

# Analyse der Survey-Daten von CHILDREN for a better World e.V.

Laura Huber

Laura Jepsen

Jonathan Kirschner

Rafael Schütz

Yannick Zurl

Studentisches Praxisprojekt zur Empirischen Wirtschaftsforschung

PaRE3To

Ludwig-Maximilians-Universität München

3. März 2020

# Table of Contents I

- 1 Einleitung
- 2 Datenaufbereitung
- 3 Zusammenfassende Statistiken
  - Überblick: Entwicklung der Anzahl der geförderten Einrichtungen
  - Entwicklung der Fördersummen über die Zeit
  - Dynamiken des Selbstwertgefühls und der Alltagskompetenzen
  - Dynamiken gesundheitsrelevanter Variablen
- 4 Explorative Faktoranalyse
- 5 Partition
- 6 Zusammenhänge zwischen CHILDRENs Zuschüssen, dem DGE-Kriterium und ausgewählten Variablen
  - Empirischer Ansatz
  - Assoziationen zwischen Fördersumme und ausgewählten Variablen
  - Gesundheit
- 7 Effekte des Entdeckerfonds
  - Idee
  - Empirische Methode
  - Ergebnisse

# CHILDREN for a better World e.V.

- Spendenfinanzierte Kinderhilfsorganisation
- Finanziert Einrichtungen in ganz Deutschland
- CHILDREN Mittagstisch: Bereitstellung von Mahlzeiten, um eine ausgewogene und gesunde Ernährung der Kinder und Jugendlichen zu fördern
- CHILDREN Entdeckerfonds: Durch Ausflüge und Aktivitäten wird es Kindern und Jugendlichen ermöglicht, neue Erfahrungen zu sammeln

# Datenerhebung von CHILDREN

- Jährliche Umfragebögen an die Einrichtungen
- Beantwortung durch Mitarbeiter
- Verschiedene Fragen zu Mittagstisch und Entdeckerfonds
- Allgemeine Variablen zu den Einrichtungen (z.B. Fördersumme, Anzahl angebotener Mahlzeiten, Anzahl Aktivitäten)
- Abhängige Variablen zu den Kindern (z.B. Selbstwertgefühl, seltener krank)

# Datenaufbereitung

- Zusammenfügen der Datensätze aus den verschiedenen Jahren
- Anpassung der Variablennamen und Hinzufügen von neuen Variablen
- Änderung der Datentypen von Variablen

⇒ Erstellung eines finalen Datensatzes

- Datenauswertung mit dem Statistik-Programm „R“
- Versionskontrolle über „Git“

# Beispiel: Datensatz

	ID	Jahr	Anzahl der Kinder MT	Selbstwertgefühl	Ausflüge
1	103	2014	30	3	4
2	103	2015	50	3	6
3	103	2016	NA	2	3
4	112	2014	27	3	NA
5	112	2015	55	3	NA
6	112	2016	35	3	NA

Tabelle: Beispielhafter Datensatz

# Zusammenfassende Statistiken

	Jahr	Begünstigte, Mittagstisch	Begünstigte, Entdeckerfonds	Einrichtungen, Mittagstisch	Einrichtungen, Entdeckerfonds
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

