Institute for Fiscal Studies

Gabriella Conti
Sören Kliem
Malte Sandner

Working paper
24/34

Early home visiting delivery model and maternal and child mental health at primary school age

Economic and Social Research Council

Early Home Visiting Delivery Model and Maternal and Child Mental Health at Primary School Age

Gabriella Conti, Sören Kliem, and Malte Sandner*

July 2, 2024

Abstract

We study the impacts of a prenatal and infancy home visiting program targeting disadvantaged families on mental health outcomes, assessed through diagnostic interviews. The program significantly reduced the prevalence of mental health conditions for both mothers and children, measured at primary-school age, and broke the intergenerational association of these conditions. The impacts are predominantly associated with a particular delivery model, wherein a single home visitor interacts with the family, as opposed to a model involving two home visitors.

JEL Classification: I12, J21, J13, J16

Keywords: Home Visiting, Mental Health, Diagnostic Interviews

*Conti: University College London (UCL) and Institute for Fiscal Studies (IFS) (email: gabriella.conti@ucl.ac.uk); Kliem: University for Applied Science Jena (email: soeren.kliem@eah-jena.de); Sandner: Nuernberg Institute for Technology (email: malte.sandner@th-nuernberg.de. RCT ID AEARCTR-0002135. Gabriella Conti gratefully acknowledges financial support from the European Research Council for grant agreement no. 819752 - DEVORHBIOSHIP - ERC-2018COG. Gabriella Conti and Malte Sandner gratefully acknowledge financial support from the Nuffield Foundation (award EDO/43715). Sören Kliem and Malte Sandner gratefully acknowledge financial support from the German Federal Ministry of Education and Research [grant numbers: 01EL1408].

1 Introduction

Mental health conditions are on the rise among both adults and children (AAP-AACAP-CHA, 2021). Public health surveillance focuses more on access to treatment than on preventable risk factors in early life, such as adverse childhood experiences (ACEs) and social determinants (Shim et al., 2022). Early life interventions have been shown to improve physical health and to prevent the emergence of disease up to adulthood (Campbell et al., 2014; Hjort et al., 2017), however the evidence is more limited on the prevention of mental health conditions (Singh et al., 2022). Existing work is mostly based on self- or parental reports (Carneiro and Ginja, 2014), diagnoses (Chorniy et al., 2020), and hospitalizations (Cattan et al., 2022), and does not investigate the key features of effective interventions.

In this paper, we contribute to this literature by exploiting two distinctive features of the German home visiting program Pro Kind (henceforth, PK). First, PK measured mental health objectively trough diagnostic interviews for both children and mothers at the child's primary school age. Second, PK was implemented according to two delivery models: one comprises home visits exclusively conducted by family midwives (Continuous Model, henceforth, CM), while the other entails home visits carried out by social workers and family midwives in tandem (Tandem Model, henceforth, TM).

2 The Pro Kind Program

PK is an adaptation of the widely implemented US Nurse Family Partnership (NFP) home visiting program (Olds, 2006). In accordance with the NFP, visits begin in pregnancy and continue until the child's second birthday. Across this time frame, the frequency of home visits varies among weekly, biweekly, and monthly according to the NFP model, for an overall maximum of 52 home visits with an average duration of 90 minutes each. Teaching materials and visit-by-visit guidelines, adapted from the NFP, structure the goals and content of each home visit.

PK was implemented from 2006 to 2012 in three German federal states (Bremen, Lower Saxony, and Saxony) in 13 sites. PK eligibility criteria were (1) women with a first-time pregnancy from the 12th to the 28th week of gestation, (2) women with an economic risk (e.g., receiving welfare benefits), (3) women with at least one further social or personal risk. Overall, 755 mothers were randomly assigned to the treatment group (TG) (n=394) or to the control group (CG) (n=361). The baseline randomization was stratified by site, maternal age (<18 vs >=18 years old), and maternal nationality (German versus non-German). Frequency of social or personal risk factors at the baseline assessment did not statistically significant differ between the groups (Sandner et al., 2018).

The PK research team collected extensive data from pregnancy to 7 years post-birth, the last follow-up when for the first time diagnostic interviews were used to measure mental health in both mothers and children. The data collected in infancy and toddlerhood has shown that PK improved cognitive development, mother-child interactions, and maternal investments, with effects concentrated on girls (Conti et al., 2021). While up to 24 months after birth the program had no effects on child or maternal physical health and health behaviors (e.g., breastfeeding or smoking), it reduced maternal depressive issues – measured by a self-reported screening questionnaire – and prescriptions of psycholeptics (Sandner et al., 2018). A previous analysis of the 7-year follow-up data has confirmed program impacts on maternal mental health, and showed also fewer behavioral problems among children in the treatment group - both outcomes measured by self-reported screening questionnaires (Kliem and Sandner, 2021). The data from the PK diagnostic interviews has not been analyzed to date.

