Effects of CHILDREN's programs

Laura Huber 11709507

Laura Jepsen 11567799

Jonathan Kirschner 11724530

Rafael Schütz 11574909

Yannick Zurl 11701793

Studentisches Praxisprojekt zur Empirischen Wirtschaftsforschung PaRE3To

Ludwig-Maximilians-Universität Munich

27th February 2020

Contents

List of Tables

List of Figures

Abstract

abstract on separate page 100 - 150 words

1 Formal requirements

all important information on title page, maybe use template
always name sources of tables and graphs
have to cite data sets
latin number pages

2 Examples

Citing

If you want to cite **Millstein.2020** inline, use this.
If you want to cite (**Millstein.2020**) inline with a page number, use this.
If you want to cite **Millstein.2020** only as authors, use this.

3 Introduction

the introduction should include: motivation, precise research questions, very short literature review, most important results, further proceedings

CHILDREN's aims for data analysis
CHILDREN supports organizations working with children and youth across Germany (in German: Einrichtungen der offenen Kinder- und Jugendarbeit) across Germany. We call them organizations in the following. They apply to CHILDREN for yearly grants. If approved, they are supposed to use them for specific purposes defined by CHILDREN. CHILDREN provided us with data from two of its flagship programs: Mittagstisch (we refer to this as Meals program) and Entdeckerfonds (Trips program). The organizations use money from the Meals program to finance meals, from breakfast to dinner, that they sell at concessionary prices to the children and youth that visit them. In the following, we call these children and youth who ultimately profit from CHILDREN's grants beneficiaries. The organizations also use money from the Trips program to make trips to nearby places usually unknown to the beneficiaries. Unless otherwise specified, we consider all variables to be metric, even if they are ordinal.

3.1 Literature overview

Previous research shows, that the childhood and youth is central in development of character traits and future opportunities. Factors like education, stable family conditions, social contacts and mobility shape children and adolescents for their further life. Heckmann and Carneiro (2003) show that especially early family factors determine the most of gaps in highschool attendance. Moreover, Deckers et. al. (2015) found by experiments in Germany, that children of higher educated parents are significantly more patient. Children from parents with lower income and lower educational attainment have typically a lower weight at birth and are born in a later week of gestation, which are indicators for bad initial conditions (Deckers et. al. 2015, Case et. al. 2002). Children which grow up in families with a low socio-economic status are often less mobile and more likely to stay amongst themselves. This could have a negative influence on their character development. For example, Chetty et. al (2014) investigate for the US, that areas with more mobility are highly correlated with better primary schools, greater social capital and more stable family conditions. Furthermore, the result of Deckers et. al. (2020) provides causal evidence on the effect of social environment on prosocial attitudes. Being supported by a mentor could result in a significant and persistent increase in the prosociality of elementary school children, with regard to prosocial role models and intense social interactions. Heckmann et. al. (2010) investigated a preschool education program in the US. He shows that an investment in the improvement of the childhood conditions could have a high rate of return, even by controlling for possible distortions.

4 Data

To measure the effect of CHILDREN´s engagement on the funded organizations, we use the data CHILDREN collected from 2011 to 2018. In each year they send a survey to the organizations with several questions about the previous year. The number of organizations varies among the years and increases over time, from 52 in the 2012 survey to 73 in the survey from 2020. In some organizations one employee fills in the survey and in others they do it as a team. Since the children and adolescents are not questioned directly, all responses are documented through the perception of the employees. The number of variables varies over time as well. Included are numbers such as the average eaters per meal or the amount of money they provide to the organizations but also general questions. For instance, CHILDREN asks the average amount of kids with a better confidence or an improved dietary knowledge in the specific organization. This part of the survey must be answered on a scale from zero (no kids) to four (all kids). If an organization does not answer a question, this is documented as a "99". We worked with the statistical program "R" and therefore changed the format from 99s to NA´s (not available) to avoid distortions. The surveyed variables change over the years, but some of them are included every year. However, we did several steps to get a full dataset we could work with (we could use for our empirical analysis??). The data was divided into one dataset for each survey from 2011 to 2019, but we only use the surveys till 2018 since in 2019 some organization-ID´s occurred several times and the data for 2019 were incomplete. Since each survey includes data about the year before, we changed the names of the dataset to the corresponding year and finally used the years from 2011 to 2018. Moreover, we outlined a hierarchical file structure, enabling us to use relative file paths throughout. This makes a quick work with R possible since we only use paths relative to the working directory. Afterwards we made sure that variables with names containing non-standard characters like German “Umlaute” are correctly read in and established naming conventions. We created a file reading the excel sheets and we reviewed and aligned new English-language variable names across the years. Moreover, we systematically compared variable names between years by creating a correspondence table, ordered first by variables of 2019, then of 2018 and so on. To ensure the comparability between the years, we gave all variables from the different years that equal each other the same name. As a next step, we merged the different datasets to one dataset, including all years and variables CHILDREN collected. For an efficient and clear data structure, we created a function that automatically changed the data type of all variables from "character" to "ordinal" and added several versions for each initially metric encoded variable afterwards. The three variants are ordinal, standardized and weighted (FUßNOTE: The variables regarding the Mittagstisch are weighted as variable*0.25*eaterspermeal, the variables that are assigned to the Entdeckerfonds as variable*0.25*tripskidsno). Furthermore, we created more new variables: We used the information CHILDREN gave us in another excel-sheet to assign the German states to the corresponding organization-ID and created dummy-variables for each ID, every year and a treatment dummy that will be explained in a later section. The final dataset we worked with is structured as follows: Each row represents one organization-ID with the answers the organization gave in the specific year. The questions are divided in two categories: the variables regarding to the meals program, answered by all organizations since they are all part of this program and the trips program variables, answered by the organizations that take part on the trips program in the respective year. Including the years from 2011 to 2018 and all variables we created, the final dataset has X observations of Y variables.

Many organizations do not answer all questions CHILDREN poses. We create a seperate data set, in which we impute missing values with an organization-specific linear trend. CHILDREN supports some very large organizations that give out hundreds of meals a day or conduct dozens of trips per year. We fit our models with outliers excluded to see if they are driving results. We define an outlier as a value that is at least 1.5 interquartile ranges below the 25th percentile or equally far above the 75th percentile. Once we exclude outliers in terms of numbers of meals and once in terms of number of trips.

5 Summary Statistics

5.1 Fundamental Dynamics

In this section, we give an overview of the dynamics of CHILDREN's two flagship programs. We focus on the number of estimated ultimate beneficiaries in both programs, median total subsidy, median subsidy per institution, and median subsidy per beneficiary. We also look at selected outcomes, i.e. those related to health as well as self-worth and day-to-day skills. We have converted all nominal monetary variables into 2015 euros, using price indices from the Federal Statistical Office of Germany (Statistisches Bundesamt). We deflate (requested) grants as well as organizations' total expenses for the Meals program with the price index related to food and non-alcoholic beverages (in German: Nahrungsmittel und alkoholfreie Getränke) and (requested) grants towards the Trips program with the price index for leisure, entertainment, and culture (in German: Freizeit, Unterhaltung und Kultur). These are only available after logging in on DESTATIS. The organizations also gave information about their total yearly budget. We inflate this with the general price index.

	Year	Beneficiaries, Meals	Beneficiaries, Trips	Organizations, Meals	Organizations, Trips
1	2011	3748.0		52	
2	2012	3556.0	2803.0	51	44
3	2013	4015.0	2823.0	55	42
4	2014	4685.0	2752.0	55	43
5	2015	5857.0	3823.0	55	49
6	2016	3075.0	3819.0	59	48
7	2017	4895.0	4150.0	64	48
8	2018	5102.5	6911.0	68	49

Table 1: Summary Statistics

Table **??** shows that, at the beginning of the time series in 2011, in the Lunch program they supported in 52 organizations. In 2018, this number has increased to 68. In 2011, CHILDREN financed meals for 3748 beneficiaries, and for 5103 in 2018. In the trips program, which launched in 2012, the number of supported organizations amounted to 44 in 2012 and grew to 49 until 2018. In 2012, their grant allowed 2803 beneficiaries to go on trips, and 6911 in 2018.

5.2 Trend of grants

For the dynamics of figure **??** we visualze the dynamics of CHILDRENs' grants by distinguishing between the sum of grants in one year, the median and the median grant per beneficiary. We also compare the grants of the Lunch and the Trips Program.