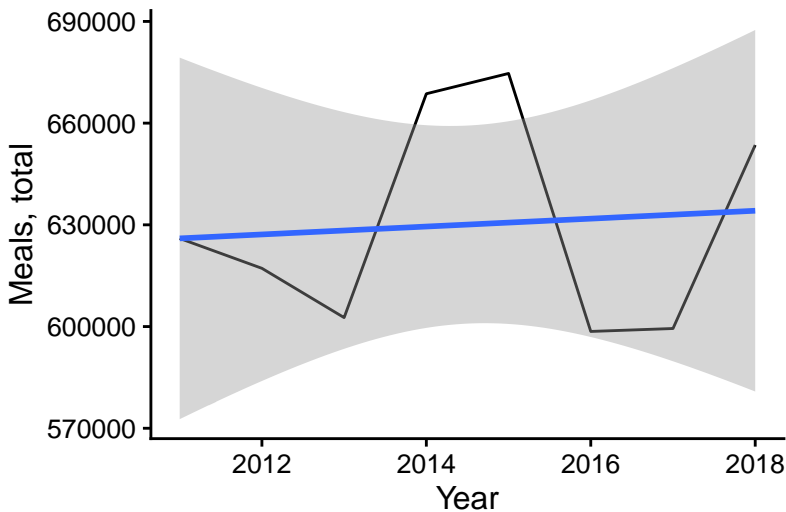
Tabelle: Zusammenfassende Statistiken

# Umrechnung der Fördersummen: Reale Werte

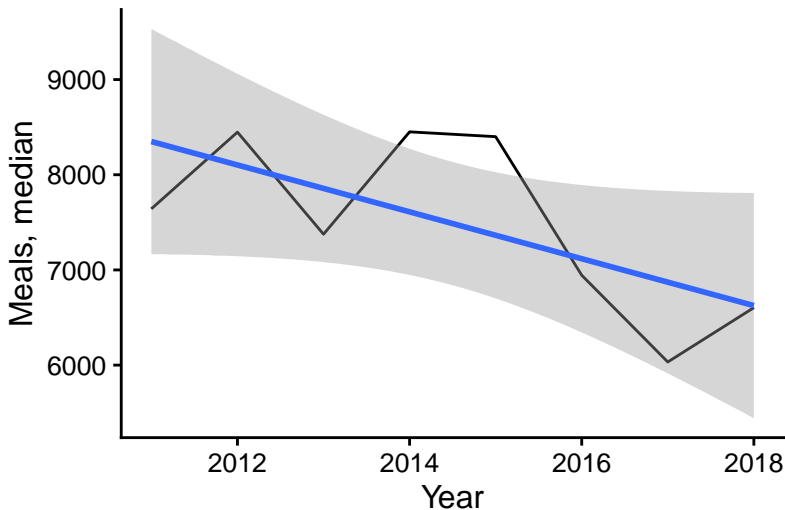
- Zur besseren Vergleichbarkeit der Fördersummen über die Zeit: Umrechnung in 2015 EUR
- Verwendung der Preisindizes des statistischen Bundesamtes
- Mittagstisch Fördersumme: Preisindex für Nahrungsmittel und alkoholfreie Getränke
- Entdeckerfonds Fördersumme: Preisindex für Freizeit, Unterhaltung und Kultur



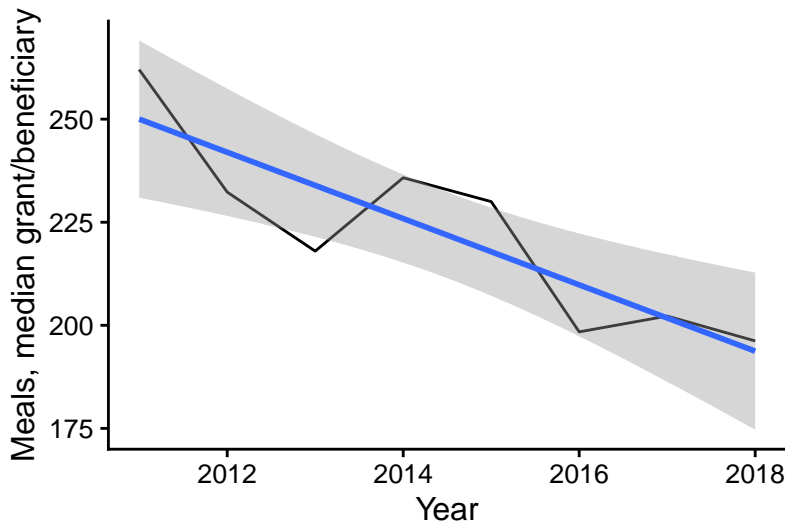
# Dynamik der Fördersumme, Mittagstisch: Summe