Innovating upon the original NFP-program, the German PK implemented two different types of delivery models, depending on the geographical site. In the CM, families were visited by the same home visitor, a midwife, throughout the whole intervention period (from pregnancy to the child's second birthday). In the TM, a midwife conducted the home visits only from pregnancy to six months of age of the child, in tandem with a social worker; she was then entirely replaced by the social worker, who conducted the home visits up to the child's second birthday (see for details Brand and Jungmann, 2012). Differences between the PK model types have not been evaluated systematically before.

3 Data & Empirical Approach

To improve the measurement of child and maternal mental health during the 7-year follow-up, the PK research team implemented the Diagnostic Interview of Mental Disorders in Children and Youths (DIPS) and the Expert System for Diagnosis of Mental Disorders (DIA-X), respectively. The DIPS and DIA-X are semi- or fully structured diagnostic procedures to measure mental disorders based on the International Classification of Diseases (ICD) 10. Trained psychological students blinded to treatment status conducted all clinical interviews at the home of the families. The DIPS was used to diagnose mental health conditions with ICD-10 codes F 40-48 (Neurotic, stress-related and somatoform disorders) and F 90-98 (Behavioural and emotional disorders with onset usually occurring in childhood and adolescence). The DIA-X covered the mental health conditions with ICD-10 codes F 30-39 (Mood [affective] disorders (Depression)) and F 40-48 (Neurotic, stress-related and somatoform disorders) (see for more details Kliem et al., 2018).

In contrast to mental health screening questionnaires commonly used in surveys, such as the Strength and Difficulties Questionnaire (SDQ), the Child Behavior Check List (CBCL) or the Depression, Anxiety, Stress

Scale (DASS), which indicate potential issues or areas of concerns, diagnostic interviews lead to a diagnosis, which indicates a mental health condition. To achieve this, it is necessary a more detailed exploration of symptoms and their duration, severity, and impact.

The DIPS and DIA-X diagnoses are only available for a subsample, for the following reasons. First, we excluded from the analysis one site, in which both the CM and the TM were implemented. Second, due to financial constraints, it was not possible to conduct the diagnostic interviews with all families. Finally, some scheduled at home interviews could not take place as planned, because the family moved or the family withdrew the consent to participate in the research.

These restrictions left us with a sample of 150 children (TG: 77, CG: 73) with a DIPS interview in CM sites and 129 children (TG: 67, CG: 62) in TM sites and 162 mothers (TG: 80, CG:82) with a DIA-X interview in CM sites and 144 mothers (TG: 74, CG: 70) in TM sites. Appendix Tables A and B demonstrate that for the DIPS and DIA-X samples TG and CG baseline characteristics are overall balanced (Columns 1 to 4). Additionally, participants in TM and CM sites are overall balanced with respect to baseline characteristics, too (Columns 5 to 6); this is the case for both individual variables and summary measures of risk.

We estimate the treatment effects using augmented inverse probability weighting (AIPW) in combination with lasso to select relevant control variables from those which are listed in Appendix Tables A and B (e.g. Chernozhukov et al., 2018).

To explain the approach, let $Y_i(d)$ denote the outcome under the treatment status $d \in \{0, 1\}$, with the expectation modeled as $m_d(X_i) = E(Y_i(d)|X_i)$. Let D_i be a binary variable denoting the treatment assignment, whose propensity score is $p(W_i) = \Pr(D_i = 1|W_i)$. W_i is a set of control variables potentially equal to X_i. Let $\widehat{m}_d(X_i)$ and $\widehat{p}(W_i)$ be the estimated models of $m_d(X_i)$ and $p(W_i)$. The AIPW estimator is then given by

$$\widehat{ATE}_{AIPW} = \frac{1}{n}\sum_{i=1}^{N}\Bigg\{\left[\frac{D_i Y_i}{\widehat{p}(W_i)} - \frac{(1-D_i)Y_i}{1-\widehat{p}(W_i)}\right] - \frac{(D_i - \widehat{p}(W_i))}{\widehat{p}(W_i)(1-\widehat{p}(W_i))} \times [(1-\widehat{p}(W_i))\widehat{m}_1(X_i) + \widehat{p}(W_i)\widehat{m}_0(X_i)]\Bigg\} \quad (1)$$

where $\widehat{m}_d(X_i)$ and $\widehat{p}(W_i)$ are specified as linear and logit functions, respectively. X_i and W_i are selected through a lasso procedure. The strata variables are always included in $\widehat{p}(W_i)$. Standard errors are clustered

at the level of the site.

