percent at the last survey at 24 months of age of the child. The *Pro Kind* research team has not systematically recorded the reasons for attrition but anecdotal evidence suggests that the most important reasons were loss of contact with the family, or relocation. Such attrition is very common in experimental and non-experimental panel surveys with low SES samples and only poses a threat to identifying a causal effect if the attrition is correlated to the treatment, something we investigate below.[7]

---

[6] For every mother participating in the program there should be a connected child, but in the administrative data a few mother-child pairs are incomplete. This was mainly the case when the mother only participated in the pregnancy survey in which the health insurance number of the child was not yet available, or when the mother's or the child's health insurance number was missing or misreported in the survey.

[7] The levels of attrition we observe in the survey data are comparable to attrition of disadvantaged populations in non-experimental panel surveys. For example, in the KIGGS study or in the Panel Study 'Labour Market and Social Security' (PASS) retention is less than 50% for low SES or welfare receiving households (Lange et al. 2014; Trappmann et al. 2013). Although a little larger, our attrition rates are comparable to other experimental home visiting studies. Doyle et al. (2015) have an attrition rate of 36% of the baseline sample after 36 months and Robling et al. (2016) have 33% of missing values for outcomes directly after birth.

Figure 1: CONSORT flow diagram



Note: The diagram shows the participants' progress through the sampling phases. Numbers of treated and contols are not statistically significantly different at all stages, with the exception of one weakly significant difference (p=0.09) in the dental care sample for children.

To identify causal effects of the intervention on maternal and child health, characteristics of mothers and children for whom administrative and survey data is available should be balanced between treatment and control group. Successful randomization into the study groups of the *Pro Kind* trial ensures that this is the case in the baseline sample (Sierau et al. 2016; Sandner, 2017).[8] However, as shown above, the samples used in our analysis are available for a subset of individuals only. In Online Appendix Table A.1 we show that participants included in the administrative data deviate somewhat from the baseline sample in terms of their observed characteristics. This might indicate that some of the most disadvantaged participants may be underrepresented in the administrative data, probably because they might have been more likely to misreport their health insurance numbers. However, this does not by itself bias our treatment effect estimates.

Differences in unobserved characteristics across samples would only lead to biased treatment effects if the attrition was related to treatment status. For example, we could get a spurious treatment effect if individuals with unobserved characteristics related to ill-health would be more likely to drop out of the sample when assigned to the treatment group rather than the control group. For the administrative AOK data, such attrition based on treatment status is extremely unlikely and could only occur if individuals change their health insurance or move away from their federal state as a result of being assigned to the treatment group. For the administrative dental health data, only moving across state borders based on treatment status could lead to differential attrition. Both, switching between health insurances and moving across state borders is extremely rare in the target group of the program, and differential attrition by treatment status is therefore almost impossible in the administrative data. This is confirmed in Table 1, in which we show that for both, the AOK and the dental sample, and for mothers and children, the characteristics are overwhelmingly balanced across treatment and control group. Only three of the 120 hypothesis tests in the table (30 outcomes in 4 samples) indicate a weak statistically significant difference.[9]

Differential attrition across treatment and control groups could be more likely in the survey data than the administrative data, in case the program affects the participants' motivation to respond to follow-up surveys. However, it has been shown elsewhere (Sandner and Jungmann 2017) that, similarly to what we show for the administrative data, individuals included in the follow-up survey data have fewer risk factors, but their characteristics are balanced across treatment and control groups.

[8] None out of the 26 baseline variables analyzed in Sandner (2017), and only one (psychiatric disorder) out of the 32 baseline variables analyzed in Sierau et al. (2016), differs statistically significantly between treatment and control groups. We include this among the control variables in our study.

[9] Some of the control group means in Table 1 differ somewhat between the AOK and dental sample (e.g., higher shares of depression and anxiety in the AOK sample). This finding is not surprising due to the different populations that these samples represent. The dental sample represents only two out of the three federal states, and the AOK sample, while representing all three federal states, represents only those individuals who are insured with AOK. As we explain below, we address such differences in composition across samples by inverse probability weighting.

Table 1: Balancing tests for baseline characteristics

	AOK				Dental			
	Child		Mother		Child		Mother	
	Mean CG	*Diff TG-CG*	*Mean CG*	*Diff TG-CG*	*Mean CG*	*Diff TG-CG*	*Mean CG*	*Diff TG-CG*
Demographic Characteristics								
Teen	0.44	0.00	0.45	0.00	0.41	-0.03	0.43	-0.03
Age in Years	21.62	-0.52	21.78	-0.94*	21.64	0.64	21.61	0.42
Week in Pregnancy	20.43	-0.45	20.49	-0.51	19.82	0.01	19.52	0.26
Foreign Nationality	0.09	-0.03	0.08	-0.02	0.12	-0.01	0.13	-0.02
Underage	0.16	0.07	0.17	0.07	0.15	0.02	0.16	0.02
Mon. HH-Inc. in EUR	986	-36	959	-22	1020	13	985	58
Debt over 3000 EUR	0.17	-0.01	0.17	-0.02	0.14	0.06	0.14	0.06
Education Risk	0.79	-0.02	0.81	-0.03	0.76	-0.04	0.77	-0.03
Income Risk	0.77	0.04	0.77	0.03	0.74	0.04	0.77	0.01
Employment Risk	0.85	-0.04	0.88	-0.06	0.86	-0.06	0.88	-0.07*
No Partner	0.25	0.00	0.25	0.01	0.27	0.02	0.28	0.01
Living with Parents	0.28	0.02	0.29	0.02	0.33	-0.08	0.30	-0.05
Persons in HH	2.55	0.07	2.58	0.04	2.54	-0.02	2.51	0.00
Psychological and Physical Characteristics								
Unwanted Pregnancy	0.08	0.08*	0.10	0.07	0.16	0.02	0.16	0.03
Daily Smoking	0.34	0.01	0.34	0.05	0.36	-0.03	0.38	-0.05
Isolation	0.09	-0.03	0.09	-0.03	0.08	-0.03	0.08	-0.04
Foster Care Exper.	0.14	0.06	0.17	0.05	0.16	0.05	0.17	0.03
Neglect Experience	0.37	-0.03	0.41	-0.07	0.35	-0.02	0.37	-0.03
Loss Experience	0.52	-0.01	0.55	-0.04	0.55	-0.06	0.55	-0.05
Violence Ever	0.08	-0.03	0.09	-0.04	0.07	-0.01	0.08	-0.02
Depression	0.13	-0.01	0.15	-0.05	0.07	0.01	0.10	0.00
Anxiety	0.20	-0.06	0.20	-0.04	0.14	0.02	0.16	0.01
Stress	0.27	0.03	0.27	0.03	0.24	0.02	0.28	-0.01
Aggression	0.17	-0.06	0.17	-0.01	0.14	-0.03	0.17	-0.03
Body-Mass-Index	25.43	0.30	25.31	0.25	25.74	-0.14	25.33	0.27
Risky Pregnancy	0.08	0.03	0.09	0.04	0.13	-0.01	0.13	0.00
Sum Risk Factors	5.65	-0.17	5.86	-0.25	5.43	-0.12	5.68	-0.27
Lower Saxony	0.32	0.02	0.29	0.03	0.50	0.05	0.54	0.02
Bremen	0.31	-0.08	0.31	-0.07	0.50	-0.05	0.46	-0.02
Saxony	0.38	0.06	0.40	0.04				
Observations	***95***	***220***	***115***	***244***	***147***	***328***	***168***	***365***

Notes: The table reports control group means (CG) and differences between treatment and control group means (TG-CG) for observed characteristics from the baseline survey separately for mothers and children who are included in the two administrative samples (AOK and Dental) used in the subsequent analysis. See Online Appendix Table A.9 for a more detailed definition of the variables. Significance levels reported for two-tailed t-tests. Statistically significant at the *** 0.01 level, ** 0.05 level, * 0.10 level.

Data Source: Pro Kind participant survey at baseline.

A further important fact shown in Table 1 is that *Pro Kind* indeed reached a highly disadvantaged target group. Besides many socioeconomic risk factors (e.g., 44 percent are teenage mothers), the women show health related risk factors and adverse health behavior (e.g., 34 percent state to smoke daily during pregnancy).

4. Estimation Strategy

For a given outcome of interest Y_i we identify the treatment effect β via the regression

$$Y_i = \alpha + \beta D_i + \boldsymbol{\gamma} \boldsymbol{x_i'} + u_i, \quad (1)$$

in which α is a regression constant, D_i is the indicator for treatment assignment, $\boldsymbol{x_i'}$ is a vector of control variables, with associated coefficient vector $\boldsymbol{\gamma}$, and u_i the error term. When treatment assignment is based on stratified randomization, treatment effects should be estimated conditional on strata fixed effects (Kahan and Morris 2012). Among the control variables we therefore include a set of dummies for the stratification characteristics (municipality, teenage mother, foreign nationality). We also include the one baseline characteristic (psychiatric disorder) for which Sierau et al. (2016) report a statistically significant pre-treatment difference between treatment and control groups of the *Pro Kind* program, and the few characteristics for which we find statistically significant pre-treatment differences in our different samples in Table 1 (mother's age in years in the AOK sample for mothers, unwanted pregnancy in the AOK sample for children, and risk of employment loss in the dental sample for mothers). To take into account that the average observed characteristics in our estimation samples differ slightly from those in the wider baseline sample (as shown in Table A.1), we run regression (1) as a weighted regression, using the inverse of the (treatment and control-group specific) probability of being included in the estimation sample as weights. To obtain the weights for the inverse probability weighting (IPW) we run logit regressions of an indicator for whether an individual is included in each of the respective estimation samples, separately for treatment and control groups, on the observed baseline characteristics shown in Table 1. We then construct the weights as the inverse of the predicted probability from these regressions. Under the assumption that the attrition pattern is determined solely by observed characteristics, this weighting ensures that the results are representative for the characteristics of the baseline sample. These results should thus have external validity in predicting the effect of the program that we can expect if the program would be implemented at a larger scale on a population with similar baseline characteristics.