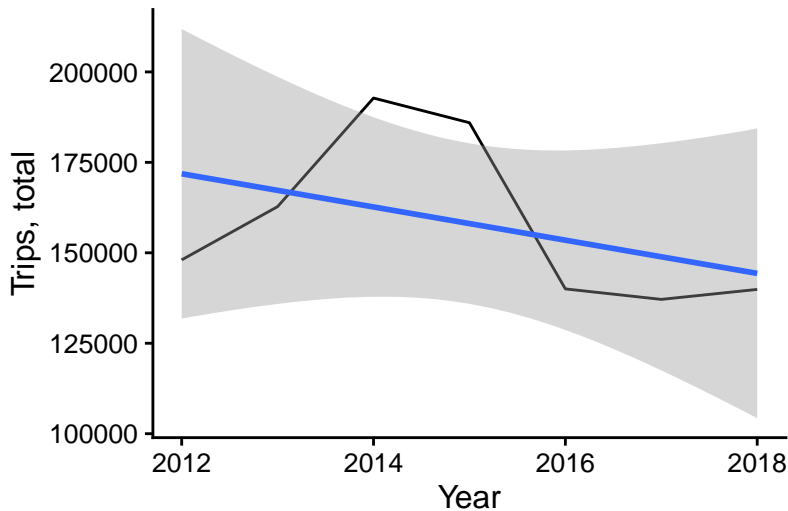
# Dynamik der Fördersumme, Mittagstisch: Median



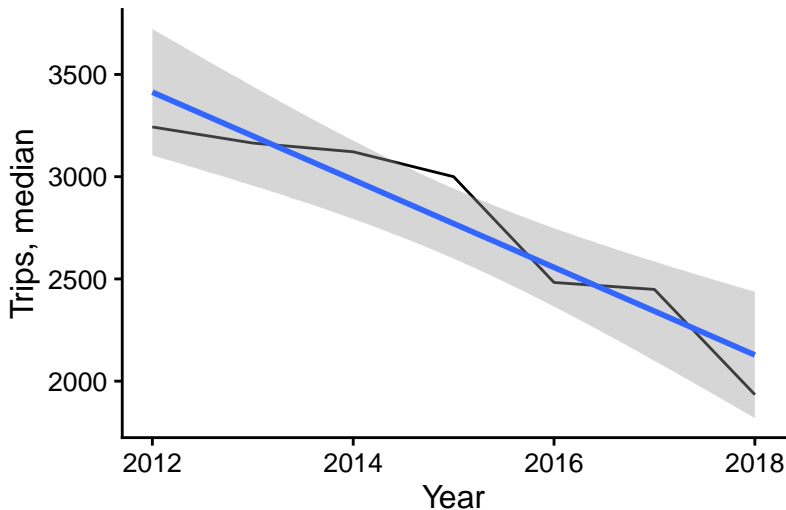
# Dynamik der Fördersumme, Mittagstisch: Median pro Begünstigter



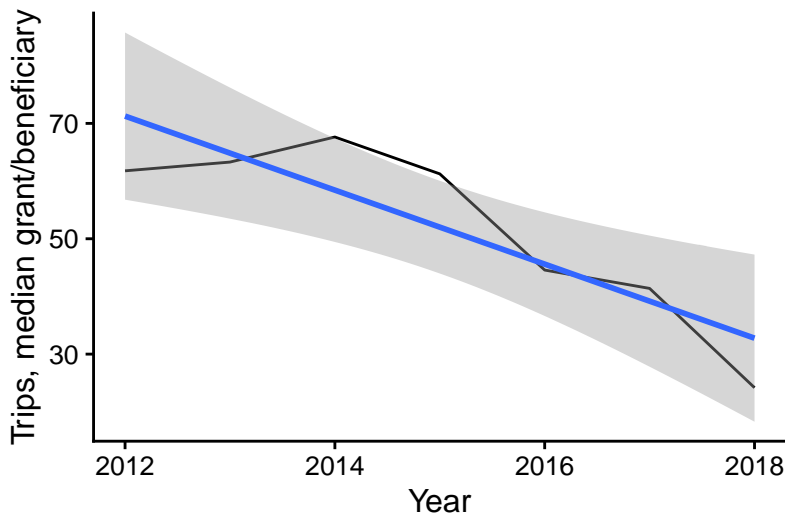
# Dynamik der Fördersumme, Entdeckerfonds: Summe