4 Results

4.1 Mental Health Disorders

Table 1 shows in Columns 2 and 4 the prevalence of mental health diagnoses for treatment and control groups, respectively. For all diagnoses, we fail to reject equality of means between CM and TM in the control group. This further reassures us about the comparability of the two delivery models.

For the CM (Panel A), the prevalence of the diagnoses F-90 for children and F-30 for mothers is lower in the treatment group than in the control group. The AIPW treatment effects (see equation 1) in Column 5 show a significant reduction by 5.1 p.p. for the diagnosis F-90 in children and a 10.3 p.p. reduction for the F-30 diagnosis in mothers. These reductions correspond to a 27% decrease for F-90 diagnoses and to a 42% decrease for F-30 diagnoses with respect to the mean of the control group. In the TM (Panel B) treatment and control groups do not differ significantly in the prevalence of any diagnoses.

Table 1: Child & Mother Mental Health - Clinical Interviews

Outcome	N TG	Mean TG	N CG	Mean CG	TE	p-value
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Continuous Model						
ICD F-40 - Child	77	0.078	73	0.068	0.006	0.837
ICD F-90 - Child	77	0.143	73	0.192	-0.051	**0.004**
ICD F-30 - Mother	80	0.150	82	0.244	-0.103	**0.029**
ICD F-40 - Mother	80	0.375	82	0.378	0.005	0.882
Panel B: Tandem Model						
ICD F-40 - Child	67	0.030	62	0.032	-0.006	0.847
ICD F-90 - Child	67	0.194	62	0.193	0.009	0.834
ICD F-30 - Mother	74	0.189	70	0.243	-0.017	0.664
ICD F-40 - Mother	74	0.405	70	0.386	0.024	0.815

Notes: Diagnoses for children are obtained by DIPS. Diagnoses for mothers are obtained by DIAX. All outcomes are binary. Treatment effects estimated via AIPW. Standard errors clustered at the site level. p-values <0.10 bolded. TG = Treatment group, CG = Control group, TE= Treatment effect.
ICD F-40 Child includes only F40.2 (Specific (isolated) phobias: Animal type; Natural environment type; Blood, injection, injury type; Situational type; Other specified phobia).
ICD F-90 Child includes F90 (Hyperkinetic disorders, ADHD), F91 (Conduct disorders), F93 (Emotional disorders), F98 (Other behavioural and emotional disorders with onset usually occurring in childhood and adolescence).
ICD F-30 Mother includes F30 (Manic episode), F31 (Bipolar affective disorder), F32 (Depressive episode), F33 (Recurrent depressive disorder), F34 (Persistent mood [affective] disorders).
ICD F-40 Mother includes F40 (Phobic anxiety disorders), F41 (Other anxiety disorders), F42 (Obsessive-compulsive disorders), F43 (Reaction to severe stress), F45 (Somatoform disorders).

4.2 Intergenerational Correlations

Besides investigating whether home visiting can prevent mental health disorders, the availability of ICD F-40-related clinical interviews for 257 mother-child pairs allows us to examine possible intervention impacts on intergenerational mental health associations (Bütikofer et al., 2023).

The results are presented in Table 2, which shows a significant and positive association between mother and child F-40 diagnoses for the pairs in the control group and in the TM. In contrast, for mother-child pairs who received the home visiting in the CM the association between maternal and child F-40 diagnoses is reversed: If a mother shows a neurotic, stress-related and somatoform disorder, it is significantly less likely that her child manifests this disorder. Hence, while the intervention did not reduce ICD F-40 diagnoses for either children or mothers, it broke the intergenerational correlation of these disorders, but only in the CM.

4.3 Implementation of the Two Models

Why was the continuous model more effective at reducing the prevalence of mental health conditions? We now investigate differences in implementation as possible explanations of the different performance of the two models.

A first reason might be differences in qualifications between the two professions: in Germany, midwives undergo a 3-year apprenticeship program, encompassing 1600 hours of theoretical training and 3000 hours of practical experience; on the other hand, social workers typically possess at least a 3-year Bachelor's degree. A second potential explanation might revolve around differences in the content of the home visits: however, as Appendix Figure A shows, using data from standardized encounter forms of each home visit, home visitors in both models covered the same topics to a comparable extent.

A third potential reason might be that mothers in the TM establish a less strong relationship with the home visitor than in the CM. This is suggested by Figure 1, which shows a substantial decline in the average number of visits in the TM after 6 months post-birth, at the time when the social worker entirely replaces the midwife. Additionally, Appendix Table C shows that, while the majority (57%) of families was visited by one home visitor in the CM, the majority (60.9%) of families in the TM was visited by three home visitors: this might have made more challenging to establish a closer relationship with mothers in the TM.