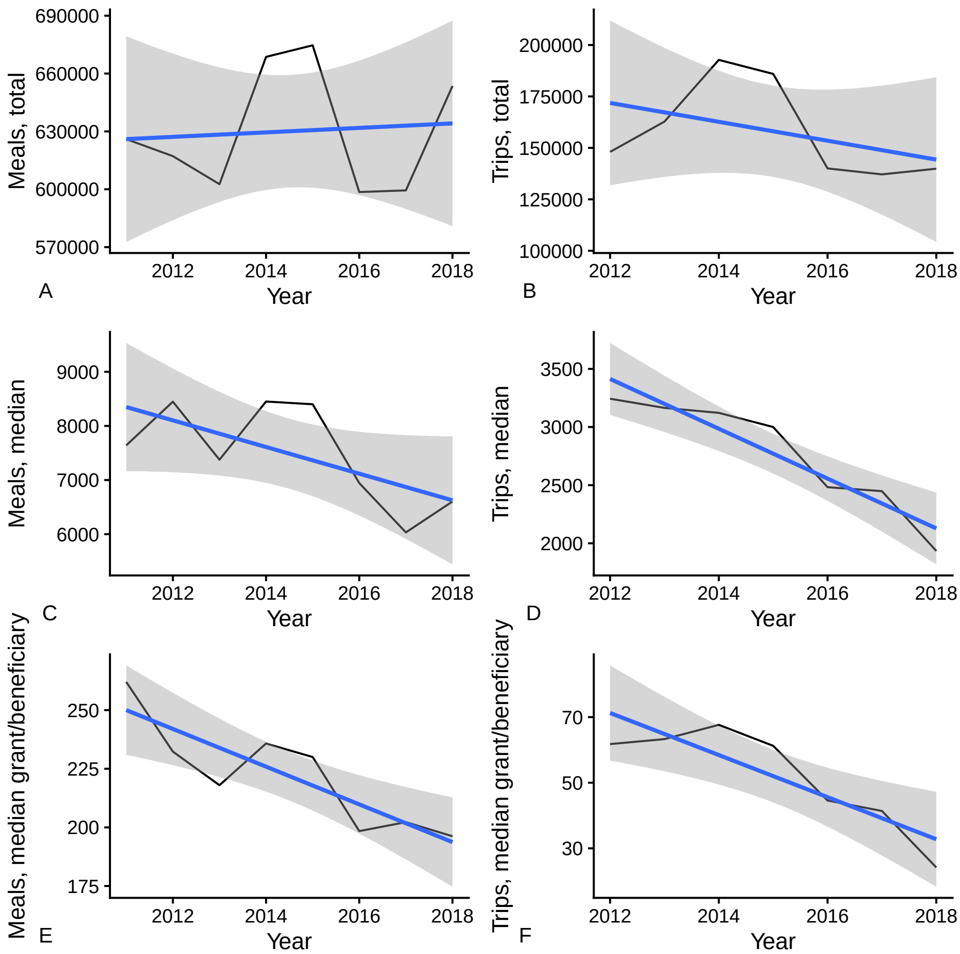


Figure 1: Yearly dynamics of total grants in Meals and Trips program

This graph shows the development of grants in the Meals compared to the Trips program. We distinguish between the sum of grants in one year, the median grant and the median grant per beneficiary. From left to right: Meals, Trips. From top to bottom: sum, median, median per beneficiary. We have deflated the values to 2015 euros using the price index related to food and non-alcoholic beverages(in German: Nahrungsmittel und alkoholfreie Getränke) for the Meals program and the price index related to Leisure, Entertainment and Culture (in German: Freizeit, Unterhaltung, Kultur) provided by the Federal Statistical Office of Germany (Statistisches Bundesamt).

Figure ?? shows no exact trend for the total real grant for the Lunch program. Between 2013 and 2015 the grant increased from about 600000 EUR to 680000 EUR, but falls behind in the following year. Since 2017 an increase is again visible. In comparison to this, there is a clearer negative trend in median Meals subsidy, falling from about 8000 EUR in 2012 to about 6500 EUR in 2018. In accordance, the median Meals grant per beneficiary shrinks from about 250 EUR to about 200 EUR.

In the total Trips grant, a slightly negative trend is visible, after the subsidy increased to approximately 200000 EUR in 2014, but decreased to about 130000 EUR in 2016, and is connstand eversince. The Trips median grant therefore also decreased from 3000 EUR to 2000 EUR over the time period, as well as the median grant per beneficiary decreased from about 60 EUR to about 30 EUR.

These results visualize, for the Lunch as well as for the Trips program, together with the fundamental dynamics of table ??, that children was able to increase the number of organizations they support, but had to distribute the available subsidy among them.

5.3 Health relevant variables over time

In its yearly surveys, CHILDREN asks about three variables closley related to a healthy diet. These are the shares of beneficiaries who are healthier, have a growing appreciation for a healthy diet, or have increased their knowledge about what constitutes a healthy diet. Figure ?? displays the share of organizarions in each category of the health outcome from 2011 to 2018. The possible values are: all, most, some, few, and none. For this figure, we use the original ordinal varibles which result from the survey structure.

In Figure ??, for the variable 'lessIll_ordered', there is much non available data (lessIll_ordered refers to the share of beneficiaries who are less frequently ill). In all years, most organizations say that most beneficiaries are more healthy. The least stated category is, that all beneficiaries are less ill. That few are less ill, appears sometimes and the leftover category, none, which is only coded for the variable 'lessIll_ordered', appears as good as never.

In the second plot of figure ?? (variable 'dietaryKnowledge_ordered', refers to the share of beneficiaries with expanded dietary knowledge) most organizations state, that predominantely most or all beneficiaries increased their knowledge about what constitutes a healthy diet. The least stated category is that some beneficiaries increased their dietary knowledge. The leftover category, few, does not appear.

The the bottom plot of figure ??, which visualizes the variable 'appreciateHealthy_ordered' (refers to the share of beneficiaries with increased appreciation for a healthy diet). Again most organizations state, that most beneficiaries have a growing appreciation for a healthier diet. The second most stated answer is that all beneficiaries have a growing appreciation for a healthier diet. The least stated category by the organizations is, as well as in the second plot, that some have a growing appreciation for a healthier diet. The leftover category, few, also does not appear.

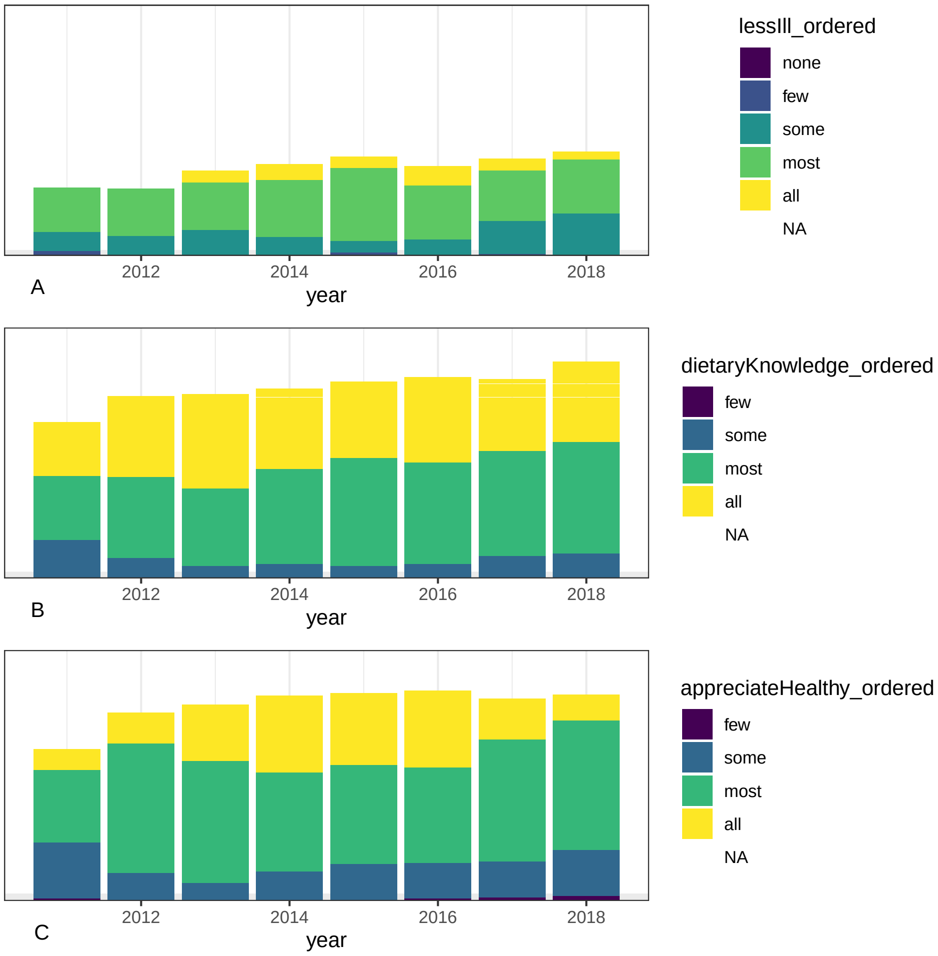
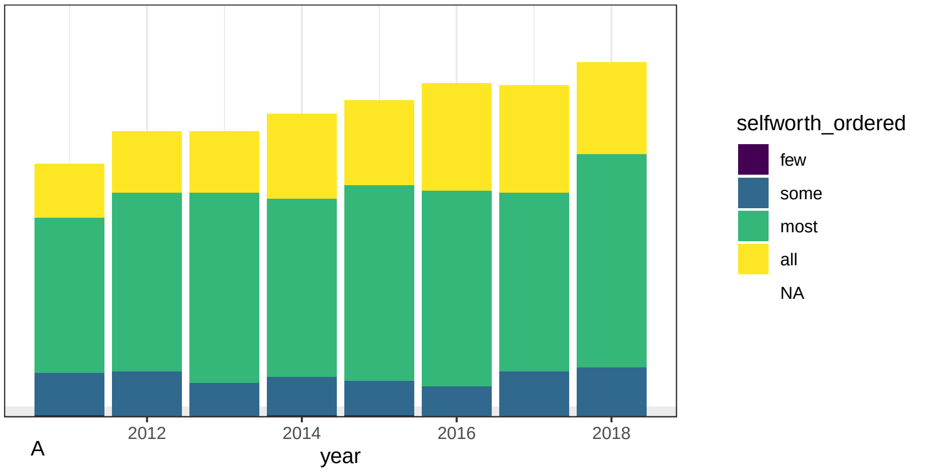


Figure 2: Health outcome over time

In its yearly surveys CHILDREN asks about three variables closley related to a healthy diet. These are the shares of beneficiaries who are less frequently ill (lessIll_ordered), with expanded dietary knowledge (dietaryKnowledge_ordered), or with increased appreciation for a healthy diet (appreciateHealthy_ordered). The x-axis plots the year. The y-axis displays the share of organizarions in each category of the health outcome. The possible values are: all, most, some, few and none.