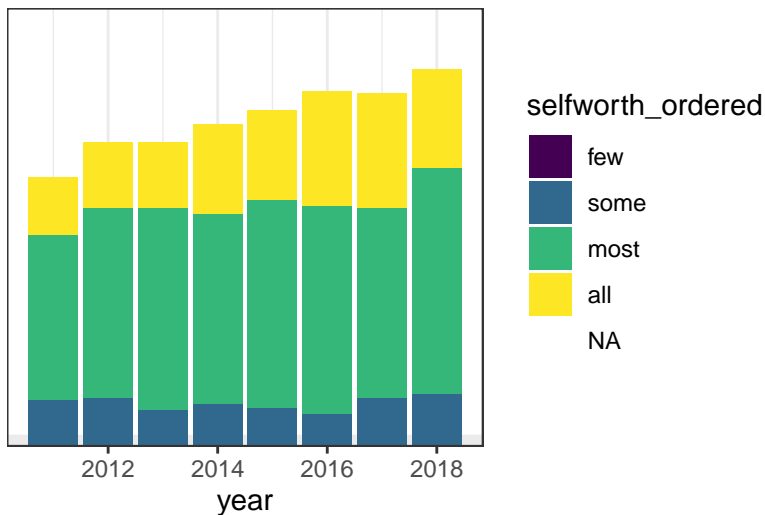
# Dynamik der Fördersumme, Entdeckerfonds: Median



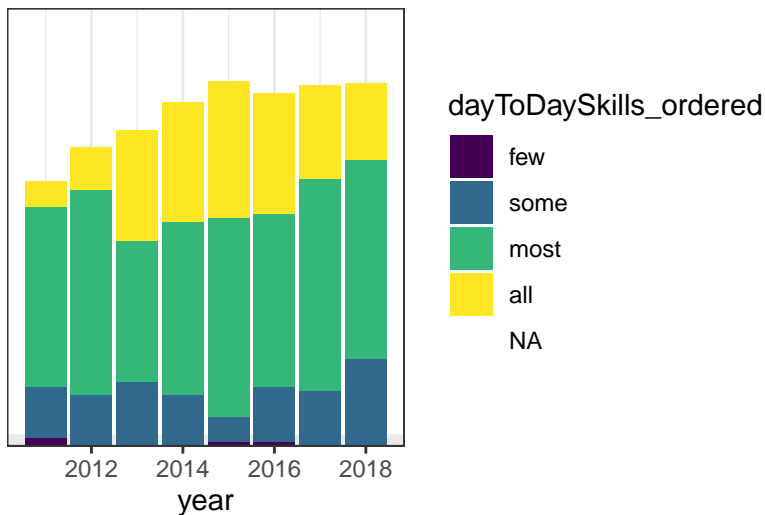
# Dynamik der Fördersumme, Entdeckerfonds: Median pro Begünstigter