A final explanation, not directly testable here, is that mothers might have more difficulty in establishing a trusting relationship with social workers than with midwives, because of the social workers' connection with the Child Protection Services.

Table 2: Intergenerational Correlations in Mental Health - Clinical Interviews

	CG	**TG Continuous Model**	**TG Tandem Model**
	ICD F-40 Child (DIPS)		
	(1)	(2)	(3)
ICD F-40 Mother (DIAX)	0.084**	-0.085**	0.083*
	(0.031)	(0.031)	(0.040)
N	127	68	62

Notes: Each column presents the coefficient of a regression of child ICD F-40 diagnosis on mother ICD F-40 diagnosis (Neurotic, stress-related and somatoform disorders), respectively for the control group (column 1), treatment group in the CM (column 2) and treatment group in the TM (column 3). TG = Treatment group. CG = Control Group. Standard errors in parentheses clustered at the site level. **=p-value<0.05, *=p-value<0.10.

5 Conclusion

We have exploited two distinctive features of a German early home visiting program - two staffing models of implementation and the use of clinical diagnostic interviews for mental health assessment - to examine the potential of preventive public health interventions to reduce the prevalence of mental disorders in both mothers and children at primary school age.

We have shown persistent effects of prenatal and infancy home visiting on mental health disorders of children and mothers, five years after the program ended, concentrated in the delivery model in which a single home visitor (a midwife) interacts with the families. The early home visiting program also broke the intergenerational association of mental health disorders between mothers and children - again, only in the single home visitor model.

Possible explanations for the differential effectiveness of the two delivery models likely involve a greater difficulty in forming a closer relationship with the mother in the presence of multiple home visitors. Our findings have relevant implications for the prevention of mental disorders and the design of public health programs.

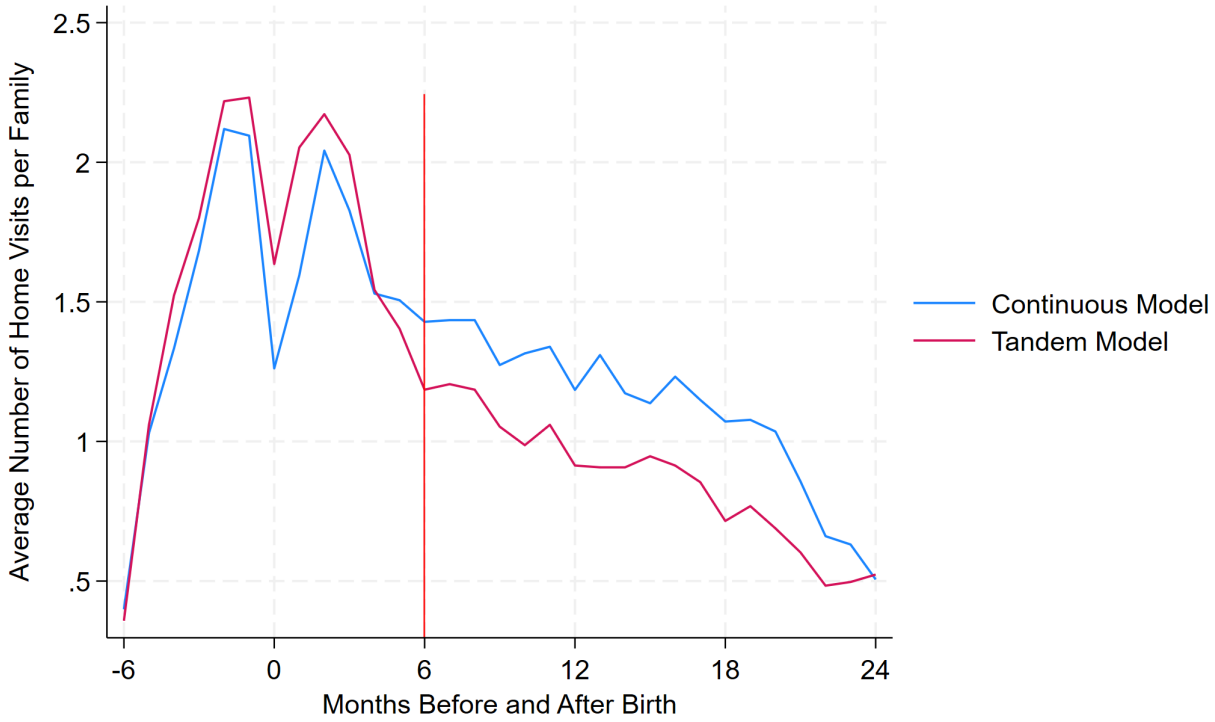


Figure 1: Average Number of Home Visits per Family per Month - Continuous and Tandem Model
The figure shows the average number of home visits in the 6 months prior to birth and in the 24 months after birth for CM and TM. The data includes a balanced sample of 319 families for each month (168 in the CM and 151 in the TM), irrespective of whether they received a visit in that specific month (no visits are coded as 0). Families from the site where both delivery models were implemented are excluded. [Source] Standardized encounter forms filled in by the home visitors for each home visit.