When interpreting the treatment effect, it should be kept in mind that the number of home visits varied across participants, so that the effect size should be interpreted against the average number of home visits received (which is 32.7, see section 2 above for more detail). In further analysis in section 6 below we look for evidence of whether effects differ by the actual number of home visits received (which is however not exogenous), and at other sub-group results.

For inference on the statistical significance of the treatment effect β, we rely on several different types of hypothesis tests. As a benchmark, we present results from a standard (two-tailed) t-test. As an alternative, we present results of a permutation test. The classical t-test obtains a p-value by comparing the test statistic to its theoretical sampling distribution derived from distributional

assumptions. These assumptions are unlikely to hold in small samples and with non-normally distributed data. In this case a permutation test is preferable (Hayes 1996; Heckman et al. 2010). It obtains a p-value by comparing the test statistic to a distribution generated from the observed data. This distribution of the test statistic is generated by randomly permuting the treatment indicator across individuals (within strata from the stratified randomization) and repeatedly estimating the test statistic, which provides a data-driven distribution of the test statistic under the null hypothesis.

Finally, to take into account that we analyze many outcomes, we present p-values that adjust our results for multiple hypothesis testing (MHT), a problem increasingly recognized in the literature on early childhood interventions (e.g., Anderson 2008; Heckman et al. 2010; Heckman et al. 2017; Doyle et al. 2015). To control the family-wise type I error rate within a family of outcomes at the desired level α, we conduct MHT based on the Romano-Wolf method (Romano and Wolf 2005a,b, 2016).[10] This method provides adjusted p-values that we display alongside our results. For the purpose of MHT, we group variables together into a family if they measure conceptually the same or a similar outcome, for example different measures of health care costs or different prescription-based and diagnosis-based measures proxying for the same underlying medical condition.

5. Results

5.1 Health care utilization

In Table 2 we report program effects on the overall utilization of health care by mothers and children based on the administrative data. Throughout the paper, we assign outcomes that measure a similar underlying concept into groups that are identified in the tables by bold titles. We carry out multiple hypothesis testing within each of these groups. In Table 2 the main groups consist of hospital utilization at different stages and overall cost, separately for mother and child, as well as midwife utilization for the mothers. The control group means give an overall picture of hospital utilization of the target group. The share of mothers admitted to hospital during pregnancy and birth is almost 100%, which is to be expected because births outside of hospital are extremely rare in Germany. The reason why children's hospital utilization at birth is lower (15%) is because newborns are only registered as admitted to hospital if they receive an own diagnosis. Mothers stay in hospital at pregnancy and birth on average for almost 12 days (four of these 12 days correspond to the actual hospital spell at birth, in line with the national average).

[10] We implement this via the Stata command **rwolf** developed by Damian Clarke which is available in Stata by typing **ssc install rwolf**. It is based on bootstrap resampling for which we use 1,000 repetitions, stratifying the bootstrap in a way that ensures a constant number of non-missing observations for each outcome across all bootstraps. We slightly modified the code of the program to allow outcome-specific IPW weights.

Table 2: Health care utilization

	Obs.	*Mean CG*	*Treatm. effect*	*Std. error*	p-values *t-test*	*perm.*	*MHT*	*Odds Ratio [Cohen's d]*
				Mother				
Hospital during Pregnancy and at Birth								
Any Admission at Birth/Pregnancy	244	1.00	-0.01	(0.01)	0.18	0.45	0.31	0.00
Nights in Hospital	244	11.86	1.90	(2.13)	0.38	0.46	0.35	[0.10]
Nights in Hospital (cond.)	242	11.86	2.02	(2.13)	0.35	0.42	0.34	[0.10]
Hospital 0-2 (without Birth and Pregnancy)								
Any Admission after Birth	244	0.39	-0.07	(0.06)	0.24	0.24	0.50	0.73
Nights in Hospital	244	4.12	-0.96	(1.15)	0.41	0.52	0.57	[-0.10]
Nights in Hospital (cond.)	80	10.60	-0.02	(2.98)	0.99	1.00	1.00	[-0.00]
Midwife Utilization (Health Insurance Financed)								
Total Contacts	244	18.32	-0.09	(1.62)	0.96	0.97	0.96	[-0.01]
Number of Contacts during Preg.	244	6.51	-0.43	(0.85)	0.61	0.64	0.85	[-0.07]
Number of Contacts after Birth	244	11.71	0.38	(1.06)	0.72	0.73	0.87	[0.05]
Any Contact during Preg.	244	0.73	-0.06	(0.06)	0.35	0.39	0.68	0.76
Any Contact after Birth	244	0.82	0.06	(0.05)	0.16	0.21	0.43	1.70
Costs								
Costs hospitalisation	244	4838	425.3	(537)	0.43	0.50	0.68	[0.09]
Costs medication	244	350	-208.9	(174)	0.23	0.11	0.55	[-0.25]
Costs Midwifes (heal. insur. financed)	244	506	15.2	(54)	0.78	0.80	0.76	[0.04]
				Child				
Hospital at Birth								
Any Admission at Birth	220	0.15	-0.03	(0.05)	0.62	0.65	0.59	0.80
Nights in Hospital	220	4.7	-3.1	(2.1)	0.14	0.16	0.31	[-0.18]
Nights in Hospital (cond.)	27	31.4	-23.5	(21.3)	0.29	0.41	0.39	[-0.58]
Hospital 0-2 (without Birth)								
Any Admission	220	0.49	-0.03	(0.07)	0.64	0.65	0.89	0.87
Nights in Hospital	220	5.4	0.0	(2.2)	1.00	1.00	1.00	[-0.00]
Nights in Hospital (cond.)	98	11.0	0.8	(5.1)	0.88	0.89	0.96	[0.03]
Costs								
Costs hospitalisation	220	4993	-1447.3	(1734)	0.41	0.45	0.61	[-0.10]
Costs medication	220	454	-1.8	(106)	0.99	0.99	0.98	[-0.00]

Notes: Obs.=No. of Observations, CG=Control Group, cond.=conditional on admission, perm.=permutation test, MHT=Multiple Hypothesis Testing. Treatment effects are estimated from IPW-weighted OLS regressions with control variables, as described in Section 4 of the text. Groups of variables that are assigned to the same family for MHT are separated by bold titles. Online Appendix Table A.9 provides definitions of the variables. All p-values are two-tailed.

Data Source: Administrative data from the AOK public health insurance.

The hospital admissions rate of 46% in the first two years after birth is largely representative for low SES children.[11] Overall, hospitalizations are an important outcome because the available U.S. evidence on home visiting points towards a reduction in hospitalizations due to injuries and accidents. Moreover, costs are a particularly attractive summary measure, because they are sensitive not only to the quantity of health care utilization (such as number of admissions), but also to the severity of the underlying conditions. However, in contrast to the U.S. evidence, none of these important outcomes in Table 2 is statistically significantly affected by the program

5.1 Diagnoses and prescriptions

Following the goals and guidelines of the *Pro Kind* program, we used hospital diagnoses codes (ICD codes) and medication prescriptions (ATC codes) to generate seven groups of child health outcomes that the program could have affected.[12] The groups are typed in bold italics in Table 3 (see Online Appendix Table A.9 for a more detailed description of the individual variables). Three outcome measures show marginally statistically significant (at the 10%-level) differences based on *individual* hypothesis tests. These relate to prescriptions of drugs for obstructive airway diseases and of antidiarrheals or anti-inflammatory agents, as well as to hospital diagnoses for injuries, burns, corrosion, or poisoning. However, after MHT, we do not find effects in any of the seven groups. Therefore, our overall conclusion from the results in Table 3 is that child health outcomes, as proxied by diagnosis codes from hospitalizations and from medication prescriptions, are not statistically significantly affected by the program.

In Table 4 we report the results for a similar range of outcomes defined for mothers. For one group of variables, maternal mental health, we find significantly (1%-level in an individual permutation test, 5%-level when adjusting for multiple hypothesis testing) reduced prescriptions of calmatives. Below we show that this finding is backed up by evidence of a reduction of survey-reported maternal mental health outcomes. For the other outcome groups we do not find any significant effect after adjusting for MHT. Overall, the combined findings of Tables 3 and 4 suggest very little effect of the program on mother and child health outcomes, with the notable exception of a positive effect on maternal mental health.