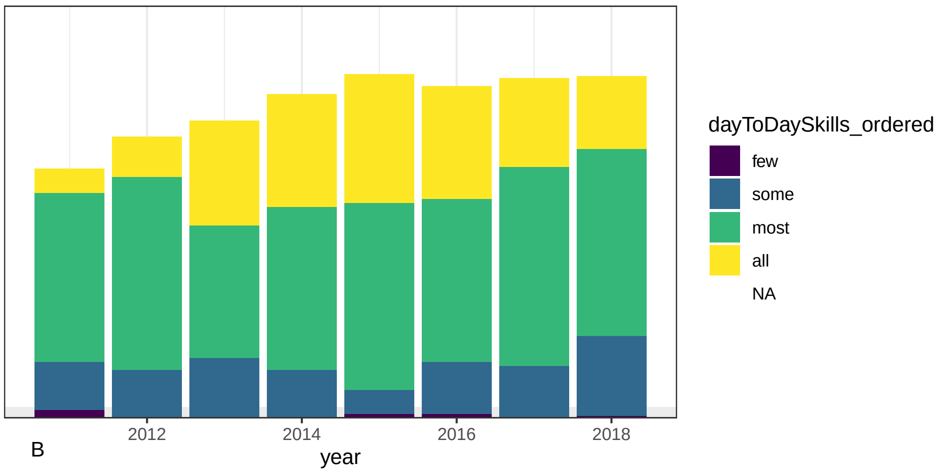


Figure 3: Selfworth and everyday expertise over time

In its yearly surveys CHILDREN has always asked about following two variables. These are the shares of beneficiaries with improved self-worth (selfworth_ordered) and with a broadened understanding for everyday expertise (dayToDaySkills_ordered). The x-axis plots the year. The y-axis displays the share of organizations in each category of the health outcome. The possible values are: all, most, some, few, and none.

5.4 Selfworth and everyday expertise

In its yearly surveys, CHILDREN has always asked about the following two variables. These are the shares of beneficiaries with improved self-worth and with broadened everyday expertise. Figure 4 displays the share of organizarions in each category of the health outcome from 2011 to 2018. The possible values are: all, most, some, few, and none. As well as in the last figure (figure 3), we use the original ordinal varibles which result from the survey structure.

Figure **??** shows in its upper plot the development of the answered categories for the variable selfworth_ordered (refers to the share of beneficiaries with improved self-worth). Mostly, the answer is that most beneficiaries in most organizations have more selfworth. The second most stated category is that all beneficiaries gained in selfworth and the least stated is that some have more selfworth. The leftover category, few, does not appear.

In the bottom plot of figure **??** the development of the share of beneficiaries with a broadened understanding for everyday expertise over time is visible. The results are about similar to the previous results of selfworth: Most beneficiaries in most organisations have a growing understanding for everyday expertise.

6 Feature Selection

6.1 Factor Analysis

Benefits of reducing dimensionality can be reduced computational demands, reduced multiple-testing burden, reduced noise, as well as better-behaved data **Millenstein.2020**; **p.677**.

We use imputed data set.

we use the orthogonal rotation technique Varimax, which is recommended in Price (2017, p. 307) when the goal of analysis is to "minimize complexity of factors by maximizing variance of loadings on each factor" We estimate factor scores with the Bartlett method, which is similar to a Maximum Likekelihood estimation, and presented as the most convential one in Eid and Schmidt (2010, p. 291) We use maximum likelihood as factoring method

6.2 Double Selection

7 Regressions

7.1 Empirical Approach

$$y_{it} = \beta_0 + \beta_1 x_{it} + \epsilon_{it} \tag{1}$$

In this section we use a simple linear model, as described in equation **??**. We look at the association between:

- the subsidy (in 2015 EUR) an organization receives through CHILDREN's Meals program and the number of meals it dispenses
- the subsidy (in 2015 EUR) an organization receives through CHILDREN's Trips program and the number of trips it conducts
- the subsidy per beneficiary (in 2015 EUR) an organization receives through CHILDREN's Meals program and the standardized share of beneficiaries with improved self-worth

	(1)	(2)	(3)	(4)	(5)
(Intercept)	-12089.14^{*}	-1814.16	3535.39^{***}	3107.70^{***}	-12250.60^{**}
	(5192.86)	(1765.93)	(498.99)	(508.94)	(4524.09)
realSubsidy	2.61^{***}	0.50^{**}	0.29^{***}	0.25^{***}	2.72^{***}
	(0.57)	(0.18)	(0.05)	(0.05)	(0.51)
eatersPerMealNo		172.83^{***}		19.00^{*}	
		(14.92)		(8.45)	
R^2	0.43	0.73	0.13	0.21	0.45
Adj. R^2	0.43	0.73	0.12	0.20	0.45
Num. obs.	329	329	250	250	440
RMSE	39992.79	27390.90	3629.72	3463.66	39601.41

Dependent variable: number of meals
realSubsidy: subsidy for Meals program in 2015 EUR
eatersPerMeal: number of beneficiaries of Lunch program
Model (1): original data set, simple linear model, estimated with OLS
Model (2): original data set, linear model with controls, estimated with OLS
Model (3): data set without outliers, simple linear model, esmitaed with OLS
Model (4): data set without outliers, linear model with controls, estimated with OLS
Model (5): imputed data set, simple linear model, estimated with OLS
All regressions are estimated with robust standard errors $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

Table 2: Association between number of meals and real subsidy

- the subsidy per beneficiary (in 2015 EUR) an organization receives through CHILDREN's Meals program and the standardized share of beneficiaries with broadened everyday expertise
- a standardized measure of the healthiness of the meals an organization dispenses and various standardized health-related outcomes of beneficiaries

We discuss all these models in turn.

7.2 Direct effects of CHILDREN's grants

If CHILDREN's grants are to have an effect on beneficiaries, they should first influence the output of organizations in terms of meals dispensed and trips conducted.

As table **??** shows, this is emphatically true. Whether we look at the original data set, the one without outliers and the one with imputed values, a strong association becomes evident. The estimated coefficients are very similar when we use the original dataset and the one with imputed values. In the case of the original data set, increasing the subsidy to an organization by one EUR is associated with 2.6 additional meals dispensed. This estimate is highly statistically significant. When we exclude outliers, i.e. those organizations that give out very many menus or very few menus, the estimated coefficient decreases by about one order of magnitude. Still, spending not much more than three EUR more goes hand in hand with one extra meal. The estimate is also highly statistically significant.

	(1)	(2)	(3)	(4)	(5)
(Intercept)	3.7049***	3.4394***	2.6236***	2.3660***	3.6237***
	(0.3313)	(0.3359)	(0.2300)	(0.2609)	(0.3253)
realTripsSubsidy	0.0002*	0.0001	0.0003***	0.0003***	0.0002*
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
tripsKidsNo		0.0059		0.0043	
		(0.0032)		(0.0027)	
R^2	0.0474	0.0729	0.0880	0.1241	0.0504
Adj. R^2	0.0444	0.0671	0.0844	0.1172	0.0476
Num. obs.	322	319	257	256	334
RMSE	2.9565	2.8967	1.6981	1.6579	2.9310

Dependent variable: number of trips
realTripsSubsidy: subsidy for Trips program in 2015 EUR
tripsKidsNo: number of beneficiaries of Trips program
Model (1): original data set, simple linear model, estimated with OLS
Model (2): original data set, linear model with controls, estimated with OLS
Model (3): data set without outliers, simple linear model, esmitaed with OLS
Model (4): data set without outliers, linear model with controls, estimated with OLS
Model (5): imputed data set, simple linear model, estimated with OLS
All regressions are estimated with robust standard errors $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

Table 3: Association between number of trips and real subsidy

In the remit of the Trips program, the picture looks much different. In the case of the original data set and the data set with imputed variables, an increase in the trips subsidy by 5000 EUR is needed for an extra trip. The coefficent of 0.0002 is statistically significant at the 5% level. As soon as outliers are excluded, the coefficent changes to 0.0003, which is highly statistically significant. This means that, according to this model, at least 3000 EUR are needed for an additional trip. In sum, there seems to be no clear connection between how much money CHILDREN gives to an organization and the number of trips it organizes.