References

AAP-AACAP-CHA (2021). Declaration of a national emergency in child and adolescent mental health. *American Academy of Pediatrics, American Academy of Child & Adolescent Psychiatry, Children's Hospital Association*.

Brand, T. and Jungmann, T. (2012). Implementation differences of two staffing models in the german home visiting program "pro kind". *Journal of Community Psychology*, 40(8):891–905.

Bütikofer, A., Ginja, R., Karbownik, K., and Landaud, F. (2023). (breaking) intergenerational transmission of mental health. *Journal of Human Resources*, forthcoming.

Campbell, F., Conti, G., Heckman, J. J., Moon, S. H., Pinto, R., Pungello, E., and Pan, Y. (2014). Early childhood investments substantially boost adult health. *Science*, 343(6178):1478–1485.

Carneiro, P. and Ginja, R. (2014). Long-term impacts of compensatory preschool on health and behavior: Evidence from head start. *American Economic Journal: Economic Policy*, 6(4):135–173.

Cattan, S., Conti, G., Christine, F., Ginja, R., and Pecher, M. (2022). The health effects of universal early childhood interventions: evidence from sure start. *IFSWP*, 43.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., and Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1):C1–C68.

Chorniy, A., Currie, J., and Sonchak, L. (2020). Does prenatal wic participation improve child outcomes? *American Journal of Health Economics*, 6(2):169–198.

Conti, G., Poupakis, S., Sandner, M., and Kliem, S. (2021). The effects of home visiting on mother-child interactions: evidence from a randomized trial using dynamic micro-level data. *Child Abuse & Neglect*, 115:105021.

Hjort, J., Sølvsten, M., and Wüst, M. (2017). Universal investment in infants and long-run health: Evidence from denmark's 1937 home visiting program. *American Economic Journal: Applied Economics*, 9(4):78–104.

Kliem, S. and Sandner, M. (2021). Prenatal and infancy home visiting in germany: 7-year outcomes of a randomized trial. *Pediatrics*, 148(2).

Kliem, S., Sandner, M., Lohmann, A., Sierau, S., Dähne, V., Klein, A. M., and Jungmann, T. (2018). Follow-up study regarding the medium-term effectiveness of the home-visiting program "pro kind" at age 7 years: study protocol for a randomized controlled trial. *Trials*, 19(1):1–20.

Olds, D. L. (2006). The nurse–family partnership: An evidence-based preventive intervention. *Infant Mental Health Journal*, 27(1):5–25.

Sandner, M., Cornelissen, T., Jungmann, T., and Herrmann, P. (2018). Evaluating the effects of a targeted home visiting program on maternal and child health outcomes. *Journal of health economics*, 58:269–283.

Shim, R., Szilagyi, M., and Perrin, J. M. (2022). Epidemic rates of child and adolescent mental health disorders require an urgent response. *Pediatrics*, 149(5).

Singh, V., Kumar, A., and Gupta, S. (2022). Mental health prevention and promotion—a narrative review. *Frontiers in Psychiatry*, 13:898009.