[11] To check this we analyzed data from German Health Interview and Examination Survey for Children and Adolescents (KIGGS) and found that 18.3% of low-SES parents of children up to the age of 2 state that their child was admitted to hospital within the last 12 months. While this may be an appropriate estimate of the share of hospital admissions in the second year of a low-SES child's life, it probably understates the admission probability during the first year of life. It is therefore plausible for the two-year admission rate to be more than twice this number.

[12] It should be noted that in Germany over-the-counter (non-prescription) medications for children can be reimbursed by health insurances if a doctor writes a note recommending their use. That explains the relatively high means, for example, for drugs such as cough and cold preparations and painkillers for children in Table 3.

Table 3: Health outcomes from diagnoses and prescriptions (Child)

	Obs.	*Mean CG*	*Treatm. effect*	*Std. error*	p-values *t-test*	*perm.*	*MHT*	*Odds Ratio*
Birth outcomes								
Birthweight Below 2500g (Hosp.)	220	0.04	-0.01	(0.02)	0.77	0.80	0.95	0.82
Other Perinatal Conditions (Hosp.)	220	0.14	-0.02	(0.05)	0.76	0.78	0.95	0.88
Respiratory tract conditions								
Diseases of the Respiratory System (Hosp.)	220	0.20	-0.03	(0.06)	0.67	0.69	0.90	0.85
Nasal Decongestants (Presp.)	220	0.91	-0.01	(0.04)	0.82	0.80	0.90	0.90
Drugs for Obstructive Airway Diseases (Presp.)	220	0.41	-0.12	(0.07)	**0.09**	0.11	0.37	0.58
Cough and Cold Preparations (Presp.)	220	0.87	0.03	(0.05)	0.55	0.53	0.90	1.31
Antiasthmatic Drugs (Presp.)	220	0.24	-0.08	(0.06)	0.18	0.22	0.52	0.61
Antiasthmatic Drugs More Than Once (Presp.)	220	0.16	-0.077	(0.05)	0.11	0.15	0.37	0.46
Diseases of digestive system								
Intestinal Infectious Diseases (Hosp.)	220	0.14	0.02	(0.05)	0.64	0.66	0.87	1.21
Diseases of the Digestive System (Hosp.)	220	0.03	0.01	(0.02)	0.81	0.82	0.87	1.21
Drugs for Gastrointestinal Disorder (Presp.)	220	0.40	-0.05	(0.07)	0.43	0.45	0.81	0.79
Drugs for Constipation (Presp.)	220	0.19	-0.06	(0.05)	0.25	0.28	0.68	0.64
Antidiarrheals or Anti-inflammatory Agents (Presp.)	220	0.41	0.13	(0.07)	**0.09**	0.11	0.35	1.66
Injuries, Poisoning								
Head Injuries (Hosp.)	220	0.09	0.04	(0.04)	0.36	0.38	0.35	1.49
Other Injuries, Burns, Corrosion, Poisoning (Hosp.)	220	0.01	0.03	(0.02)	**0.098**	**0.08**	0.26	5.86
Emollients and Protectives (Presp.)	220	0.35	-0.11	(0.07)	0.13	0.12	0.26	0.60
Dermatological conditions								
Antifungals for Dermatological Use (Presp.)	220	0.74	-0.11	(0.07)	0.11	0.11	0.28	0.60
Preparations for Wounds and Ulcers (Presp.)	220	0.17	-0.04	(0.05)	0.49	0.51	0.74	0.76
Corticosteroids, Dermatological Prep. (Presp.)	220	0.18	0.02	(0.06)	0.69	0.67	0.74	1.18
Mental health								
Mental and Behavioural Disorders (Hosp.)	220	0.02	-0.01	(0.02)	0.37	0.46	0.63	0.36
Psycholeptics - Calmatives (Presp.)	220	0.04	0.01	(0.03)	0.77	0.80	0.78	1.20
Prescription of multipurpose drugs								
Painkillers (Presp.)	220	0.96	0.02	(0.03)	0.42	0.43	0.68	2.07
Antibacterials, Antibiotics for Systemic Use (Presp.)	220	0.56	0.09	(0.07)	0.21	0.22	0.51	1.48
Vitamins (Presp.)	220	0.46	0.09	(0.07)	0.16	0.18	0.51	1.45
Antiinflammatory Products (Presp.)	220	0.60	0.01	(0.07)	0.86	0.86	0.86	1.05

Notes: Obs.=No. of Observations, CG=Control Group, perm.=permutation test, MHT=Multiple Hypothesis Testing. Treatment effects are estimated from IPW-weighted OLS regressions with control variables, as described in Section 4 of the text. Outcomes are derived from hospital diagnoses (Hosp. = Any admission because of the indicated diagnosis) and medication prescriptions (Presp. = Any prescription in the first two years of life). Groups of variables that are assigned to the same family for MHT are separated by bold titles. Online Appendix Table A.9 provides definitions of the variables. p-values below 0.1 are in bold. All p-values are two-tailed.

Data Source: Administrative data from the AOK public health insurance.

Table 4: Health outcomes from diagnoses and prescriptions (Mother)

	Obs.	*Mean CG*	*Treatm. effect*	*Std. error*	p-values *t-test*	p-values *perm.*	p-values *MHT*	*Odds Ratio*
Birth complications								
Indications Related to Preg./Birth (Hosp.)	244	0.95	0.04	(0.03)	0.17	0.17	0.30	2.95
Complications at Delivery (Hosp.)	244	0.41	0.13	(0.07)	**0.05**	**0.05**	0.13	1.69
Complications of Fetus/Newborn (Hosp.)	244	0.11	-0.01	(0.04)	0.76	0.76	0.76	0.87
Fertility								
Subsequent Birth (Hosp.)	244	0.04	0.05	(0.03)	0.14	0.16	0.26	2.32
Contraceptives (Presp.)	244	0.07	0.08	(0.08)	0.35	0.48	0.38	2.29
Mental health								
Mental and Behavioural Disorders (Hosp.)	244	0.05	-0.01	(0.03)	0.77	0.78	0.76	0.84
Psycholeptics - Calmatives (Presp.)	244	0.07	-0.07	(0.03)	**0.02**	**0.004**	**0.04**	0.00
Psychoanaleptics - Antidepressants (Presp.)	244	0.11	-0.04	(0.04)	0.28	0.26	0.44	0.59
Respiratory tract conditions								
Diseases of the Respiratory System (Hosp.)	244	0.00	0.02	(0.01)	0.17	0.36	0.49	
Nasal Decongestants (Presp.)	244	0.06	-0.06	(0.03)	**0.02**	**0.01**	0.11	0.00
Drugs for Obstructive Airway Diseases (Presp.)	244	0.08	0.03	(0.04)	0.54	0.54	0.70	1.36
Cough and Cold Preparations (Presp.)	244	0.07	0.03	(0.04)	0.45	0.46	0.70	1.46
Antiasthmatic (Presp.)	244	0.06	0.05	(0.04)	0.20	0.19	0.49	1.97
Antiasthmatic more than once (Presp.)	244	0.02	0.04	(0.03)	0.13	0.11	0.39	3.35
Prescription of multipurpose drugs								
Painkillers (Presp.)	244	0.18	0.00	(0.06)	0.96	0.97	0.96	0.98
Antibacterials or Antibiotics (Presp.)	244	0.54	0.04	(0.07)	0.59	0.60	0.84	1.16
Vitamins (Presp.)	244	0.05	-0.03	(0.03)	0.39	0.41	0.84	0.43
Anti-inflammatory Products (Presp.)	244	0.27	0.06	(0.06)	0.37	0.35	0.84	1.31

Notes: Obs.=No. of Observations, CG=Control Group, perm.=permutation test, MHT=Multiple Hypothesis Testing. Treatment effects are estimated from IPW-weighted OLS regressions with control variables, as described in Section 4 of the text. Outcomes are derived from hospital diagnoses (Hosp. = Any admission because of the indicated diagnosis) and medication prescriptions (Presp. = Any prescription in the first two years of life). Groups of variables that are assigned to the same family for MHT are separated by bold titles. Online Appendix Table A.9 provides definitions of the variables. For binary outcomes with a control group mean of 0 the odds ratio is not defined and therefore not reported. p-values below 0.1 are in bold. All p-values are two-tailed.

Data Source: Administrative data from the AOK public health insurance.

5.1 Dental Health

To check whether the program affected dental health outcomes and utilization, we now switch to the dental health administrative dataset. In Table 5 we present results for mothers (top panel) and children (bottom panel). For each point in time we distinguish between diagnostic/prophylactic dental care and therapeutic dental care, because the aim of the program was to increase prophylactic care that has the potential to reduce therapeutic care.[13] For mothers,

[13] Based on the cost codes associated with each dental visit, we classified the services carried out at each visit as either diagnostic, prophylactic, or therapeutic (see Online Appendix Table A.9 for more details). For children, diagnostic and prophylactic services were difficult to distinguish and therefore combined into one category.

the results show a significant increase in diagnostic and prophylactic dentist visits during pregnancy, but no effect on therapeutic visits. We find an increase of 16 percentage points in the probability to attend at least one diagnostic visit, compared to a control group mean of 10 percent, and an increase of 9 percentage points in the probability to attend a prophylactic visit (control group mean 17 percent). These increases are individually significant and remain statistically significant at the 1%-level and 5%-level respectively when applying multiple hypothesis testing. In the second group of maternal outcomes, measured in the period from after pregnancy up to the child's second birthday, there is an individually significant effect on both, the probability of at least one diagnostic visit, and at least one therapeutic visit, but these effects are not significant in multiple hypothesis testing.