7.3 Variables of interest: selfworth and everyday expertise

In their surveys, CHILDREN asks about two variables in both programs. These are the share of beneficiaries with improved self-worth and share of beneficiaries with broadened everyday expertise. This feature could potentially allow us to compare the two programs regarding their realtive effectiveness. Like always, we standardize the two outcome variables. As before, we can only make associational claims. We invariably find no clear-cut relationship between the subsidy per beneficiary and the two outcomes. In fact, no coefficent is statistically significantly differnent from zero at any of the usual levels. The four variables are recorded for each year in which the program was operative. Tables **??** and **??** present the null results.

graphicx

graphicx

	(1)	(2)	(3)	(4)	(5)
(Intercept)	0.08	0.12	0.09	0.12	0.23^{*}
	(0.09)	(0.12)	(0.09)	(0.11)	(0.11)
realSubsidyPerBeneficiary	-0.00		-0.00		-0.00
	(0.00)		(0.00)		(0.00)
realTripsSubsidyPerBeneficiary		-0.00		-0.00	
		(0.00)		(0.00)	
ML1					0.24^{***}
					(0.06)
ML2					0.37^{***}
					(0.05)
ML3					0.15^{***}
					(0.04)
R^2	0.00	0.01	0.00	0.01	0.30
Adj. R^2	0.00	0.01	0.00	0.01	0.28
Num. obs.	428	184	430	187	161
RMSE	1.00	1.00	1.00	1.00	0.79

realSubsidyPerBeneficiary: subsidy per beneficiary of Meals program in 2015 EUR
realTripsSubsidyPerBeneficiary: subsidy per beneficiary of Trips program in 2015 EUR
Model (1): dependent variable: share of beneficiaries with improved self-worth in the Lunch program, original data set, simple linear model, estimated with OLS
Model (2): dependent variable: share of beneficiaries with improved self-worth in the Trips program, original data set, simple linear model, estimated with OLS
Model (3): dependent variable: share of beneficiaries with improved self-worth in the Lunch program, imputed data set, simple linear model, estimated with OLS
Model (4): dependent variable: share of beneficiaries with improved self-worth in the Trips program, imputed data set, simple linear model, estimated with OLS
Model (5): dependent variable: share of beneficiaries with improved self-worth in the Lunch program, imputed data set, linear model with extracted factor scores as controls, estimated with OLS
All regressions are estimated with robust standard errors $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

Table 4: Association between selfworth and subsidy per beneficiary

	(1)	(2)	(3)	(4)	(5)	(6)
(Intercept)	0.15	0.13	0.14	0.11	0.28^{*}	0.08
	(0.09)	(0.10)	(0.09)	(0.10)	(0.11)	(0.09)
realSubsidyPerBeneficiary	-0.00		-0.00		-0.00	
	(0.00)		(0.00)		(0.00)	
realTripsSubsidyPerBeneficiary		-0.00		-0.00		-0.00
		(0.00)		(0.00)		(0.00)
ML1					0.31^{***}	0.03
					(0.06)	(0.07)
ML2					0.40^{***}	0.16^{*}
					(0.06)	(0.07)
ML3					0.16^{**}	0.19^{**}
					(0.05)	(0.06)
ML4						0.49^{***}
						(0.06)
R^2	0.01	0.01	0.01	0.01	0.37	0.37
Adj. R^2	0.01	0.01	0.01	0.01	0.36	0.35
Num. obs.	426	177	429	181	161	169
RMSE	1.00	0.98	1.00	0.99	0.78	0.80

realSubsidyPerBeneficiary: subsidy per beneficiary of Meals program in 2015 EUR
realTripsSubsidyPerBeneficiary: subsidy per beneficiary of Trips program in 2015 EUR
Model (1): dependent variable: share of beneficiaries with with broadened everyday expertise in the Lunch program, original data set, simple linear model, estimated with OLS
Model (2): dependent variable: share of beneficiaries with with broadened everyday expertise in the Trips program, original data set, simple linear model, estimated with OLS
Model (3): dependent variable: share of beneficiaries with with broadened everyday expertise in the Lunch program, imputed data set, simple linear model, estimated with OLS
Model (4): dependent variable: share of beneficiaries with with broadened everyday expertise in the Trips program, imputed data set, simple linear model, estimated with OLS
Model (5): dependent variable: share of beneficiaries with with broadened everyday expertise in the Lunch program, imputed data set, linear model with extracted factor scores as controls, estimated with OLS
Model (6): dependent variable: share of beneficiaries with with broadened everyday expertise in the Trips program, imputed data set, linear model with extracted factor scores as controls, estimated with OLS
All regressions are estimated with robust standard errors $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

Table 5: Association between everyday expertise and subsidy per beneficiary

7.4 Variables of interest: Health variables

Now, we turn to a predictor that CHILDREN does not directly influence until now, but that it easily could. In 2014, 2016, 2017 and 2018 the organizations had to send CHILDREN a sample of their menus. An ecotrophologist collaborating with CHILDREN assessed those menus with regards to how healthy they were. This assessment was based on criteria of the German Nutrition Society (Deutsche Gesellschaft für Ernährung). https://www.schuleplusessen.de/startseite/

In its yearly surveys, CHILDREN asks about three variables closely related to a healthy diet. These are the shares of beneficiaries who are less frequently ill, with expanded dietary knowledge and with increased appreciation for a healthy diet.

Figure **??** gives an overview of the relationship between the healthy food criterion (DGECriteriaNo, index of healthy diet criteria fulfilled in organization's menu) and each of the three health-related variables. For these results, we use their ordinal values. The x-axis displays the index for a healthy diet. The y-axis displays the share of organizarions in each category of the health outcome. The possible values are: all, most, some, few and none. For example, if an organization says that most beneficiaries are healthier, then this would be coded as most. In the last reporting period, there are 13 healthy diet criteria. In previous years there were up to 15. The ecotrophologist assessed each criterion individually and decided wether it was fulfilled or not. The healthy food incex is calculated as the number of fulfilled criteria. Examples for criteria are daily carbohydrates or daily vegetables.

All plots of Figure **??** show that most organizations suffice about 5 to 10 healthy food criteria. The largest part of criterion values are associated with most beneficiaries being less Ill, having a growing appreciation for a healthy diet or having increased their knowledge about what constitutes a healthy diet. It secondly becomes visible that the healthy criterion only exchanges with a growing appreciation for a healthy diet or an increasing dietary knowlege for all beneficiaries (which is the second most stated criterion in the two bottom plots), but not with all beneficaries being more healthy. However, for the variable less ill, the least data is availabe. Overall, the criteria, even a few, seem to have a positive relationship with beneficiaries health relevant variables. This connection is empiricaly examined in regressions **??**, **??** and **??**.

We use standardized versions of the health related variables as outcomes in simple linear models with a standardized version of the healthy food criterion as predictor in each. By estimating these models with ordinary least squares (OLS), we ascribe the same weight to an organization where a thousand beneficiaries regularly eat as to one where only ten people do so. To control for this difference in size, we additionally fit the models with weighted least squares (WLS), using the number of beneficiaries as weights.

In Table **??**, after weighting the regression for the original data set, the coefficent rises slightly but is becoming less statistically significant. An increase of the healthy food

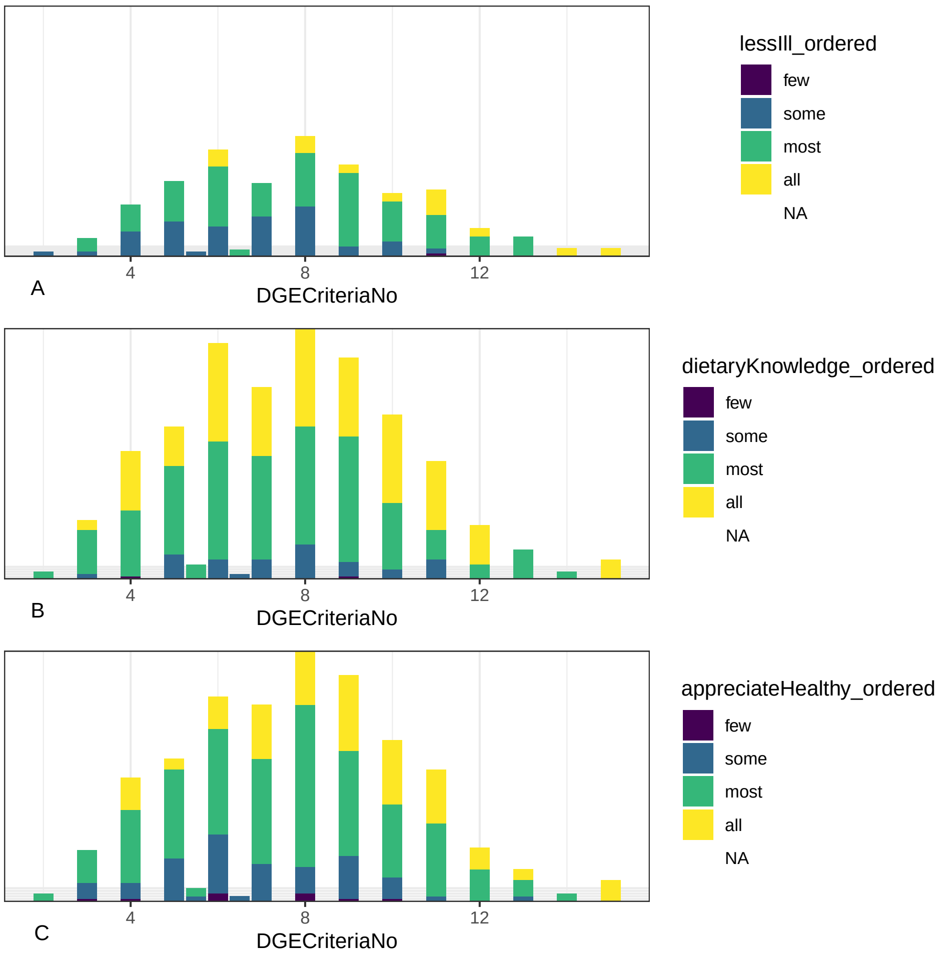


Figure 4: Health Outcomes versus Healthy Meals

DGECriteriaNo is an index of healthy diet criteria fulfilled in organization's menu. It is based on criteria of the German Nutrition Society (Deutsche Gesellschaft für Ernährung). According to information from CHILDREN, they ask the organizations to send them a sample of their menus. An ecotrophologist collaborating with CHILDREN assessed the menus. In its yearly surveys CHILDREN asks about three variables closley related to a healthy diet. These are the shares of beneficiaries who are less frequently ill (lessIll_ordered), with expanded dietary knowledge (dietaryKnowledge_ordered), or with increased appreciation for a healthy diet (appreciateHealthy_ordered). The x-axis plots the index for a healthy diet. The y-axis displays the share of organizarions in each category of the health outcome. The possible values are: all, most, some, few, and none

	(1)	(2)	(3)	(4)	(5)
(Intercept)	0.02	0.46**	0.09	0.39***	0.05
	(0.08)	(0.16)	(0.07)	(0.12)	(0.07)
DGECriteriaNoScaled	0.33***	0.35*	0.25***	0.24	0.18*
	(0.08)	(0.16)	(0.07)	(0.14)	(0.07)
ML1					0.12*
					(0.06)
ML2					0.27***
					(0.06)
R^2	0.12	0.29	0.07	0.16	0.19
Adj. R^2	0.11	0.29	0.07	0.16	0.17
Num. obs.	121	120	177	177	161
RMSE	0.91	7.83	0.94	7.95	0.87

Dependent variable: share of beneficiaries who are less frequently ill
DGECriteriaNo: index of healthy diet criteria fulfilled in organization's menu
Model (1): original data set, simple linear model, estimated with OLS
Model (2): original data set, simple linear model, estimated with WLS
Model (3): imputed data set, simple linear model, estimated with OLS
Model (4): imputed data set, simple linear model, estimated with WLS
Model (5): imputed data set, linear model with extracted factor scores as controls, estimated with OLS
All regressions are estimated with robust standard errors $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

Table 6: Association between healthy meals criterion and beneficiaries being less ill

criterion by one standard deviation is associated with an increase in beneficiaris being less frequently ill by 0.35 standard deviations. Regressing with the imputed dataset, results decrease by about 0.1. Using weighted least squares additionaly results in a slighty decreased coefficent, which is not significant anymore.