Appendix Table A: Balance Table Sample DIPS

	Baseline Diff. TG vs CG Continuous Model		**Baseline Diff. TG vs CG Tandem Model**		**Baseline Diff. CM vs TM**	
Baseline Characteristics	Difference	*p*-value	Difference	*p*-value	Difference	*p*-value
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Risk factor: underage	0.01	0.87	0.00	0.88	-0.02	0.67
Risk factor: low income	0.03	0.63	-0.09	0.29	0.02	0.78
Risk factor: unwanted pregnancy	0.02	0.43	-0.03	0.67	0.00	0.93
Risk factor: social isolation	-0.03	0.69	0.00	0.90	0.00	0.83
Risk factor: custodial care	0.08	0.08	0.02	0.78	0.07	0.08
Risk factor: loss of significant other	-0.12	0.13	-0.13	0.21	0.01	0.81
Risk factor: violence in pregnancy	-0.02	0.76	0.00	0.93	0.04	0.14
Risk factor: mental illness	-0.09	0.27	-0.23	0.01	-0.04	0.46
Risk factor: aggression	-0.09	0.21	-0.06	0.15	-0.02	0.64
Risk factor: depression (DASS)	0.03	0.43	0.03	0.47	0.02	0.65
Risk factor: anxiety (DASS)	0.02	0.81	-0.07	0.35	0.06	0.08
Risk factor: stress (DASS)	-0.08	0.04	0.10	0.38	0.01	0.78
Risk factor: low education status	0.08	0.41	0.07	0.56	0.04	0.59
Risk factor: no partnership	0.06	0.25	0.15	0.12	-0.01	0.83
Teenager (below 20 years)	-0.02	0.86	0.00	1.00	-0.02	0.83
Lives alone	0.08	0.05	-0.17	0.06	-0.05	0.33
Lives with partner or husband	-0.02	0.69	0.02	0.85	0.05	0.43
Marital status: married	0.09	0.04	0.09	0.09	0.02	0.59
First pregnancy	-0.02	0.76	-0.01	0.54	0.03	0.46
Smokes	-0.07	0.66	0.02	0.91	0.04	0.70
Smokes daily	-0.06	0.54	-0.01	0.94	0.04	0.41
Student	-0.05	0.49	-0.03	0.62	-0.01	0.86
Medical high risk pregnancy	-0.02	0.55	0.06	0.48	-0.05	0.15
Self-efficacy	0.01	0.85	0.00	0.98	-0.02	0.73
Weight before pregnancy	0.45	0.86	3.53	0.19	2.00	0.20
Height	0.01	0.15	-0.02	0.17	0.01	0.47
BMI	-0.17	0.83	1.70	0.10	0.33	0.58
Obesity, BMI ¿ 30	-0.01	0.85	0.08	0.15	0.02	0.66
Underweight, BMI ¡ 17	0.00	0.97	-0.04	0.26	-0.03	0.16
Attachment	0.02	0.79	-0.08	0.20	0.03	0.48
Parental self-efficacy	0.04	0.18	-0.01	0.73	0.04	0.43
Household size	-0.12	0.59	0.19	0.38	0.00	0.99
Social class index	0.17	0.61	-0.33	0.60	-0.08	0.71
Year affiliated: 2006	0.00		0.04	0.07	-0.05	0.05
Year affiliated: 2007	0.03	0.39	-0.03	0.68	0.00	0.97
Year affiliated: 2008	-0.02	0.52	0.06	0.46	-0.01	0.69
Year affiliated: 2009	-0.01	0.88	-0.07	0.21	0.06	0.61
Month affiliated: January	0.00	0.84	0.05	0.37	-0.01	0.73
Month affiliated: February	-0.02	0.55	0.01	0.78	0.03	0.36
Month affiliated: March	-0.02	0.60	-0.01	0.92	0.02	0.51
Month affiliated: April	0.10	0.05	-0.10	0.09	0.04	0.41
Month affiliated: May	-0.01	0.92	0.01	0.92	-0.01	0.77
Month affiliated: June	0.00	0.76	0.08	0.02	-0.04	0.26
Month affiliated: July	0.06	0.56	0.00	0.90	0.08	0.04
Month affiliated: August	-0.07	0.14	0.03	0.53	0.00	0.93
Month affiliated: September	-0.06	0.04	0.04	0.49	-0.03	0.40
Month affiliated: October	0.04	0.18	0.01	0.69	0.01	0.67
Month affiliated: November	0.02	0.52	-0.07	0.03	-0.04	0.25
Month affiliated: December	-0.04	0.20	-0.05	0.32	-0.04	0.33
Number of weeks in pregnancy	0.02	0.67	0.02	0.83	-0.03	0.67
Mother non-German nationality	0.01	0.77	-0.02	0.39	0.08	0.06
Risk factors Total	**-0.38**	**0.10**	**-0.23**	**0.73**	**0.26**	**0.18**
Observations	150		129		279	

Notes: Columns 2, 4 and 6 present the coefficients of regressions of the variables listed in the first column on: the treatment dummy for the CM sample (column 2); the treatment dummy for the TM sample (column 4); a dummy for the CM (versus the TM, column 6). Columns 3, 5 and 7 present the corresponding *p*-values. Standard errors are clustered at the site level. Samples restricted to observations with DIPS outcomes. TG = Treatment Group. CG = Control Group. Diff. = Difference. CM = Continuous Model. TM = Tandem Model. All risk factors are binary variables. *Risk factor: low income* is 1 if participants' net household income is below 1250 Euro. The characteristic *Smokes* is a categorical variable with the outcomes never, sometimes, daily. *Attachment* is a characteristic measuring feelings about the expected baby by the Maternal Antenatal/Postnatal Attachment Questionnaire using a 4-point Likert scale with 19 items. *Parental self-efficacy* is a continuous characteristic obtained by the parental expectations survey (Reece and Harkless, 1998) using a 4-point Likert scale with 25 items. *Self-efficacy* characteristic obtained by the German adaptation of the Generalized Self-Efficacy Scale (Dumont et al., 2000) using a 4-point Likert scale with 10 items.