# Variable „Selbstwertgefühl“: Dynamik

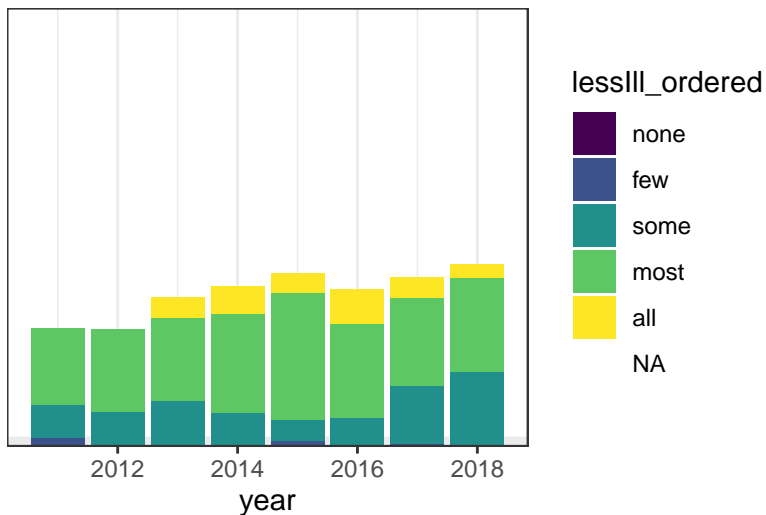


# Variable „Alltagskompetenzen“: Dynamik

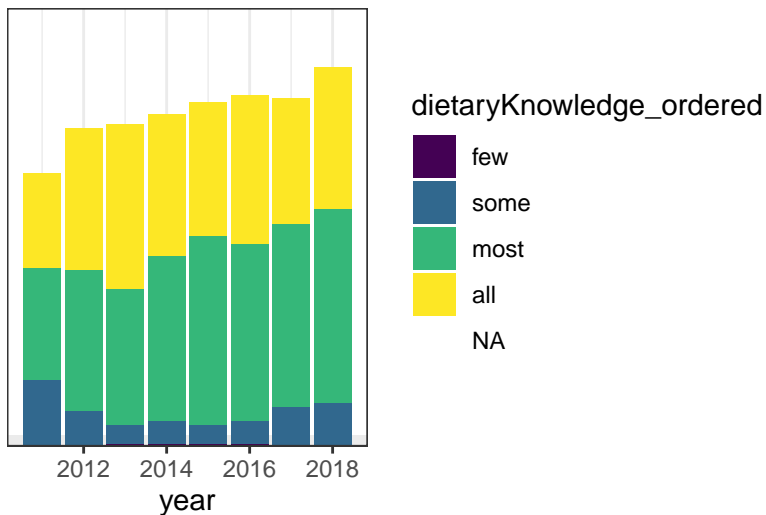




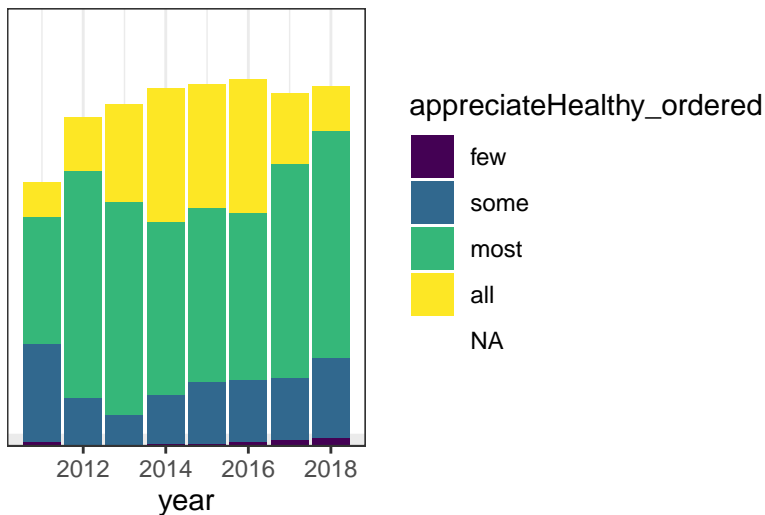
# Variable „seltener krank“: Dynamik



# Variable „erweitertes Ernährungswissen“: Dynamik



# Variable „Wertschätzung gesunder Ernährung“: Dynamik



# Situation

- Viele Zielvariablen

# Situation

- Viele Zielvariablen
- Teilweise sehr ähnliche Zielvariablen, z.B. Begünstigte „kochen mindestens einmal im Monat in der Einrichtung“ und „kochen mindestens einmal in der Woche in der Einrichtung“

# Situation

- Viele Zielvariablen
- Teilweise sehr ähnliche Zielvariablen, z.B. Begünstigte „kochen mindestens einmal im Monat in der Einrichtung“ und „kochen mindestens einmal in der Woche in der Einrichtung“
- Welche Variablen liegen zugrunde?

# Situation

- Viele Zielvariablen
- Teilweise sehr ähnliche Zielvariablen, z.B. Begünstigte „kochen mindestens einmal im Monat in der Einrichtung“ und „kochen mindestens einmal in der Woche in der Einrichtung“
- Welche Variablen liegen zugrunde?
- Variablenreduktion zur Bestimmung möglicher Kontrollen

# Technisches Vorgehen

- Imputierter Datensatz



# Technisches Vorgehen

- Imputierter Datensatz
- Separate Faktoranalyse für jedes später zu schätzende Modell:  
Abhängige Variable und Prädiktor ausgeschlossen

# Technisches Vorgehen

- Imputierter Datensatz
- Separate Faktoranalyse für jedes später zu schätzende Modell:  
Abhängige Variable und Prädiktor ausgeschlossen
- Jede Variable als Linearkombination von wenigen Faktoren