An expanded model with the dependent variable less Ill is presented in Appendix 1 in table **??**.

In Table **??** weighted least squares constanly leads to smaller coefficents which are not statistically significant. For the original data set, the coeffiecent is slightly negative. Using the imputed data set, an increase in the healthy meals criterion by one standard deviation is associated with an increase in beneficiaries dietary knowledge by 0.1 standard deviations. This coefficent is as well insignifacant.

In the regressions presented in table **??**, the changes between ordinary and weíghted least squares are much stronger. For the original data set, the OLS coefficent amounts to 0.27, which is highly significant. When using WLS, we find a slightly negative association. A one standard deviation increase in the healthy meals criterion is associated with a decrease in the appreciation of a healthy diet by 0.02. This coefficent is not significant. For the imputed dataset we find a similar picture.

The regressions with WLS suggest that the results obtained using OLS, are driven by large oraganizations.

graphicx

graphicx

	(1)	(2)	(3)	(4)	(5)
(Intercept)	0.02	0.08	0.02	0.21	0.02
	(0.07)	(0.19)	(0.06)	(0.18)	(0.07)
DGECriteriaNoScaled	0.11	−0.02	0.12^{*}	0.10	−0.00
	(0.06)	(0.12)	(0.05)	(0.14)	(0.06)
ML1					0.26^{***}
					(0.06)
ML2					0.24^{***}
					(0.06)
ML3					0.37^{***}
					(0.06)
R^2	0.01	0.00	0.02	0.01	0.31
Adj. R^2	0.01	−0.00	0.01	0.01	0.29
Num. obs.	214	212	275	275	161
RMSE	0.98	8.49	0.96	9.45	0.83

Dependent variable: share of beneficiaries with expanded dietary knowledge
DGECriteriaNo: index of healthy diet criteria fulfilled in organization's menu
Model (1): original data set, simple linear model, estimated with OLS
Model (2): original data set, simple linear model, estimated with WLS
Model (3): imputed data set, simple linear model, estimated with OLS
Model (4): imputed data set, simple linear model, estimated with WLS
Model (5): imputed data set, linear model with extracted factor scores as controls, estimated with OLS
All regressions are estimated with robust standard errors $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

Table 7: Association between healthy meals criterion and beneficiaries dietary knowledge

graphicx

The previous simple linear models are regressed with metric versions of the survey varibales. Originally, variables which range from none to all are ordinal. To account for this, we present a method to regress those variables according to their original traits. Therefore, we use a cumulative logit model. Selected results for the Meals program are presented in appendix 2 in the tables**??**, **??**, **??**, **??** and **??**.

¡¡¡¡¡¡ HEAD =======

8 Feature Selection

8.1 Factor Analysis

Benefits of reducing dimensionality can be reduced computational demands, reduced multiple-testing burden, reduced noise, as well as better-behaved data **Millstein.2020**; **p.677**.

We use imputed data set.

we use the orthogonal rotation technique Varimax, which is recommended in Price (2017, p. 307) when the goal of analysis is to "minimize complexity of factors by maximizing variance of loadings on each factor" We estimate factor scores with the Bartlett method, which is similar to a Maximum Likekelihood estimation, and presented as the

	(1)	(2)	(3)	(4)	(5)
(Intercept)	-0.03	0.26	0.02	0.37^{*}	0.05
	(0.07)	(0.18)	(0.06)	(0.17)	(0.07)
DGECriteriaNoScaled	0.27^{***}	-0.02	0.25^{***}	0.01	0.03
	(0.07)	(0.15)	(0.06)	(0.13)	(0.06)
ML1					0.03
					(0.07)
ML2					0.47^{***}
					(0.05)
ML3					0.24^{***}
					(0.05)
R^2	0.06	0.00	0.06	0.00	0.37
Adj. R^2	0.06	-0.00	0.06	-0.00	0.35
Num. obs.	213	211	274	274	161
RMSE	1.02	8.61	1.01	9.00	0.82

Dependent variable: share of beneficiaries with increased appreciation for a healthy diet
DGECriteriaNo: index of healthy diet criteria fulfilled in organization's menu
Model (1): original data set, simple linear model, estimated with OLS
Model (2): original data set, simple linear model, estimated with WLS
Model (3): imputed data set, simple linear model, estimated with OLS
Model (4): imputed data set, simple linear model, estimated with WLS
Model (5): imputed data set, linear model with extracted factor scores as controls, estimated with OLS
All regressions are estimated with robust standard errors $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

Table 8: Association between healthy meals criterion and beneficiaries appreciation of a healthy diet

most convential one in Eid and Schmidt (2010, p. 291) We use maximum likelihood as factoring method

8.2 Double Selection

¿¿¿¿¿¿¿ af55806ca517c74fcdbf9b9542cb65743282f0b7

9 The Effects of the Trips Program on Participating Beneficiaries

How do children and adolescents benefit from visiting social institutions that CHILDREN supports financially? So far this question could not be empirically validated. Hence, one of our biggest challenges was determining a possible solution for measuring potential causal effects of the supported programs on the participating beneficiaries. During the first meeting with CHILDREN, Wiltrud de Haan presented relevant information that CHILDREN supports all organizations with their meals program. However, not all organizations receive additional funding to provide the trips program. This fact could be used for applying an empirical approach which determines the effects of the trips program by comparing a treatment with a control group. The aim of this analysis is to test whether the activities offered by the trips program have a positive effect on the participating children, measured through a change in selfworth and everyday expertise.

9.1 Empirical Approach

The basis of the empirical approach is the specification of the treatment and the control group. Using the data provided by children we determine the treatment group as all organizations that receive funding for both the trips and the meals program. On the other hand, the control group represents all organizations that only receive funding for the meals program. Therefore, the social institutions in the control group do not participate in the trips program.

9.1.1 Specification of the Treatment Group

When analyzing the available dataset, it was not certain which organizations actually received funding for the trips program. As previously mentioned, the CHILDREN survey consists of two parts. The first part contains questions that are specific to the meals program and the second part includes all questions which are relevant for organizations that receive funding for the trips program. The dataset shows that there are several organizations that did not provide information regarding the trips program. In consultation with Wiltrud de Haan, she informed us that these specific organizations did not receive funding to offer trips and activities. Hence, we assumed that all organizations that did not provide information regarding the trips program in a given year, did not receive funding for the program and would be part of the control group. When analyzing the possible treatment group, we realized that some organizations did not completely answer the trips program survey. There are organizations, for example organization 103 in 2015, that did not provide information about the funding amount but answered the Entdeckerfonds survey questions such as the number of trips in a given year. Wiltrud de Haan informed us, that one possible explanation for this occurrence is that these organizations did not completely use the funding in the previous year. Therefore, they did not receive additional funding in the following year but were able to organize trips for the children as part of the trips program with the remaining funding of the year before.

Due to these uncertainties we determine the treatment group in two different ways. Firstly, the treatment variable is specified as follows:

$$TreatEF_{it} = \begin{cases} 1, & \text{if organization } i \text{ participates in the trips program in year } t \text{ or any year before} \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

In this setting, once an organization receives funding for the trips program, we consider this organization as treated in the year of the first funding and all proceeding years. This implies that an organization remains in the control group as long as it does not receive funding for the trips program. Due to this definition, an organization cannot change from the treatment into the control group.

In general, the constellation of the treatment and control group varies across years. As the number of organizations supported by CHILDREN increased over time, both the treatment and control group increased as well. At first, new organizations are generally funded to provide the meals program and might receive funding for the trips program during later funding periods. In this case, organizations change from the control group into the treatment group resulting in an overall adjustment of the constellation of both groups. However, the dataset also includes organizations that switch from the treatment to the control group. The first definition of the treatment variable does not take this case into account. Therefore, we additionally construct a second treatment indicator which is defined as follows:

$$TreatEF_{it} = \begin{cases} 1, & \text{if organization } i \text{ participates in the trips program in year } t \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

Using this alternative treatment variable, the actual funding for the trips program in a given year determines the treatment. Therefore, the treatment status could change every year. With both definitions of the treatment variable, the number of observations units in the control group is substantially lower than in the treatment group. Consequently, the estimates cannot be very robust. Therefore, the results should not be overstated.