Appendix Table B: Balance Table Sample DIA-X

	Baseline Diff. TG vs CG Continuous Model		**Baseline Diff. TG vs CG Tandem Model**		**Baseline Diff. CM vs TM**	
Baseline Characteristics (1)	Difference (2)	*p*-value (3)	Difference (4)	*p*-value (5)	Difference (6)	*p*-value (7)
Risk factor: under age	-0.01	0.75	0.03	0.41	-0.03	0.55
Risk factor: low income	0.02	0.66	-0.07	0.50	0.02	0.80
Risk factor: unwanted pregnancy	0.00	0.74	-0.01	0.87	0.02	0.81
Risk factor: social isolation	-0.02	0.49	0.02	0.61	-0.01	0.72
Risk factor: custodial care	0.01	0.81	0.04	0.59	0.09	0.08
Risk factor: loss of significant other	-0.12	0.08	-0.08	0.33	0.02	0.40
Risk factor: violence in pregnancy	-0.01	0.81	0.03	0.50	0.05	0.11
Risk factor: mental illness	-0.11	0.18	-0.16	0.03	-0.02	0.72
Risk factor: aggression	-0.09	0.09	-0.06	0.05	-0.01	0.88
Risk factor: depression (DASS)	-0.01	0.86	0.00	0.90	0.03	0.47
Risk factor: anxiety (DASS)	0.05	0.24	-0.06	0.31	0.09	0.03
Risk factor: stress (DASS)	-0.09	0.01	0.07	0.59	-0.04	0.24
Teenager (below 20 years)	-0.08	0.32	-0.02	0.73	-0.03	0.70
Risk factor: low education status	0.07	0.43	0.03	0.71	0.03	0.74
Risk factor: no partnership	0.00	0.93	0.12	0.05	-0.03	0.60
Lives alone	0.02	0.61	-0.20	0.01	-0.06	0.22
Lives with partner or husband	0.05	0.32	0.06	0.42	0.07	0.36
Marital status: married	0.10	0.05	0.08	0.11	0.03	0.44
First pregnancy	-0.03	0.54	-0.07	0.13	0.02	0.57
Smokes daily	-0.03	0.64	-0.05	0.65	0.03	0.55
Smokes	-0.03	0.66	-0.07	0.73	0.06	0.62
Student	-0.06	0.03	-0.05	0.20	-0.02	0.59
Medical high risk pregnancy	-0.02	0.64	0.09	0.28	-0.06	0.20
Self-efficacy	0.03	0.64	0.00	0.97	0.02	0.78
Weight before pregnancy	0.37	0.88	3.12	0.19	2.95	0.03
Height	0.01	0.14	-0.01	0.29	0.01	0.46
BMI	-0.23	0.75	1.46	0.07	0.71	0.18
Obesity, BMI ¿ 30	-0.01	0.81	0.06	0.20	0.03	0.32
Underweight, BMI ¡ 17	-0.01	0.66	-0.06	0.01	-0.02	0.29
Attachment	0.07	0.24	-0.09	0.14	0.06	0.12
Parental self-efficacy	0.08	0.01	-0.04	0.63	0.06	0.23
Household size	-0.10	0.46	0.12	0.56	0.08	0.74
Social class index	0.13	0.68	-0.01	0.99	0.03	0.90
Year affiliated: 2006	0.00		0.04	0.23	-0.06	0.04
Year affiliated: 2007	0.07	0.02	-0.01	0.86	0.00	1.00
Year affiliated: 2008	-0.05	0.33	-0.01	0.89	0.01	0.71
Year affiliated: 2009	-0.02	0.81	-0.02	0.77	0.05	0.64
Month affiliated: January	0.00	0.93	0.04	0.48	-0.04	0.35
Month affiliated: February	-0.02	0.56	0.04	0.29	0.04	0.17
Month affiliated: March	-0.05	0.30	-0.05	0.52	0.01	0.81
Month affiliated: April	0.06	0.04	-0.07	0.14	0.05	0.29
Month affiliated: May	-0.01	0.84	0.01	0.86	-0.01	0.74
Month affiliated: June	-0.01	0.78	0.06	0.07	0.00	0.98
Month affiliated: July	0.06	0.45	-0.02	0.64	0.07	0.04
Month affiliated: August	-0.06	0.10	0.05	0.19	0.01	0.73
Month affiliated: September	-0.01	0.75	0.05	0.46	-0.03	0.41
Month affiliated: October	0.03	0.27	0.00	0.93	-0.03	0.16
Month affiliated: November	0.04	0.29	-0.06	0.04	-0.04	0.18
Month affiliated: December	-0.04	0.21	-0.05	0.37	-0.02	0.55
Number of weeks in pregnancy	0.04	0.52	0.07	0.12	0.00	0.97
Mother non-German nationality	0.00	0.95	-0.02	0.66	0.06	0.13
Risk factors total	**-0.60**	**0.03**	**-0.07**	**0.92**	**0.28**	**0.29**
Observations	162		144		306	