# Technisches Vorgehen

- Imputierter Datensatz
- Separate Faktoranalyse für jedes später zu schätzende Modell:  
Abhängige Variable und Prädiktor ausgeschlossen
- Jede Variable als Linearkombination von wenigen Faktoren
- Schätzung an sich uneindeutig  $\Rightarrow$  rotiere Faktoren so, dass orthogonal zueinander

# Beispiel Faktoranalyse Mittagstisch

	ML2	ML1
participateMore_scaled	0.14	0.07
monthlyCooks_scaled	0.22	0.79
weeklyCooks_scaled	0.17	0.85
shoppers_scaled	0.16	0.55
easyDishes_scaled	0.47	0.44
dietaryKnowledge_scaled	0.64	0.20
appreciateHealthy_scaled	0.74	0.00
foodCulture_scaled	0.56	0.14
dayToDaySkills_scaled	0.57	0.34
selfworth_scaled	0.54	0.26
influenceHome_scaled	0.46	0.17

**Tabelle:** Faktoranalyse Mittagstisch ohne 'seltener krank'

# Alternative zu Faktoranalyse?

- Neue Methode aus Biostatistik zur Dimensionalitätsreduktion: Partition (Millstein et al., 2020)

# Alternative zu Faktoranalyse?

- Neue Methode aus Biostatistik zur Dimensionsreduktion: Partition (Millstein et al., 2020)
- Aufteilung der Menge aller Variablen in Gruppen

# Alternative zu Faktoranalyse?

- Neue Methode aus Biostatistik zur Dimensionsreduktion: Partition (Millstein et al., 2020)
- Aufteilung der Menge aller Variablen in Gruppen
- Jeder Variable wird genau eine zugrundeliegende Variable zugewiesen
- In jeder Gruppe muss Intraklassen-Korrelationskoeffizient mindestens bestimmten, willkürlich festgelegten Wert annehmen

# Alternative zu Faktoranalyse?

- Neue Methode aus Biostatistik zur Dimensionsreduktion: Partition (Millstein et al., 2020)
- Aufteilung der Menge aller Variablen in Gruppen
- Jeder Variable wird genau eine zugrundeliegende Variable zugewiesen
- In jeder Gruppe muss Intraklassen-Korrelationskoeffizient mindestens bestimmten, willkürlich festgelegten Wert annehmen
- Bestimmung der Gruppen durch deterministischen Nächste-Nachbarn-Algorithmus



# Partition Mittagstisch

	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

Tabelle: Partition der Zielvariablen, Mittagstisch

# Partition Entdeckerfonds

	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

Tabelle: Partition der Zielvariablen, Entdeckerfonds

# Empirischer Ansatz

$$y_{it} = \beta_0 + \beta_1 x_{it} + \epsilon_{it} \quad (1)$$

- Schätzung der Modelle mit OLS (Methode der kleinsten Quadrate)
  - Problem: Einrichtungen werden gleich gewichtet, unabhängig von Größe
  - Lösung: Modellschätzung mit WLS (Gewichtete kleinste Quadrate)
- ⇒ Gewichte: Anzahl der regelmäßigen Teilnehmer am Mittagstisch und im Entdeckerfonds

# Variante: Imputieren fehlender Werte

- Problem: Viele Einrichtungen beantworten nicht alle Fragen
  - Lösung: Erstellung eines separaten Datensatzes in welchem fehlende Werte imputiert werden
    - ⇒ Imputieren der Daten mit einrichtungsspezifischem linearen Trend
- Vergleich von Regressionen mit den Daten des originalen Datensatzes mit den imputierten Daten

# Variante: Ausschließen von Ausreißern

- Problem: CHILDREN fördert Einrichtungen mit sehr vielen und sehr wenigen ausgegebenen Essen und unternommenen Ausflügen
- Lösung: Datensätze ohne Ausreißer
- Mittagstisch: Ausreißer bei Anzahl von Essen
- Entdeckerfonds: Ausreißer bei Anzahl von Ausflügen
- Definition eines Ausreißers:
  - Werte, die 1,5 Interquartilsabstände unter dem 25%-Perzentil liegen
  - Werte, die 1,5 Interquartilsabstände über dem 75%-Perzentil liegen