With respect to dentist visits of the child, there is an interesting pattern of an increased probability of prophylactic visits (although not statistically significant) and a marked decrease in

Table 5: Dental health outcomes

	Obs.	*Mean CG*	*Treatm. effect*	*Std. error*	p-values *t-test*	p-values *perm.*	p-values *MHT*	*Odds Ratio [Cohen's d]*
					Mother			
Pregnancy - Diagnostic and Prophylactic care								
Any visit (diagnostic)	365	0.10	0.16	(0.05)	**0.002**	**0.01**	**0.005**	3.19
Any visit (prophylactic)	365	0.17	0.09	(0.04)	**0.04**	**0.04**	**0.07**	1.71
Number of visits	365	0.71	0.37	(0.19)	**0.06**	**0.07**	**0.07**	[0.19]
Pregnancy - Therapeutic care								
Any visit	365	0.17	0.01	(0.04)	0.76	0.73	0.93	1.09
Number of visits	365	0.36	0.01	(0.12)	0.91	0.91	0.93	[0.01]
Child Age 0-2 - Diagnostic and Prophylactic care								
Any visit (diagnostic)	365	0.58	0.11	(0.06)	**0.05**	**0.07**	0.11	1.62
Any visit (prophylactic)	365	0.50	0.03	(0.05)	0.65	0.66	0.85	1.11
Number of visits	365	3.36	-0.10	(0.42)	0.81	0.79	0.85	[-0.03]
Child Age 0-2 - Therapeutic care								
Any visit	365	0.48	0.10	(0.05)	**0.07**	**0.09**	0.13	1.48
Number of visits	365	2.53	0.27	(0.56)	0.63	0.63	0.64	[0.06]
					Child			
Child Age 0-2 - Prophylactic care								
Any visit	328	0.20	0.05	(0.05)	0.35	0.34	0.48	1.31
Number of visits	328	0.35	-0.01	(0.09)	0.94	0.95	0.94	[-0.01]
Child Age 0-2 - Therapeutic care								
Any visit	328	0.03	-0.03	(0.01)	**0.05**	**0.03**	**0.07**	0.00
Number of visits	328	0.15	-0.15	(0.08)	**0.06**	**0.02**	**0.07**	[-0.20]

Notes: Obs.=No. of Observations, CG=Control Group, perm.=permutation test, MHT=Multiple Hypothesis Testing. Treatment effects are estimated from IPW-weighted OLS regressions with control variables, as described in Section 4 of the text. Groups of variables that are assigned to the same family for MHT are separated by bold titles. Online Appendix Table A.9 provides definitions of the variables. p-values below 0.1 are in bold. All p-values are two-tailed.

Data Source: Administrative data from the KZV public dentists' association.

therapeutic visits, which is statistically significant at the 10%-level after multiple hypothesis testing. Our overall reading of the results in Table 5 is that the *Pro Kind* program increased maternal prophylactic dentist visits during pregnancy, and that it may also have increased maternal and child prophylactic dentist visits from birth up to age 2 of the child, and seems to have decreased therapeutic dentist visits of the children.

5.1 Survey Data

All previous results were based on administrative data. We now turn to survey data to check whether the previous results are in line with maternal self-reports on health behavior and other health outcomes. We again arrange variables into groups of outcomes according to the program goals the time at which they were surveyed. The top panel of Table 6 reports results on important maternal health behaviors that were targeted by the program: smoking, breastfeeding and nutrition of the child. At pregnancy, a strikingly high percentage of women in the control group smoke daily (30 percent), and even more state that they smoke in general (39 percent). The program, however, seems to have had no effect on this behavior. Six, twelve, and twenty-four months after birth, maternal smoking in the control group increases from 39 percent during pregnancy to above 60 percent, but with no effect of the program on this behavior. Breastfeeding behavior is surveyed at 6 months after birth. Here, 50 percent of mothers in the control group report not having breastfed for longer than one month, with no statistically significant treatment effect on this outcome. Among all maternal health behaviors reported in the table, the only marginally significant effect in individual hypothesis tests, but not after MHT, is a reduction in the share of mothers offering sugary drinks to their children.

The bottom panel of Table 6 reports results on preventive health utilization according to maternal survey responses. Outcomes include prenatal screening, vaccinations, and postnatal checkups. While there are no significant effects of the program on these groups of outcomes, the control group means suggest that the target population as a whole shows good preventive health behavior, with 83 percent of mothers attending all recommended prenatal checkups, 98 percent attending all 4 postnatal checkups up to age 1 of the child, a number that gradually falls to 68 percent for all 7 postnatal checkups recommended up to age 2 of the child.[14] Furthermore, almost all mothers have had their children vaccinated.

[14] The reduction at the 7th postnatal checkup may be explained with a timing effect. The recommended age for the 7th postnatal checkup is between the 22 and 24 months after birth. A small delay in the checkup would therefore mean that it wouldn't be recorded in the interview 24 months after birth.

Table 6: Maternal health behaviour and preventive utilization (Survey Data)

	Obs.	*Mean CG*	*Treatm. effect*	*Std. error*	p-values *t-test*	p-values *perm.*	p-values *MHT*	*Odds Ratio [Cohen's d]*
Pregnancy			***Maternal Health Behavior***					
Smoking	523	0.39	0.01	(0.04)	0.75	0.78	0.98	1.06
Daily Smoking	523	0.30	0.00	(0.04)	0.97	0.96	1.00	0.99
Expecting Mo. Passive Smoke	521	0.57	0.00	(0.04)	0.99	0.99	1.00	1.00
6 Months after Birth								
Smoking	499	0.64	-0.03	(0.04)	0.53	0.55	0.98	0.89
Daily Smoking	499	0.54	-0.01	(0.05)	0.82	0.82	0.98	0.96
Child in Room where Smoking	496	0.11	0.02	(0.03)	0.56	0.56	0.98	1.19
No Breastfeeding	503	0.23	0.03	(0.04)	0.50	0.49	0.98	1.16
Breastfeeding less than 31 Days	502	0.50	0.00	(0.05)	0.97	0.97	0.98	1.01
Any Sugary Drinks	489	0.24	-0.02	(0.04)	0.56	0.53	0.98	0.88
Any Sweets	489	0.33	-0.02	(0.04)	0.72	0.71	0.98	0.93
12 Months after Birth								
Smoking	431	0.65	-0.02	(0.05)	0.62	0.61	0.76	0.90
Daily Smoking	431	0.51	0.04	(0.05)	0.39	0.38	0.76	1.19
Child in Room where Smoking	427	0.22	-0.04	(0.04)	0.37	0.37	0.76	0.79
Any Sugary Drinks	431	0.46	-0.09	(0.05)	**0.09**	**0.10**	0.35	0.70
Any Sweets	431	0.85	-0.04	(0.04)	0.34	0.37	0.76	0.76
24 Months after Birth								
Smoking	342	0.61	0.02	(0.06)	0.75	0.76	0.94	1.08
Daily Smoking	342	0.54	0.04	(0.06)	0.51	0.52	0.81	1.16
Child in Room where Smoking	333	0.28	-0.01	(0.05)	0.81	0.81	0.94	0.94
Up to 6 Months after Birth			***Preventive Utilization***					
All Prenatal Screening Examinations	521	0.83	-0.03	(0.03)	0.40	0.43	0.71	0.83
Sum Vaccinations	443	6.52	0.01	(0.14)	0.92	0.92	0.99	[0.010]
Any Vaccination	443	0.97	0.001	(0.02)	0.95	0.95	0.99	1.04
All 4 Postnatal Check Ups	518	0.98	0.01	(0.01)	0.25	0.30	0.62	2.30
12 Months after Birth								
Sum Vaccinations	393	9.39	-0.001	(0.25)	1.00	1.00	1.00	[-0.00]
Any Vaccination	393	1.00	-0.01	(0.01)	0.56	0.65	0.90	0.45
All 6 Postnatal Check Ups	397	0.89	-0.01	(0.03)	0.77	0.77	0.94	0.91
24 Months after Birth								
Sum Vaccinations	296	10.88	0.23	(0.22)	0.30	0.27	0.56	[0.13]
Any Vaccination	296	0.99	0.01	(0.01)	0.33	0.34	0.69	
All 7 Postnatal Check Ups	275	0.68	0.01	(0.07)	0.92	0.92	0.91	1.03

Notes: Obs.=No. of Observations, CG=Control Group, perm.=permutation test, MHT=Multiple Hypothesis Testing. Treatment effects are estimated from IPW-weighted OLS regressions with control variables, as described in Section 4 of the text. Groups of variables that are assigned to the same family for MHT are separated by bold titles. Online Appendix Table A.9 provides definitions of the variables. For binary outcomes with a treatment group mean of 1 the odds ratio is not defined and therefore not reported. p-values below 0.1 are in bold. All p-values are two-tailed.