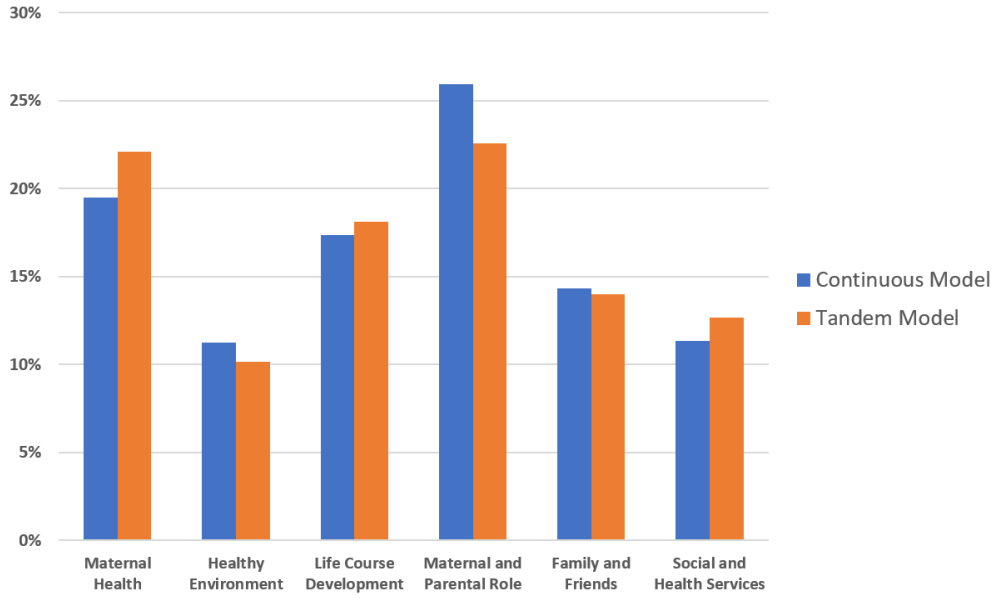
Notes: Columns 2, 4 and 6 present the coefficients of regressions of the variables listed in the first column on: the treatment dummy for the CM sample (column 2); the treatment dummy for the TM sample (column 4); a dummy for the CM (versus the TM, column 6). Columns 3, 5 and 7 present the corresponding *p*-values. Standard errors are clustered at the site level. Samples restricted to observations with DIPS outcomes. TG = Treatment Group. CG = Control Group. Diff. = Difference. CM = Continuous Model. TM = Tandem Model. All risk factors are binary variables. *Risk factor: low income* is 1 if participants' net household income is below 1250 Euro. The characteristic *Smokes* is a categorical variable with the outcomes never, sometimes, daily. *Attachment* is a characteristic measuring feelings about the expected baby by the Maternal Antenatal/Postnatal Attachment Questionnaire using a 4-point Likert scale with 19 items. *Parental self-efficacy* is a continuous characteristic obtained by the parental expectations survey (Reece and Harkless, 1998) using a 4-point Likert scale with 25 items. *Self-efficacy* characteristic obtained by the German adaptation of the Generalized Self-Efficacy Scale (Dumont et al., 2000) using a 4-point Likert scale with 10 items.

Appendix Table C: Number of Home Visitors, by Delivery Model

Tandem Model			**Continuous Model**		
Number of Home visitors in the family	Freq.	Percent	Number of Home visitors in the family	Freq.	Percent
1	4	2.65	1	96	57.14
2	13	8.61	2	56	33.33
3	92	60.93	3	8	4.76
4	21	13.91	4	6	3.57
5	14	9.27	5	2	1.19
6	5	3.31			
7	2	1.32			
N	151		N	168	

Notes: The table shows how many families were visited by how many visitors, separately for each deliver model. The data includes all families allocated to the Continuous Model or to the Tandem Model at baseline, excluding families from the site where both delivery models were implemented. Source: Standardized encounter forms filled in by the home visitors for each home visit.

Appendix Figure A: Difference in Visit Content by Model Type



Notes: The figure shows the percentage of home visits on average devoted to each of the six topics covered by the program curriculum, separately for each delivery model. The dataset includes every home visit for all 319 families (168 in the CM and 151 in the TM) allocated to the Continuous Model or to the Tandem Model at baseline, excluding families from the site where both delivery models were implemented. Source: Standardized encounter forms filled in by the home visitors for each home visit.