# Assoziationen zwischen:

- Fördersumme für Mittagstisch (in 2015 EUR) und Anzahl ausgegebener Essen
- Fördersumme für Entdeckerfonds (in 2015 EUR) und Anzahl unternommener Ausflüge
- Sowohl für Mittagstisch als auch für Entdeckerfonds:
  - Fördersumme pro Begünstigter (in 2015 EUR) und standardisierter Anteil der Begünstigten mit gestiegenen Selbstwertgefühl
  - Fördersumme pro Begünstigter (in 2015 EUR) und standardisierter Anteil der Begünstigten mit erweiterten Alltagskompetenzen

# Anzahl Mahlzeiten und Fördersumme

Tabelle: Zusammenhang zwischen Anzahl der Mahlzeiten und realer Fördersumme

	(1)	(2)	(3)	(4)	(5)
(Intercept)	-12089.14* (5192.86)	-1814.16 (1765.93)	3535.39*** (498.99)	3107.70*** (508.94)	-12250.60** (4524.09)
realSubsidy	2.61*** (0.57)	0.50** (0.18)	0.29*** (0.05)	0.25*** (0.05)	2.72*** (0.51)
eatersPerMealNo		172.83*** (14.92)		19.00* (8.45)	
R <sup>2</sup>	0.43	0.73	0.13	0.21	0.45
Adj. R <sup>2</sup>	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

Abhängige Variable: Anzahl der Mahlzeiten

realSubsidy: Fördersumme für Mittagstisch (EUR von 2015)

eatersPerMeal: Anzahl der durch Mittagstisch Begünstigten

Modell (1): einfaches lineares Modell, geschätzt mit Methode der kleinsten Quadrate

Modell (2): ursprünglicher Datensatz, lineares Modell mit Kontrollen, geschätzt mit Methode der kleinsten Quadrate

Modell (3): Datensatz ohne Ausreißer, einfaches lineares Modell, geschätzt mit Methode der kleinsten Quadrate

Modell (4): Datensatz ohne Ausreißer, lineares Modell mit Kontrollen, geschätzt mit Methode der kleinsten Quadrate

Modell (5): Datensatz mit durch lineare Interpolation pro Einrichtung imputierten Daten, einfaches lineares Modell, geschätzt mit Methode der kleinsten Quadrate

Alle Standardfehler sind robust. \*\*\* $p < 0.001$ , \*\* $p < 0.01$ , \* $p < 0.05$ .

# Anzahl Ausflüge und Fördersumme

**Tabelle:** Zusammenhang zwischen Anzahl Ausflüge und Fördersumme

	(1)	(2)	(3)	(4)	(5)
(Intercept)	3.7049*** (0.3313)	3.4394*** (0.3359)	2.6236*** (0.2300)	2.3660*** (0.2609)	3.6237*** (0.3253)
realTripsSubsidy	0.0002* (0.0001)	0.0001 (0.0001)	0.0003*** (0.0001)	0.0003*** (0.0001)	0.0002* (0.0001)
tripsKidsNo		0.0059 (0.0032)		0.0043 (0.0027)	
R <sup>2</sup>	0.0474	0.0729	0.0880	0.1241	0.0504
Adj. R <sup>2</sup>	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, estimated 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 \*\*\* $p < 0.001$ , \*\* $p < 0.01$ , \* $p < 0.05$ .



# Selbstwertgefühl und Fördersumme pro Begünstigter

**Tabelle:** Zusammenhang zwischen Selbstwertgefühl und Fördersumme pro Begünstigtem

	(1)	(2)	(3)	(4)	(5)
(Intercept)	0.08 (0.09)	0.12 (0.12)	0.09 (0.09)	0.12 (0.11)	0.23* (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 <sup>2</sup>	0.00	0.01	0.00	0.01	0.30
Adj. R <sup>2</sup>	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 \*\*\* $p < 0.001$ , \*\* $p < 0.01$ , \* $p < 0.05$ .