Data Source: Four waves of the *Pro Kind* participant survey

Table 7 reports the results for additional measures from the survey data, in particular maternal mental health (top panel), as well as birth outcomes and information on accidents of the child at 12 and 24 months after birth (bottom panel). Maternal mental health is captured by the outcomes depression, anxiety and stress, constructed from the Depression, Anxiety, and Stress scales (DASS, Lovibond and Lovibond 1995).[15] Interestingly, *all* of the maternal mental health indicators show negative point estimates, suggesting that the *Pro Kind* program improved maternal mental health. The reduction is individually statistically significant at the 5%-level for depression at 24 months after birth, and this effect remains significant at the 10%-level after adjusting for MHT.[16] The program reduced the probability of maternal depression by 11 percentage points, compared to a control group mean of 28 percent. This corroborates the finding from the administrative data in Table 4, where we found a strong reduction in prescriptions of psycholeptics.

The results on child health, reported in the bottom half of Table 7, confirm our earlier findings from the administrative data of no statistically significant effects of the program on birth outcomes after MHT. For both surveys, at 12 months and at 24 months after birth, the results also confirm that the program did not affect the probability of accidents of the child. Thus, the individually statistically significant positive effect in the administrative data on other injuries, burns, corrosion, or intoxication reported in Table 3 is not confirmed by any statistically significant effect in the survey data.

5.5. Additional results

As we show in Online Appendix Tables A.2 to A.7, our main results of no effects on physical health, but positive effects on maternal mental health and utilization of dental health care continue to hold when we do not use IPW and exclude control variables. With a 6-percentage-point reduction in prescriptions of psycholeptics, and a 9-percentage-point reduction of maternal depression measured at age 2 of the child, effect sizes are slightly smaller but very close to the IPW-weighted results. The increased maternal dental visits for diagnostic and prophylactic purposes are also statistically significant and of a comparable size as in the IPW weighted results, but the decreased therapeutic dental visits of the child lose statistical significance after MHT in the unweighted regressions.

Online Appendix Table A.8 reports separate results by gender for the child health outcomes included in Tables 2 and 3. One noteworthy result is that the individually marginally significant reduction in medication for obstructive airway diseases shown in Table 3 seems to be entirely

[15] The DASS consists of 42 negative emotional symptoms. Participants rate the extent to which they have experienced each symptom over the past week, on a 4-point severity and frequency scale. Subscale scores are determined by summing up the values for the relevant 14 items. Internal consistencies (alpha coefficient) for each scale for the DASS normative sample were: Depression 0.91; Anxiety 0.84; Stress 0.90.

[16] In a robustness check not reported here, this effect also remains significant at the 10%-level when we define the MHT groups by mental health outcome instead of by time point.

Table 7: Maternal mental health and child health (Survey Data)

	Obs.	*Mean CG*	*Treatm. effect*	*Std. error*	p-values *t-test*	p-values *perm.*	p-values *MHT*	*Odds Ratio [Cohen's d]*
Pregnancy			***Maternal Mental Health***					
Depression	523	0.26	-0.01	(0.04)	0.76	0.77	0.73	0.94
Anxiety	523	0.33	-0.06	(0.04)	0.17	0.15	0.37	0.77
Stress	523	0.49	-0.05	(0.05)	0.29	0.28	0.47	0.82
6 Months after Birth								
Depression	501	0.17	-0.01	(0.04)	0.83	0.83	0.97	0.95
Anxiety	501	0.16	-0.01	(0.03)	0.87	0.88	0.97	0.96
Stress	501	0.36	-0.07	(0.04)	0.13	0.13	0.31	0.74
24 Months after Birth								
Depression	344	0.28	-0.11	(0.05)	**0.03**	**0.02**	**0.07**	0.54
Anxiety	344	0.17	-0.04	(0.04)	0.40	0.37	0.40	0.75
Stress	344	0.37	-0.07	(0.05)	0.17	0.16	0.28	0.72
Pregnancy			***Child Health***					
Birthweight	603	3237	83	(43.8)	**0.06**	**0.07**	0.19	[0.16]
Birth Before 37 Week of Gestation	600	0.04	0.03	(0.02)	**0.08**	0.10	0.23	1.83
Low Birth Weight (<2500g)	603	0.06	0.00	(0.02)	0.86	0.87	0.96	1.06
Apgar Score < 10	495	0.84	-0.01	(0.03)	0.80	0.79	0.96	0.94
12 Months after Birth								
Any Accident	429	0.19	-0.01	(0.04)	0.80	0.80	0.93	0.94
Number of Accidents	429	0.23	0.00	(0.06)	0.98	0.99	0.99	1.01
Accident with Hospital Stay	429	0.06	0.02	(0.03)	0.58	0.57	0.88	1.28
24 Months after Birth								
Any Accident	310	0.36	-0.03	(0.06)	0.61	0.62	0.90	0.88
Number of Accidents	310	0.50	-0.04	(0.10)	0.68	0.69	0.90	0.85
Accident with Hospital Stay	310	0.11	0.02	(0.05)	0.73	0.71	0.90	1.17

Notes: Obs.=No. of Observations, CG=Control Group, perm.=permutation test, MHT=Multiple Hypothesis Testing. Treatment effects are estimated from IPW-weighted OLS regressions with control variables, as described in Section 4 of the text. Groups of variables that are assigned to the same family for MHT are separated by bold titles. Online Appendix Table A.9 provides definitions of the variables. p-values below 0.1 are in bold. All p-values are two-tailed.

Data Source: Four waves of the *Pro Kind* participant survey

driven by girls, for whom the effect is now much larger and significant at the 5%-level after adjusting for multiple hypothesis testing across outcomes and gender subgroups. We are not aware of any research on gender differences in the physiological effect of passive smoking on children. If the gender difference in the treatment effect on obstructive airway diseases is indeed related to maternal smoking, the only plausible mechanism would be that mothers of girls reduce smoking in presence of their child in response to the program while mothers of boys do not. When we break up the results of maternal smoking behavior in Table 6 by gender of the child (result not reported

here), however, we do not find any evidence of a differential program effect on smoking by gender of the child. It is thus difficult to argue that the effect on obstructive airway diseases for girls is due to a reduction in maternal smoking. The only other effects in Table A.8 that gain weak statistical significance at the 10%-level with MHT are a reduction in prescriptions of antifungals for dermatological use for girls, and a reduction of drugs for gastrointestinal disorders for boys.

6. Discussion and Further Analysis

In this section we shed some light on possible mechanisms through which the *Pro Kind* program may have affected maternal mental health, and then discuss potential reasons why it may not have affected child health or maternal health behaviors.

One reason why the program may have affected mental health is that the target population of disadvantaged first-time mothers has a particularly elevated risk of mental health problems. There is a well-documented SES gradient in mental health (Aneshensel 2009; Muntaner et al. 2004; Wildman 2003) which is thought to be caused by stressors related to social disadvantage (Aneshensel 2009; Dohrenwend 2000; Turner et al. 1995) combined with a reduced access to psychosocial support and reduced effectiveness of such support (Aneshensel 2009). Stressful life events and lack of social support are also among the main risk factors for developing postnatal depression (Dennis 2005; Cooper and Murray 1998). To test whether these factors are related to the effectiveness of the program, we run subgroup analyses for the mental health effects by stress levels, social risk factors and social isolation at baseline. As we report in rows (i)-(iv) of Table 8, for both, psycholeptics in the administrative data (Panel A) and depression in the survey data (Panel B), the positive mental health effects are particularly driven by mothers with more elevated stress levels, more social risk factors, less friends, and less social support at baseline. These findings suggest that the *Pro Kind* program may affect mental health because the home visits compensate for the lack of social support of socially disadvantaged mothers who face a stressful life environment. A constituent idea of the program is that the personal relationship and attachment between mother and home visitor built up during the frequent and intensive visits fosters maternal self-efficacy. Higher self-efficacy, defined as the mother's belief of being able to achieve goals and influence outcomes, may increase the ability to cope with a stressful life environment. In support of the idea that the attachment between mother and home visitor matters as a channel for the mental health effects, row (v) of Table 8 indicates that the beneficial effects on prescriptions of psycholeptics (Panel A) and depression (Panel B) are greater among mothers who had the same home visitor throughout the program.[17] In row (vi) of Panels A and B of Table 8 we further break up the effect by whether the participant received an above-median or a below-median number of

[17] Home visitors changed in about 10% of cases, for example if home visitors moved or went into maternity leave.

home visits. Here, the mental health effect does not markedly vary with the treatment dosage. 8Finally, in row (vii) we break up the effects by the staffing model, i.e., whether home visits were done only by a medical professional (continuous model), or whether a medical professional alternated with a social pedagogue (tandem model). The results show no clear picture, with the effect on psycholeptics being larger in the tandem model, and the effect on depression not depending on the staffing model.