9.1.2 Dependent Variables

The fact that all organizations within the control group did not answer the trips program survey questions was helpful to categorize the control and the treatment group. However, the impact of the trips program cannot be measured with variables that are specific to the trips program as these were only observed for the treatment group. Therefore, we have to use variables of the meals program as outcome variables in our analysis. For this purpose, we identified selfworth and everyday expertise as appropriate dependent variables. Firstly, both variables are applicable to the trips program as well as the meals program. In addition, the chosen variables are observed in every year throughout the whole observation period and seem to be influenced by the fact that an organization receives funding to provide trips and other activities.

9.1.3 Graphical Evidence

To check for differences in treatment and control group, we created descriptive statistics regarding to the variables of interest. Figure 5 and Figure 6 represent the development of average selfworth and average everyday expertise over time for both the treatment and control group. The two graphs illustrate a difference in levels as well as in trends between the treatment and control group in either average selfworth or average everyday

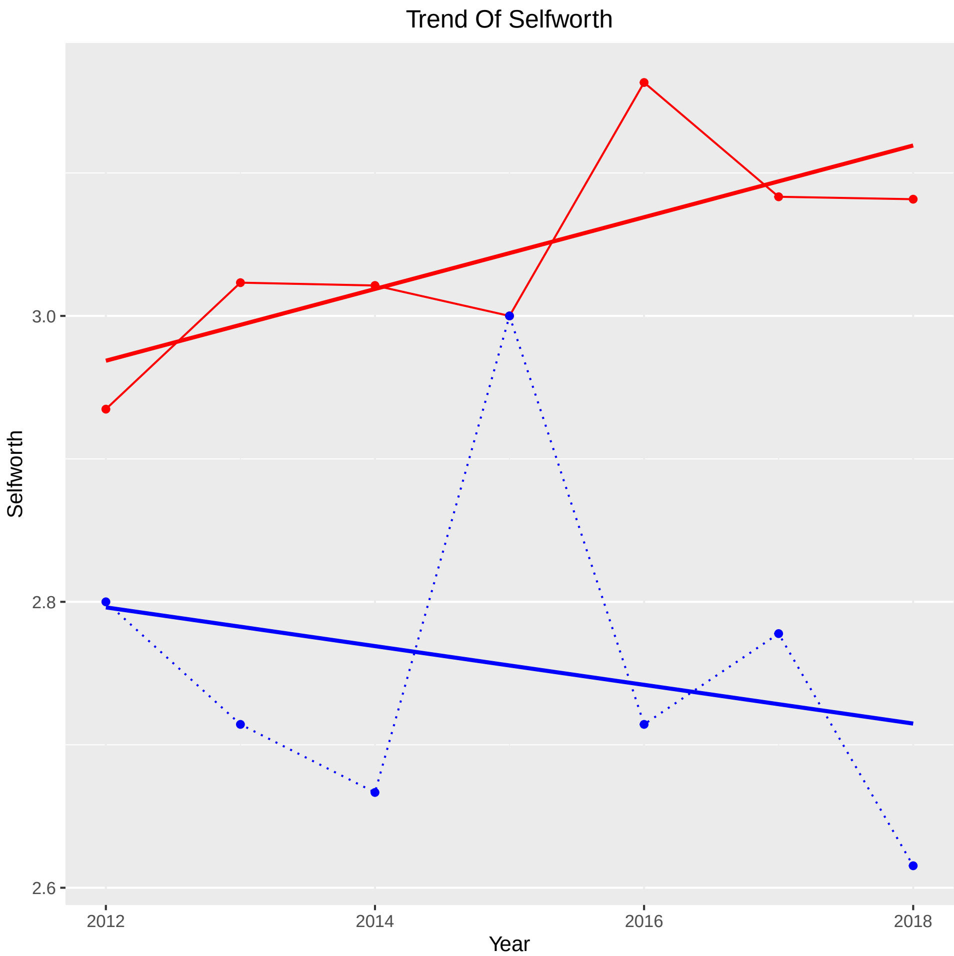


Figure 5: Trend in share of beneficiaries with improved selfworth

Note: The x-axis represents the years from 2012 to 2018, the year 2011 is left out because the trips program starts in 2012. The y-axis represents the average answers from the organizations regarding to selfworth. The time trend of the average answers of the organizations in the treatment group is characterised by the solid red line, the answers of the control group by the dotted blue line. Additionally, the linear trends of both groups are included as the straight lines.

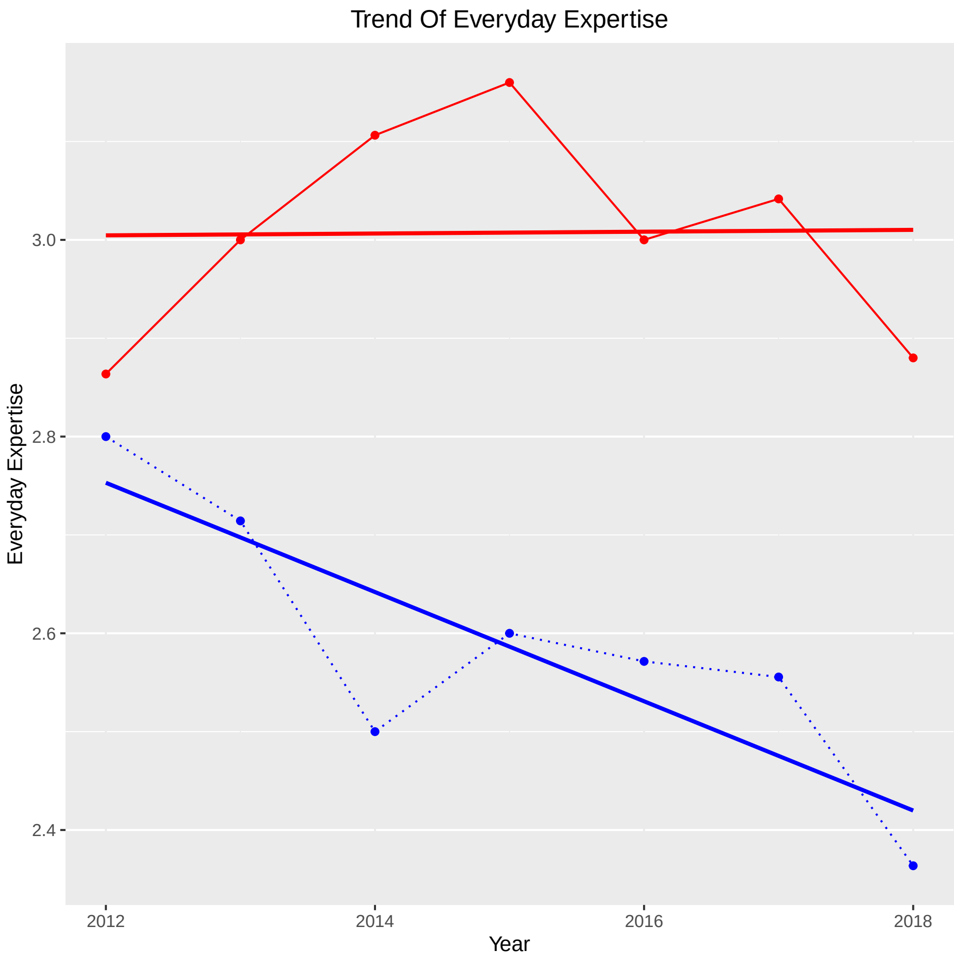


Figure 6: Trend in share of beneficiaries with broadened everyday competence

Note: The x-axis represents the years from 2012 to 2018, the year 2011 is left out because the trips program starts in 2012. The y-axis represents the average answers from the organizations regarding to everyday expertise. The time trend of the average answers of the organizations in the treatment group is characterised by the solid red line, the answers of the control group by the dotted blue line. Additionally, the linear trends of both groups are included as the straight lines.

expertise. In the treatment group the average selfworth increased over time while it decreased in the control group (Figure 5). Moreover, average everyday expertise declines in the control group over time while it remains constant in the treatment group (Figure 6). The divergence shown in both graphs suggests that funding for the trips program might positively influence the selfworth and everyday expertise of participating children and adolescents. This evidence could support the hypothesis that the trips program positively influences the beneficiaries. However, this graphical analysis is only descriptive and therefore cannot be interpreted as a causal relationship.

9.1.4 Differences-in-Differences Estimation

For the empirical analysis, we implement a differences-in-differences (DID) strategy to test whether the trips program has a positive influence on the selfworth and everyday expertise of the supported children. The DID estimator measures the effect of participating in the trips program by comparing the changes in dependent variables over time between the treatment and control group. The key identifying assumption for the DID strategy is the common trend assumption. The assumption states that, in the absence of the trips program, both the treatment and control group would have evolved with the same trend meaning that the difference between the groups would have stayed the same. In case of a violation of the parallel trend assumption, the estimated treatment effect would be biased. As the dataset contains 2011 as the only year in the pre-period, we are not able to observe a pre-trend. Therefore, we cannot argue that the common trend assumption is fulfilled.

Using the panel structure of the dataset we implement the DID estimation with the following regression equation:

$$Y_{it} = \alpha + \beta \cdot TreatEF_{it} + \gamma_i + \delta_t + \epsilon_{it} \tag{4}$$

where i indexes the identification number of each organization supported by CHILDREN and t indexes the year of observation. The outcome variable, denoted by Y_{it}, is either selfworth or everyday expertise. As mentioned in the previous section, $TreatEF_{it}$ represents the treatment status of organization i in year t. The corresponding regression coefficient β represents the DiD estimator, which measures the average treatment effect of participating in the trips program.

The panel data set allows us to implement fixed effects. In our analysis we introduce individual fixed effects and time fixed effects. The ID fixed effects γ_i control for organization specific observable and unobservable characteristics that are constant over time but differ across social institutions. For example, the state of an organization does not vary over time but might differ across supported social institutions. Additionally, the year fixed effects δ_t capture all variables that change across years but are the same for all organizations and within a specific year. For all following regression estimations, we use robust standard errors to take potential heteroscedasticity into account.