Table 8: Subgroup analysis on the mental health effects

		Subgroup I=0		Subgroup I=1		
	Obs.	*Treatm. Effect*	*p-value t-test*	*Treatm. Effect*	*p-value t-test*	*p-value for difference*
Subgroup indicator (I):		**Panel A: Psycholeptics**				
(i) Risk factor stress at baseline	244	-0.042	0.15	-0.138	**0.03**	0.14
(ii) More than 5 risk factors at baseline	244	-0.031	0.26	-0.114	**0.04**	0.20
(iii) Less than 5 friends	244	-0.051	0.17	-0.088	**0.05**	0.53
(iv) Social support scale: low	243	-0.055	**0.099**	-0.115	**0.07**	0.41
(v) No change in Home Visitor	244	-0.044	0.08	-0.072	**0.02**	0.11
(vi) Above-median no. of home visits	244	-0.063	**0.02**	-0.075	**0.03**	0.45
(vii) Single visitor model	244	-0.095	**0.03**	-0.050	**0.02**	0.19
		Panel B: Depression				
(i) Risk factor stress at baseline	344	-0.044	0.38	-0.265	**0.01**	**0.047**
(ii) More than 5 risk factors at baseline	344	-0.003	0.96	-0.201	**0.01**	**0.03**
(iii) Less than 5 friends	344	-0.105	**0.04**	-0.110	0.33	0.97
(iv) Social support scale: low	342	-0.070	0.15	-0.240	**0.07**	0.23
(v) No change in Home Visitor	344	-0.022	0.87	-0.115	**0.02**	0.46
(vi) Above-median no. of home visits	344	-0.128	**0.02**	-0.091	0.11	0.48
(vii) Single visitor model	344	-0.106	**0.08**	-0.107	**0.08**	0.98

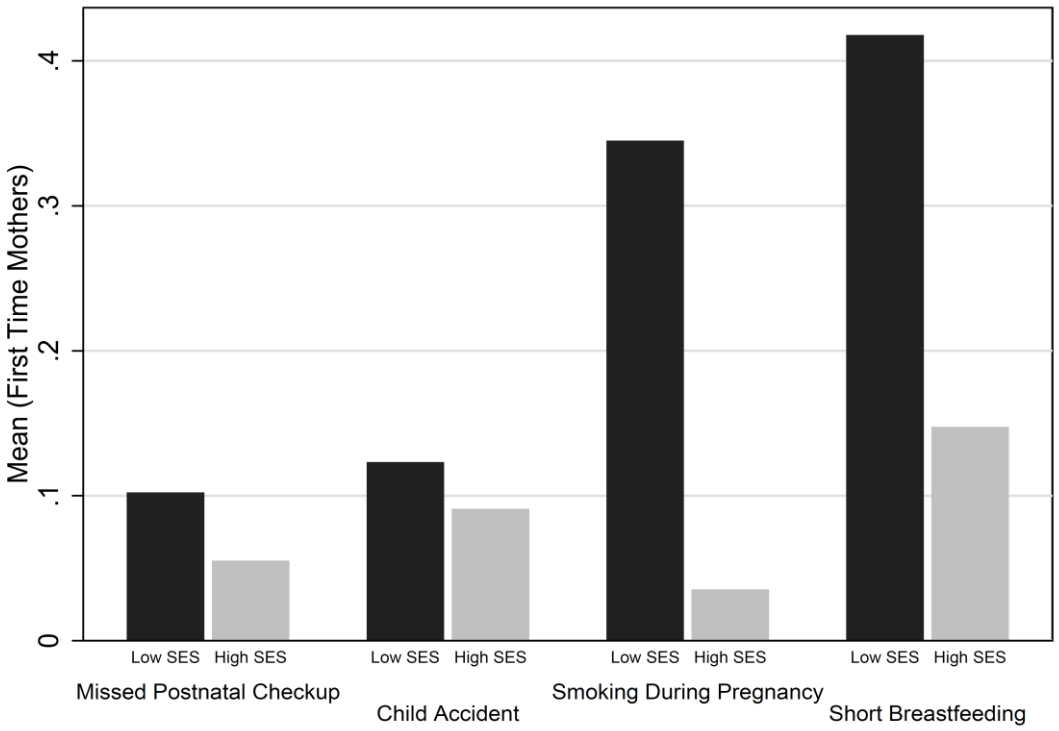
Notes: For a range of binary subgroup indicators *I* this table shows separate treatment effects for the groups defined by *I=0* and *I=1* for the outcomes of psycholeptics (Panel A) and depression at 24 months (Panel B). These are estimated by replacing the treatment indicator in regression equation (1) by its interactions with an indicator for *I=0* and an indicator for *I=1*. Treatment effects are estimated from IPW-weighted OLS regressions with control variables (including the subgroup indicator *I* among the controls where applicable). p-values below 0.1 are in bold. All p-values are two-tailed.

Data Source: Administrative data from the AOK public health insurance and *Pro Kind* participant survey.

It is worth noting that the analysis in Table 8 is exploratory in the sense that none of the effect differences can be interpreted as causal because, while treatment assignment is random, the interaction variables used in Table 8 are endogenous. For example, an alternative explanation of the finding that the mental health effects are bigger for women who had the same home visitor throughout the program could be that home visitors are more likely to quit from working with those women who respond less well to the program (i.e., who have worse mental health despite treatment). In that sense these explorative results should be viewed as a descriptive rather than causal decomposition of the treatment effects.

Why did the program fail to affect child health and several important domains that were explicitly targeted, such as maternal smoking behavior, breastfeeding, and prevention of accidents? A possible reason could be that in a European style public health insurance system, the disadvantaged target group already receives high levels of preventive and acute health care. As we illustrate in Figure 2, only 10 percent of low SES mothers in Germany miss any of the six recommended postnatal checkups in the first year of life, not dramatically different from the 5 percent of high SES mothers who do so.[18]

Figure 2: Maternal health behavior by socio-economic status



Notes: The figure reports means for indicators of maternal health behavior by socio-economic status (SES). SES is measured as sum of three metric components: education and occupational qualification, educational status, and net income (Lange et al. 2007).

Data Source: German Health Interview and Examination Survey for Children and Adolescents (KIGGS), baseline survey (Kurth 2007)

In addition, mothers can receive home visits during pregnancy and after birth by an obstetric nurse or midwife of her own choice, financed through the public health insurance. In our analysis we find that 73 percent of mothers in the control group utilize midwife consultations before birth, and 82 percent after birth (Table 2). These high take-up rates and the associated care and advice the

[18] Figure 2 is based on data from the German Health Interview and Examination Survey for Children and Adolescents (KIGGS) (Kurth 2007). Rates in the *Pro Kind* control group are very similar to the low SES-KIGGS sample as shown in the result section. Postnatal checkups are paid by health insurance but take-up is voluntary. Checkups 1 to 3 are recommended in the first month after birth, 4 to 6 in the next 11 months. The focus of the checkups is on age specific topics, such as mental and physical development, immunization or preventing sudden infant death syndrome.

mothers receive may be part of the reason why we find little effects on utilization, child health, and child accidents (which also do not differ markedly by SES—see Figure 2). In line with recent evidence from the UK and Ireland (Robling et al. 2016; Doyle et al. 2015), home visiting might thus be less effective when embedded in a European style public health insurance system that grants disadvantaged mothers comprehensive access to health care services. In the U.S., in contrast, where over one fifth (21.3 percent) of reproductive-aged females are uninsured, and 42 percent of women with less than high school diploma do not seek prenatal care in the first trimester of pregnancy (U.S. Department of Health and Human Services 2013), home visiting programs appear to have clearer beneficial effects on child health outcomes than in the European context (Avellar and Supplee 2013; Peacock et al. 2013).

Nevertheless, despite good health care access of our target group, Figure 1 also shows a clear SES gradient in adverse maternal health behavior. A high share of low-SES mothers in Germany smokes during pregnancy (35 percent), while only few high-SES mothers do so (4 percent). Low-SES mothers are also more likely to breastfeed for only a short period (less than one month) than high-SES mothers (42 percent versus 15 percent). This raises the question why the program did not lead to improvements in these areas. The main channels through which the *Pro Kind* program intended to affect maternal behavior are the transmission of information and the fostering of maternal self-efficacy promoted by the attachment between mother and home visitor. Table 6 shows a somewhat reduced level of maternal daily smoking during pregnancy (30%) in comparison to baseline (34%, see Table 1), a relatively low share of child passive smoking (11%), and relatively high rates of breast feeding *initiation* (only 23% in the control group do not breastfeed at all). These figures might suggest that the harm of smoking during pregnancy and the benefits of breastfeeding are widely known in the target population. Therefore, the program may not have transmitted new information in these areas. Moreover, changing behaviors such as smoking or breastfeeding is costly to the individual, as smoking is an addiction, and breastfeeding is time-consuming and requires perseverance in case of difficulties. Even if the program fostered self-efficacy through the attachment between mother and home visitor, as the mental health effects suggest, this may not be powerful enough to change these behaviors, in particular as the explicit policy of the program was not to put pressure on participants in order to maintain a close relationship between mother and home visitor to avoid withdrawal from the program.[19] The case is different for the utilization of dental health treatments, on which we find an effect. Here, the program may have transmitted new information on the risk of maternal oral infections to the unborn child (Sanz and Kornmann 2013; Vergnes and Sixou 2007), and taking up dental services

[19] The program was guided by the idea that participants are masters of their own life, with the home visitors merely offering support. For example, home visitors would only support participants in giving up smoking if participants explicitly stated this goal and asked for help. It was feared that any form of pressure would make the program less effective by harming the attachment between home visitor and participant.

in a publicly funded health care system which is largely free of charge at the point of use does not seem overly costly.

7. Conclusion

Based on a unique combination of survey and administrative data, this paper exploits a randomized controlled trial to provide a comprehensive evaluation of the effects of a home visiting program targeted towards disadvantaged pregnant women in Germany on a wide range of health outcomes. Overall, we find no effects on most types of health utilization, health behaviors, and a wide range of physical health measures, with the exception of some effects on the utilization of oral health care. Most interestingly, we find that the intervention had a positive effect on maternal mental health as evidenced by both, a reduction in prescriptions of psycholeptics recorded in the administrative data by 7 percentage points (against a control group mean of 7%), and a reduction in depressions reported in the survey data at age 2 of the child by 11 percentage points (against a control group mean of 28%). These effects seem to be particularly driven by mothers who entered the program with more elevated stress levels, more social risk factors, and less social support, and are stronger for women who had the same home visitor throughout all visits. This suggests that the mental health effect could be related to the personal relationship between mother and home visitor substituting for a lack of social support of the target group in a stressful life situation.

Our effects on mental health are in line with the few home visiting studies that also find positive effects on mental health. For example, the recent reanalysis of the NFP program by Heckman et al. (2017) reports improvement across several maternal mental health dimensions (anxiety, emotional stability, self-esteem, mastery skills) with effect sizes of 0.2 to 0.3 standard deviations. In one of their implementation counties of the Healthy Families New York (HFNY) program Mitchell-Herzfeld et al. (2005) find a reduction in the percentage of mothers classified with depression by fifteen percentage points, in a similar range as our 11 percentage-point effect on depression. Interestingly our results on mental health are also comparable to broader policies which did not primarily focus on this outcome such as the Moving to Opportunity program (MTO), in which disadvantaged families were offered vouchers to move to low-poverty neighborhoods (Ludwig et al. 2013), and the Oregon Health Insurance Experiment, which expanded access to health insurance (Finkelstein et al. 2012). In the second study, Medicaid coverage decreased the probability of a positive screening for depression by 9 percentage points (Baicker et al. 2013), very close to the effect sizes in our study.[20]

Overall, we conclude that home visiting targeted towards disadvantaged families has limited effectiveness in closing the SES gradient in child health, in particular if embedded within a

[20] Child benefit programs have also been shown to affect maternal mental health positively (Milligan and Stabile, 2011).

comprehensive public health insurance system with generally good health care access of the target group. This should be taken into consideration when a program is transferred from one institutional context to another. It is also unlikely to have strong effects on maternal health behaviors that are difficult to change, such as smoking and breastfeeding. Since the most important beneficial effect is on maternal mental health, this outcome should be routinely investigated in similar types of policy interventions.

The fact that the program failed to reach many of its goals despite its intensity (forty-five 90-minute meetings over two and a half years, with average costs of € 8,705, or approximately $ 9,575 per intervention—Maier-Pfeiffer et al. 2013) raises the question of how it could be improved. It seems likely that more limited interventions with specific goals (such as giving up smoking) could be more effective than a broad home-visiting program with multiple domains of intervention. At the same time, given that the participants are in frequent contact with the health care system, some elements of the program, such as the information on preventive dental health which seems to have been effective, could potentially be incorporated more strongly into routine pre- and postnatal checkups. It is likely, however, that the positive effect on maternal well-being and mental health was facilitated by the personal relationship between participant and home visitor fostered by the intensity of the program, and this might therefore be more difficult to achieve with alternative less intensive interventions.

References

Aizer, Anna; Stroud, Laura; Buka, Stephen (2016): Maternal Stress and Child Outcomes. Evidence from Siblings. In: J. Human Resources 51 (3), p. 523–555.

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

Aneshensel, Carol S. (2009): Toward Explaining Mental Health Disparities. In: Journal of Health and Social Behavior 50(4), p. 377–394.

Apouey, Bénédicte; Geoffard, Pierre-Yves (2013): Family income and child health in the UK. In: Journal of Health Economics 32(4), p. 715–727.

Avellar, Sarah A.; Supplee, Lauren H. (2013): Effectiveness of Home Visiting in Improving Child Health and Reducing Child Maltreatment. In: Pediatrics 132 (2), p. 90 -99.

Baicker et al. (2013): The Oregon Experiment — Effects of Medicaid on Clinical Outcomes. In: New England Journal of Medicine 368, p. 1713-1722.

Brand, Tilman; Jungmann, Tanja (2012): Implementation Differences of two Staffing Models in the German Home Visiting Programm "Pro Kind". In: Journal of Community Psychology 40 (8), p. 891-905.

Brand, Tilman; Adamaszek, Kristin; Hartmann, Susanne (2013): Dokumentation als Bestandteil der Qualitätssicherung Früher Hilfen. In: T. Brand und T. Jungmann (Eds.): Kinder schützen, Familien stärken. Erfahrungen und Empfehlungen für die Ausgestaltung Früher Hilfen aus der Pro Kind-Praxis und -Forschung. Weinheim: Beltz Juventa, p. 98–110.

BFSFJ (Bundesministerium für Familie, Senioren, Frauen und Jugend)(2016): Bundeshaushaltsplan 2016, Einzelplan 17 https://www.bundeshaushalt-info.de/fileadmin/de.bundeshaushalt/content_de/dokumente/2016/soll/epl17.pdf (Accessed October 26, 2017).

Carlson, Kyle (2015): Fear itself: The effects of distressing economic news on birth outcomes. In: Journal of Health Economics 41, p. 117–132.

Casamassimo, Paul S.; Thikkurissy, Sarat; Edelstein, Burton L.; Maiorini, Elyse (2009): Beyond the dmft: The Human and Economic Cost of Early Childhood Caries. In: The Journal of the American Dental Association 140 (6), p. 650-657.

Case, Anne; Lubotsky, Darren; Paxson, Christina (2002): Economic Status and Health in Childhood: The Origins of the Gradient. In: American Economic Review 92 (5), p. 1308–1334.

Condliffe, Simon; Link, Charles (2008): The Relationship between Economic Status and Child Health: Evidence from the United States. In: American Economic Review 4 (98), p. 1605–1618.

Cooper, Peter J.; Lynne Murray (1998): Postnatal depression. In: BMJ : British Medical Journal; London 316(7148), p. 1884.

Currie, Janet (2000): Child health in developed countries. In: Culyer, A., Newhouse, J.(Eds.), Handbook of Health Economics. North-Holland, Amsterdam

Currie, Janet (2009): Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development. In: Journal of Economic Literature 47 (1), p. 87–122.

Currie, Janet; Lin, Wanchuan (2007): Chipping Away at Health: More on the Relationship Between Income and Child Health. In: Health Affairs 26 (2), p. 331–344.

Currie, Janet; Stabile, Mark (2003): Socioeconomic Status and Child Health: Why Is the Relationship Stronger for Older Children? In: American Economic Review 93 (5), p. 1813–1823.

Currie, Janet; Stabile, Mark; Manivong, Phongsack; Roos, Leslie L. (2010): Child Health and Young Adult Outcomes. In: Journal of Human Resources 45 (3), p. 31–47.

Deaton, Angus (2002): Policy Implications Of The Gradient Of Health And Wealth. In: Health Affairs 21(2), p. 13–30.

Dennis, Cindy-Lee (2005): Psychosocial and psychological interventions for prevention of postnatal depression: systematic review. In: British Medical Journal 331(7507), p. 15.

Dohrenwend, Bruce P. (2000): The Role of Adversity and Stress in Psychopathology: Some Evidence and Its Implications for Theory and Research. In: Journal of Health and Social Behavior 41(1), p. 1–19.

Doorslaer, Eddy van; Koolman, Xander (2004): Explaining the differences in income-related health inequalities across European countries. In: Health Economics 13(7), p. 609–628.

Doyle, Orla; Fitzpatrick, Nick; Lovett, Judy; Rawdon, Caroline (2015): Early intervention and child physical health: Evidence from a Dublin-based randomized controlled trial. In: Economics and Human Biology 19, p. 224–245.

Duggan, Anne; McFarlane, Elizabeth; Fuddy, Loretta; Burrell, Lori; Higman, Susan M.; Windman, Amy; Sia, Calvin (2004): Randomized trial of a statewide home visiting program: impact in preventing child abuse and neglect. In: Child Abuse and Neglect, 28(6), p. 597-622.

Efron, Bradley (1971): Forcing a Sequential Experiment to be Balanced. In: Biometrika 58(3), p. 403-417.

Figlio, David; Guryan, Jonathan; Karbownik, Krzysztof; Roth, Jeffrey (2014): The Effects of Poor Neonatal Health on Children's Cognitive Development. In: American Economic Review 104 (12), p. 3921–3955.

Finkelstein, A.; Taubman, S.; Wright, B.; Bernstein, M.; Gruber, J.; Newhouse, JP.; Allen, H.; Baicker, K. (2012): The Oregon Health Insurance Experiment: Evidence From the First Year. In: The Quaterly Journal of Economics 127(3), p. 1057-1106

Gebhardt, Daniel (2010): Codebook and Documentation of the Panel Study 'Labour Market and Social Security' (PASS). In: FDZ-Datenreport.

Hayes, Andrew F. (1996): Permutation test is not distribution-free: Testing H_0: $\rho = 0$. In: Psychological Methods 1(2), p. 184–198.

Heckman, James J.; Holland, Margaret L.; Makino Kevin K.; Pinto, Rodrigo; Rosales-Rueda, Maria (2017): An Analysis of the Mempihs Nurse-Family Partnership Program. National Bureau of Economic Research. Working Paper http://www.nber.org/papers/w23610.pdf (Accessed October 26, 2017).