	Dependent variable:			
	Everyday Expertise			
	(1)	(2)	(3)	(4)
treatEF	−0.010	−0.166	0.247	0.255
	(0.326)	(0.405)	(0.299)	(0.310)
subsidy		0.019		0.016
		(0.014)		(0.014)
totalCost		0.001^{**}		0.001^{*}
		(0.000)		(0.000)
weeklyCooks		0.166^{**}		0.162^{**}
		(0.072)		(0.073)
ID fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Number of observations	428	410	428	410
R^2	0.475	0.490	0.476	0.491

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 9: DID Estimation: Results for Everyday Expertise

One concern with our empirical approach is that the allocation to treatment and control group might not be perfectly randomized because the selection into the treatment could be driven by time-variant characteristics. Therefore, we include a set of control variables to deal with potential selection bias. For this, we select variables of the available dataset that might influence both the treatment status and either of the outcome variables. Furthermore, we control for organization specific characteristics which include the subsidy received for the meals program, the corresponding total costs of providing meals and the number of days in a week on which children cook in a specific social institution.

9.2 Results

9.2.1 Everyday Expertise

The estimated coefficients of equation XX are represented in Table **??** with everyday expertise as the dependent variable. Columns (1)-(2) report estimates using the first definition of the treatment variable, while columns (3)-(4) apply the second treatment specification. Columns (1) and (3) only include ID fixed effects and year fixed effects without any controls. Columns (2) and (4) expand the regression equation by adding the organization specific control variables: Subsidy, total costs and the number of days in a week on which children cook in a social institution.

Using the first definition of the treatment variable, the average treatment effect of participating in the trips program on everyday expertise is negative, but not statistically significant. Therefore, the estimated coefficient in column (1) implies that the trips program would not have a significant effect on everyday expertise of the supported children and adolescents. However, this negative estimate could result from the fact that we consider several organizations as treated even though they did not receive funding for the trips program in a given year. As a result, the number of observation units in the control group is reduced so that the difference in size between the treatment and control group is considerably high. Therefore, this finding should not be overstated. On the other hand, using the alternative specification of the treatment indicator in column (3), the average treatment effect is positive, but remains insignificant. Thus, the sign of the estimated treatment effect changes if we use the alternative definition of the treatment variable. With the second treatment variable, the number of organizations in the control group increases, as the treatment status depends on actually receiving funding in a given year. Nevertheless, the difference in size of the treatment and control group remains relatively large. However, the estimate in Column (3) could provide evidence that the treatment effect might be positive and potentially turn significant if the sample size increases.

Adding the organization specific control variables in Column (2) and Column (4) does not influence the size of both effects and the significance substantially.

9.2.2 Selfworth

Using selfworth as the outcome variable, the estimates of regression equation **??** are reported in Table **??**, whose structure is equal to Table **??**.

Strinkinlgy, all columns of Table **??** report a negative treatment effect, regardless of the definition of the treatment variable or the inclusion of organization specific controls. In column (1)-(3), the average treatment effect is insignificant. However, the estimated effect is significant if we use the second specification of the treatment status and include controls in column (4). This indicates that participating in the trips program would negatively influence the selfworth of the children and adolescents. This surprising result contradicts our hypothesis that the participation in the trips program positively influences the selfworth of the beneficiaries. As previously mentioned, the number of observation units in both treatment and control group is relatively small. Therefore, this unexpected negative effect of the trips program on selfworth should not be overstated. Moreover, another possible explanation for the observed results could be that the survey questions are answered by employees of an organization and not by the children and adolescents themselves. It might be difficult for the respondents of the survey to assess children specific characteristics, such as selfworth. During our visit in Augsburg, the employees confirmed that it is challenging to evaluate the variables asked in the survey for all participants in the entire social institution, especially with a lot of variation in attendance. One potential

	Dependent variable:			
	Selfworth			
	(1)	(2)	(3)	(4)
treatEF	−0.474	−0.481	−0.328	−0.442*
	(0.309)	(0.312)	(0.247)	(0.256)
subsidy		0.011		0.014
		(0.018)		(0.017)
totalCost		0.000		0.000
		(0.001)		(0.001)
weeklyCooks		0.036		0.037
		(0.069)		(0.069)
ID fixed effects	*Yes*	*Yes*	*Yes*	*Yes*
Year fixed effects	*Yes*	*Yes*	*Yes*	*Yes*
Number of observations	428	410	428	410
R^2	0.475	0.484	0.474	0.485

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 10: DID Estimation: Results for Selfworth

solution for this problem could be to directly ask the children and adolescents. Data on the individual level might be more precise because the children could assess themselves better. When using variables on the level of beneficiaries, we might observe a positive effect of the trips program on selfworth. As reported in the summary statistics, the outcome variable shows a low variation. This results from the fact that, most of the time, organizations give the answer “all children” and “most children” when asked about the improvement of children’s selfworth. In general, it is more unlikely to find significant effects if the variation of the dependent variable is low. Therefore, one challenge for further analysis is to ask for variables that have a potentially higher variation.

10 Conclusion

Until now CHILDREN only influences whether organizations are supported and how much funding they receive to provide both the trips and the meals program. However, we find a strong positive correlation between the healthiness of the meals that organizations serve and health related outcomes of children and adolescents. During our visit of an organization in Augsburg, the employees told us that both the children and adolescents appreciate the healthy food and demand even more healthier meals. This might be one possible way for CHILDREN to improve the beneficiaries’ health without the necessity of spending particularly more money. In the yearly meetings, they could point out our results and mention their appreciation of the provision of healthy food. Moreover, CHILDREN could request that the organizations fulfill the DGE criteria for the provided meals. Due to the good relationship between CHILDREN and the supported social institutions, we think that the organizations are open to implementing this proposal.

On the basis of the processed dataset, we cannot measure the causal effect of the trips program on everyday expertise and selfworth of beneficiaries. However, measuring significant effects of CHILDREN's trips program might be possible if the data structure will be adjusted. At this point, we recommend to directly ask the beneficiaries to answer the children specific questions of the survey. During our first meeting with CHILDREN we were informed that they already have collected data from children and adolescents directly, that caused additional workload. For instance, the organizations need to obtain the permission of the parents. Nevertheless, the individual level data are perhaps more precise.

The dataset contains several helpful features that made our analysis easier. For example, CHILDREN has asked for many survey questions from the supported organizations in every year across the whole observation period. Furthermore, the dataset includes variables regarding increased selfworth and broadened everyday expertise for both programs allowing a comparison of the meals and the trips program. Hence, we recommend to continue asking the same questions in every year and selected variables for both programs.

11 Appendix

11.1 A1: Expanded Health Regression

In addition to table **??** table **??** presents the results of an OLS approach with scaled outcomes as well as a WLS approach with scaled outcomes with following control variables: to which extent an organization uses regional products, since when an organization is part of the Lunch program, the real subsidy per beneficiary and the state an organization is located. In comparison to the simple linear model, there are only small changes in coefficents.

graphicx

11.2 A2: Cumulative Odds Regression

As an option to run a regression on the original ordinal variables, we use the proportional odds version of the cumulative logit. Equation **??** describes the model setup:

$$P(Y_i \leq\ r) = \frac{exp(y_{0r} + x_i^T y)}{1 + exp(y_{0r} + x_i^T y)} \tag{5}$$

	OLS	WLS	OLS Impute	WLS Impute
(Intercept)	-1.08^{***}	-0.99^{**}	-0.04	-0.08
	(0.24)	(0.30)	(0.18)	(0.20)
DGECriteriaNoScaled	0.34^{***}	0.36^{***}	0.23^{***}	0.27
	(0.08)	(0.06)	(0.07)	(0.14)
regionalProducts_scaled	0.02	-0.03	0.11	0.02
	(0.08)	(0.11)	(0.07)	(0.11)
yearsSupportSince	0.02	-0.03	0.01	0.05^{*}
	(0.03)	(0.04)	(0.02)	(0.02)
realSubsidyPerBeneficiary	-0.00	0.00	0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)
stateBayern	0.96^{**}	1.88^{***}		
	(0.33)	(0.49)		
stateBerlin	1.09^{**}	1.34^{***}		
	(0.37)	(0.37)		
stateBrandenburg	2.21^{***}	2.46^{***}		
	(0.43)	(0.40)		
stateBremen	1.00	1.03		
	(0.58)	(0.59)		
stateHamburg	0.95	1.96^{***}		
	(0.67)	(0.40)		
stateHessen	0.87^{***}	1.16^{***}		
	(0.23)	(0.31)		
stateMV	0.09	0.14		
	(0.21)	(0.26)		
stateNiedersachsen	2.48^{***}	2.30^{***}		
	(0.40)	(0.43)		
stateNRW	0.68^{*}	1.00^{**}		
	(0.27)	(0.33)		
stateSaarland	1.50^{***}	1.75^{***}		
	(0.32)	(0.25)		
stateSachsen	1.34^{***}	1.23^{***}		
	(0.27)	(0.27)		
stateSchleswig-Holstein	1.33^{***}	1.63^{***}		
	(0.37)	(0.42)		
stateThüringen	0.99	1.29^{*}		
	(0.59)	(0.55)		
R^2	0.51	0.73	0.09	0.22
Adj. R^2	0.39	0.66	0.07	0.20
Num. obs.	88	88	175	175
RMSE	0.73	5.62	0.94	7.77

Dependent variable: share of beneficiaries who are less frequently ill
DGECriteriaNo: index of healthy diet criteria fulfilled in organization's menu
regionalProducts: wether the organization offers meals with local ingredients or not
yearsSupportSince: the number of years an organization is already part of the Meals program
realSubsidyPerBeneficiary: subsidy per beneficiary of Meals program in 2015 EUR
state: german 'Bundeland'
Model (1): original data set, simple linear model with controls, estimated with OLS
Model (2): original data set, simple linear model with controls, estimated with WLS
Model (3): imputed data set, simple linear model with controls, estimated with OLS
Model (4): imputed data set, simple linear model with controls, estimated with WLS
All regressions are estimated with robust standard errors $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

Table 11: Regression Results: Less ill expanded model

To demonstrate the results of this method, we show in the tables **??**, **??**, **??** the regressions of the health related variables on the healthy meals criterion and in the tables **??** and **??** the regressions of the real meals subsidy on selfworth and everyday expertise.