Heckman, James J.; Moon, Seong Hyeok; Pinto, Rodrigo; Savelyev, Peter; Yavitz, Adam (2010): Analyzing Social Experiments as Implemented: A Reexamination of the Evidence From the HighScope Perry Preschool Program. In: Quantitative Economics 1 (1), p. 1–46.

Hoynes, Hilary; Miller, Doug; Simon, David (2015): Income, the Earned Income Tax Credit, and Infant Health. In: American Economic Journal: Economic Policy 7(1), p. 172–211.

Johnston, Brian D.; Huebner, Colleen E.; Tyll, Lyinda L.; Barlow, William E.; Thompson, Robert S. (2004): Expanding developmental and behavioral services for newborns in primary care: effects on parental well-being, practice, and satisfaction. In: American Journal of Preventive Medicine 26(4), p. 356- 366.

Junge, Carolin; Garthus-Niegel, Susan; Slinning, Kari; Polte, Carolin; Simonsen, Tone Breines; Eberhard-Gran, Malin (2016): The Impact of Perinatal Depression on Children's Social-Emotional Development: A Longitudinal Study. In: Maternal and child health journal 21(3), p. 607-615

Jungmann, T., Kurtz, V., Brand, T., Sierau, S., & Von Klitzing, K. (2010). Präventionsziel Kindergesundheit im Rahmen des Modellprojektes „Pro Kind". Bundesgesundheitsblatt-Gesundheitsforschung-Gesundheitsschutz, 53(11), 1180-1187.

Jungmann, Tanja; Ziert, Yvonne; Kurtz, Vivien; Brand, Tilman (2009): Preventing Adverse

Developmental Outcomes and Early Onset Conduct Problems through Prenatal and Infancy Home Visitation: The German Pilot Project Pro Kind. In: European Journal of Developmental Science 3 (3), p. 292–298.

Kahan, Bennan C. and Morris, Tim P. (2012): Improper analysis of trials randomised using stratified blocks or minimisation. In: Statistics in Medicine, 31 (4), p. 328–340.

Koniak-Griffin, D.; Anderson, N.L.; Brecht, M.; Verzemnieks, I.; Lesser, J.; Kim, S. (2002): Public health nursing care for adolescent mothers: Impact on infant health and selected maternal outcomes at one year post birth. In: Journal of Adolescent Health 30, p. 44–54.

Kuehnle, Daniel (2014): The causal effect of family income on child health in the UK. In: Journal of Health Economics 36, p. 137–150.

Kurth, Bärbel-Maria (2007): The German Health Interview and Examination Survey for Children and Adolescents (KiGGS): an overview of its planning, implementation and results taking into account aspects of quality management. In: Bundesgesundheitsblatt 50 (5), p. 533–546.

Lange M.; Kamtsiuris P.; Lange C.; Schaffrath Rosario A,; Stolzenberg H.;, Lampert T. (2007): Der Kinder- und Jugendgesundheitssurvey (KiGGS): Messung soziodemographischer Merkmale im Kinder- und Jugendgesundheitssurvey und ihre Bedeutung am Beispiel der Einschätzung des allgemeinen Gesundheitszustands. Bundesgesundheitsblatt – Gesundheitsforschung – Gesundheitsschutz 2007 50 (5-6), p. 578–589.

Lange, Michael; Butschalowsky, H. G.; Jentsch, F.; Kuhnert, R.; Schaffrath Rosario, A.; Schlaud, M.; Kamtsiuris, P. (2014): The first KiGGS follow-up (KiGGS Wave 1): study conduct, sample design, and response. In: Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz 57 (7), p. 747–761.

List, John A.; Shaikh, Azeem M.; Xu, Yang (2016): Multiple Hypothesis Testing in Experimental Economics. National Bureau of Economic Research. Working Paper. http://www.nber.org/papers/w21875 (Accessed March 28, 2017).

Lovibond, S.H..; Lovibond, P.F. (1995): Manual for the Depression Anxiety Stress Scales. 2nd ed., Sydney: Psychology Foundation of Australia.

Ludwig, Jens; Duncan, Greg J.; Gennetian, Lisa A.; Kratz, Lawrence F.; Kessler, Ronald C.; Kling, Jeffrey R.; Sanbonmatsu, Lisa (2013): Long-Term Neigborhood Effects on Low-Income Families: Evidence from Moving to Opportunity. In: American Economic Review 103(3), p. 226-231.

Maier-Pfeiffer, Anna; Kutz, Kristina; Sandner, Malte (2013): Kosten, Nutzen und Finanzierung in den Frühen Hilfen. In: T. Brand und T. Jungmann (Hg.): Kinder schützen, Familien stärken. Erfahrungen und Empfehlungen für die Ausgestaltung Früher Hilfen aus der Pro Kind-Praxis und -Forschung. Weinheim: Beltz Juventa, p. 111–124.

Meyer, Karen; Khorshidi-Böhm, Mansoureh; Geurtsen, Werner; Günay, Hüsamettin (2014): An early oral health care program starting during pregnancy – a long-term study-Phase V.. In: Clinical Oral Investigations 18(3), p. 863–872.

Milligan, Kevin; Stabile, Mark (2011): Do Child Tax Benefits Affect the Well-being of Children? Evidence from Canadian Child Benefit Expansions. In: American Economic Journal: Economic Policy 3(3), p. 175–205.

Mitchell-Herzfeld, Susan; Izzo, Charles; Greene, Rose; Lee, Eunju; Lowenfels, Ann: Evaluation of Healthy Families New York (HFNY): First year program impacts New York State Office of Children & Family Services, Rensselaer, NY (2005).

Muntaner, Carles; Eaton, William W.; Miech, Richard; O'Campo, Patricia (2004): Socioeconomic Position and Major Mental Disorders. In: Epidemiologic Reviews 26(1), p. 53–62.

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

Peacock, Shelley; Konrad, Stephanie; Watson, Erin; Nickel, Darren; Muhajarine; Nazeem (2013): Effectiveness of home visiting programs on child outcomes: a systematic review. In: BMC Public Health 13, p. 17-31

Robling, Michael; Bekkers, Marie-Jet; Bell, Kerry; Butler, Christopher C.; Cannings-John, Rebecca; Channon, Sue et al. (2016): Effectiveness of a Nurse-Led Intensive Home-Visitation Programme for First-Time Teenage Mothers (Building Blocks): A Pragmatic Randomised Controlled Trial. In: The Lancet 387 (10014), p. 146–155.

Romano, Joseph P.; Wolf, Michael (2005a): Exact and Approximate Stepdown Methods for Multiple Hypothesis Testing. In: Journal of the American Statistical Association 100(469), p. 94–108.

Romano, Joseph P.; Wolf, Michael (2005b): Stepwise Multiple Testing as Formalized Data Snooping. In: Econometrica 73(4), p. 1237–1282.

Romano, Joseph P.; Wolf, Michael (2016). Efficient computation of adjusted p-values for resampling-based stepdown multiple testing. Statistics & Probability Letters 113, p. 38-40.

Sandner, Malte (2017): Effects of Early Childhood Intervention on Fertility and Maternal Employment: Evidence from a Randomized Controlled Trial. HCEO Working Paper https://econresearch.uchicago.edu/sites/econresearch.uchicago.edu/files/Sandner_2017_ECI-fertility-maternal-emp.pdf (Accessed October 26, 2017)

Sandner, Malte; Jungmann, Tanja (2017): Gender-specific effects of early childhood intervention. Evidence from a randomized controlled trial. In: Labour Economics, 45 (4), p. 59-78.

Sanz, Mariano; Kornmann, Kenneth (2013): Periodontitis and adverse pregnancy outcomes: consensus report of the Joint EFP/AAP Workshop on Periodontitis and Systemic Diseases. In: Journal of Periodontology 84, p. 164–169.

Sierau, Susan; Dähne, Verena; Brand, Tilman; Kurtz, Vivien; Klitzing, Kai von; Jungmann, Tanja (2016): Effects of Home Visitation on Maternal Competencies and Family Environment, and Child Development: a Randomized Controlled Trial. In: Prevention Science 16 (7), p. 1–12.

Trappmann, Marc; Bode, Jonas; Bethmann, Arne; Müller, Gerrit (2013): The PASS panel survey after six waves. In: Journal for Labour Market Research, 46:275-281.

Turner, R. Jay; Wheaton, Blair ; and Lloyd, Donald A. (1995): The Epidemiology of Social Stress. In: American Sociological Review 60(1), p. 104–125.

U.K. Department of Health (2013): Family Nurse Partnership programme to be extended: https://goo.gl/OZZbnJ (Accessed October 26, 2017).

U.S. Department of Health and Human Services (2015): HHS Awards $386 Million to Support Families Through the Home Visiting Program: http://www.hhs.gov/news/press/2015pres/02/20150219a.html (Accessed October 26, 2017).

Vergnes, Jean-Noël; Sixou, Michel (2007): Preterm low birth weight and maternal periodontal status: a meta-analysis. In: American Journal of Obstetrics and Gynecology 196, p. 135.e1-135.e7.

Wildman, John (2003): Income related inequalities in mental health in Great Britain: analysing the causes of health inequality over time. In: Journal of Health Economics 22(2), p. 295–312.