As an example of interpreting such models, we consider the model of table **??**: An increase of the healthy meals criterion by 1 is associated with an increase in chances, that a proportion of a maximum of r beneficiaries is healthier in relation to that a proportion of more than r beneficiaries are healthier, by the factor of exp(-0.29127) = 0.75.

A summary of the remaining models:

- Dietary Knowledge: exp(-0.089) = 0.91
- Appreciate Healthy: exp(-0199) = 0.82
- Selfworth: exp(-0.00001) = 1
- Everyday Expertise: (-0.00003) = 1

```
## Warning:  namespace 'VGAM' is not available and has been replaced
## by .GlobalEnv when processing object ''
```

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept):1	−2.799	1.109	−2.523	0.012
(Intercept):2	1.653	0.582	2.841	0.004
(Intercept):3	4.667	0.738	6.322	0
DGECriteriaNo	−0.291	0.075	−3.883	0.0001

Table 12: Propodss Regression Results: Association of index of healthy diet criteria fulfilled in organization's menu and the share of beneficiaries who are less frequently ill

```
## Warning:  namespace 'VGAM' is not available and has been replaced
## by .GlobalEnv when processing object ''
```

```
## Warning:  namespace 'VGAM' is not available and has been replaced
## by .GlobalEnv when processing object ''
```

```
## Warning:  namespace 'VGAM' is not available and has been replaced
## by .GlobalEnv when processing object ''
```

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept):1	−4.009	0.803	−4.996	0.00000
(Intercept):2	−1.047	0.425	−2.465	0.014
(Intercept):3	1.445	0.430	3.365	0.001
DGECriteriaNo	−0.089	0.052	−1.712	0.087

Table 13: Propodss Regression Results: Association of index of healthy diet criteria fulfilled in organization's menu and share of beneficiaries with expanded dietary knowledge

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept):1	−1.603	0.486	−3.298	0.001
(Intercept):2	0.586	0.413	1.419	0.156
(Intercept):3	3.052	0.471	6.483	0
DGECriteriaNo	−0.199	0.053	−3.744	0.0002

Table 14: Propodss Regression Results: Association of index of healthy diet criteria fulfilled in organization's menu and the share of beneficiaries with increased appreciation for a healthy diet

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept):1	−4.855	0.584	−8.315	0
(Intercept):2	−1.268	0.143	−8.893	0
(Intercept):3	1.514	0.150	10.109	0
realSubsidy	−0.00001	0.00001	−1.332	0.183

Table 15: Propodss Regression Results: Association of subsidy for Meals program in 2015 EUR and the share of beneficiaries with improved self-worth

```
## Warning:  namespace 'VGAM' is not available and has been replaced
## by .GlobalEnv when processing object ''
```

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept):1	−3.543	0.343	−10.330	0
(Intercept):2	−0.736	0.132	−5.563	0.00000
(Intercept):3	1.764	0.157	11.249	0
realSubsidy	−0.00003	0.00001	−4.070	0.00005

Table 16: Propodss Regression Results: Association of subsidy for Meals program in 2015 EUR and the share of beneficiaries with broadened everyday expertise

11.3 A3: Partition

In addition to the factor analysis which is described in section 6, we would like to introduce to a another dimensionality reduction method called partition. This method was developped by Millstein et. al in the context of genome analysis.

The partition approach partitions data into subsets of related features. It then summarizes each subset into one new feature. By that, it defines a surjective mapping (**Millstein.2020**). A function g is implemented using the arithmetic mean. In contrast to methods like the factor analysis, if a feature is identified as related to a response of interest, the exact subset of features that are implicated to the surjective mapping is known without ambiguity. What is as well different is the information loss constraint. It lets the extent of dependences among features on a local level guide the extent to which the number of features is reduced (**Millstein.2020**).

In the following a threshold of 0.4 is used, meaning that the reduced variable consists of variables which explain each other to at least 40 percent. It might be meaningful to decide to use only one of the featured variables in a reduced variable (or a summarizing one) for future surveys to avoid redundancy.

Table **??** shows the obtained results of the dimensionality reduction for the variables of the Lunch program. It displays 10 variables which haven't been reduced, and 5 reduced variables with a number of featured variables.

Table **??** shows the partitioned results for the variables of the Trips program. Here, 12 variables haven't been reduced, and a number of featured variables are summarized to 3 reduced variables.

11.4 A4: OLS Regressions

11.5 A5: DID Regression Results

The estimated coefficients of equation **??** are represented in Table **??** with everyday expertise as the dependent variable and in Table **??** with selfworth as the dependent variable. P-values are in parentheses. We observe decreasing p-values when using the second definition of the treatment variable.

12 Ehrenwörtliche Erklärung aller Teilnehmer

	Variable, Meals	Mapping, Meals	Information, Meals
1	participateMore	participateMore	1.00
2	tasksLunch	tasksLunch	1.00
3	ownIdeas	ownIdeas	1.00
4	stayLonger	stayLonger	1.00
5	dietaryKnowledge	dietaryKnowledge	1.00
6	appreciateHealthy	appreciateHealthy	1.00
7	foodCulture	foodCulture	1.00
8	lessIll	lessIll	1.00
9	betterTeamwork	betterTeamwork	1.00
10	moreRegularSchoolVisits	moreRegularSchoolVisits	1.00
11	addressProblems	addressProblems	1.00
12	reduced_var_1	moreConcentrated	0.66
13	reduced_var_1	moreBalanced	0.66
14	reduced_var_2	monthlyCooks	0.42
15	reduced_var_2	weeklyCooks	0.42
16	reduced_var_2	shoppers	0.42
17	reduced_var_2	easyDishes	0.42
18	reduced_var_3	dayToDaySkills	0.43
19	reduced_var_3	moreIndependent	0.43
20	reduced_var_3	selfworth	0.43
21	reduced_var_3	moreOpen	0.43
22	reduced_var_3	moreConfidence	0.43
23	reduced_var_3	proud	0.43
24	reduced_var_4	betterReading	0.53
25	reduced_var_4	betterNumbers	0.53
26	reduced_var_4	betterGrades	0.53
27	reduced_var_5	influenceHome	0.41
28	reduced_var_5	cookAtHome	0.41
29	reduced_var_5	askRecipes	0.41

Table 17: Partition of outcomes, Meals

	Variable, Trips	Mapping, Trips	Information, Trips
1	tripsSuggestions	tripsSuggestions	1.00
2	tripsDecisions	tripsDecisions	1.00
3	tripsOrganization	tripsOrganization	1.00
4	tripsCostCalculation	tripsCostCalculation	1.00
5	tripsBudget	tripsBudget	1.00
6	tripsMoney	tripsMoney	1.00
7	tripsReview	tripsReview	1.00
8	tripsPublicTransport	tripsPublicTransport	1.00
9	tripsMobility	tripsMobility	1.00
10	tripsAdditionalActivities	tripsAdditionalActivities	1.00
11	tripsSelfworth	tripsSelfworth	1.00
12	tripsFrustrationTolerance	tripsFrustrationTolerance	1.00
13	reduced_var_1	tripsSuccess	0.68
14	reduced_var_1	tripsSelfEfficacy	0.68
15	reduced_var_2	tripsNewPlaces	0.60
16	reduced_var_2	tripsNewCommunities	0.60
17	reduced_var_2	tripsNewIdeas	0.60
18	reduced_var_2	tripsSocialSkills	0.60
19	reduced_var_3	tripsSpecificSkills	0.46
20	reduced_var_3	tripsDayToDaySkills	0.46

Table 18: Partition of outcomes, Trips

	Dependent variable:			
	Everyday Expertise			
	(1)	(2)	(3)	(4)
treatEF	−0.010	−0.166	0.247	0.255
	(0.975)	(0.681)	(0.409)	(0.410)
subsidy		0.019		0.016
		(0.181)		(0.258)
totalCost		0.001^{**}		0.001^{*}
		(0.048)		(0.060)
weeklyCooks		0.166^{**}		0.162^{**}
		(0.022)		(0.027)
ID fixed effects	YES	YES	YES	YES
Year fixed effects	YES	YES	YES	YES
Number of observations	428	410	428	410
R^2	0.475	0.490	0.476	0.491

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 19: DID Estimation 1

	Dependent variable:			
	Selfworth			
	(1)	(2)	(3)	(4)
treatEF	−0.474	−0.481	−0.328	-0.442^{*}
	(0.126)	(0.125)	(0.185)	(0.085)
subsidy		0.011		0.014
		(0.540)		(0.422)
totalCost		0.000		0.000
		(0.575)		(0.446)
weeklyCooks		0.036		0.037
		(0.609)		(0.591)
ID fixed effects	YES	YES	YES	YES
Year fixed effects	YES	YES	YES	YES
Number of observations	428	410	428	410
R^2	0.475	0.484	0.474	0.485

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 20: DID Estimation 2