# Alltagskompetenzen und Fördersumme pro Begünstigter

**Tabelle:** Zusammenhang zwischen Alltagskompetenzen und Fördersumme pro Begünstigtem

	(1)	(2)	(3)	(4)	(5)	(6)
(Intercept)	0.15 (0.09)	0.13 (0.10)	0.14 (0.09)	0.11 (0.10)	0.28* (0.11)	0.08 (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*** (0.06)	0.03 (0.07)
ML2					0.40*** (0.06)	0.16* (0.07)
ML3					0.16** (0.05)	0.19** (0.06)
ML4						0.49*** (0.06)
R <sup>2</sup>	0.01	0.01	0.01	0.01	0.37	0.37
Adj. R <sup>2</sup>	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 \*\*\* $p < 0.001$ , \*\* $p < 0.01$ , \* $p < 0.05$ .

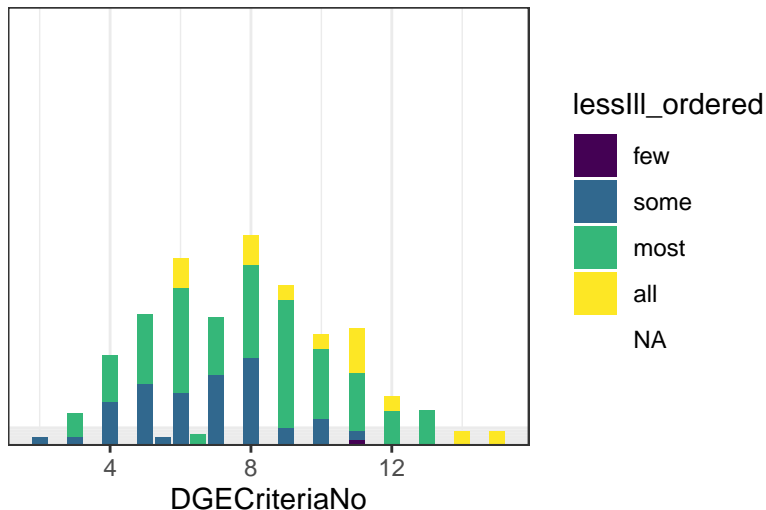
# Zusammenfassung der Ergebnisse

- Mittagstisch ohne Ausreißer: Circa 3 EUR höhere Fördersumme geht einher mit einer weiteren Mahlzeit
- Entdeckerfonds ohne Ausreißer: Circa 3000 EUR höhere Fördersumme geht einher mit einem weiteren Ausflug
- Selbstwertgefühl & die Alltagskompetenzen: kein Zusammenhang mit Fördersumme

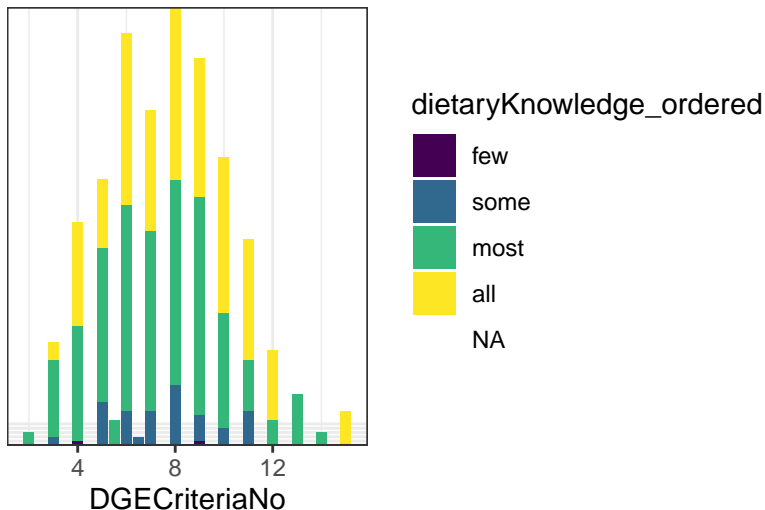
# Fragestellungen, Assoziationen zwischen:

- Standardisiertes Maß für gesundes Essen (DGE-Kriterium) und ausgewählte standardisierte gesundheitsrelevante Variablen
  - Anteil an Begünstigten, die seltener krank sind
  - Anteil an Begünstigten, die ihr Ernährungswissen erweitert haben
  - Anteil an Begünstigten, die gesunde Ernährung stärker wertschätzen
- Schätzung mit WLS

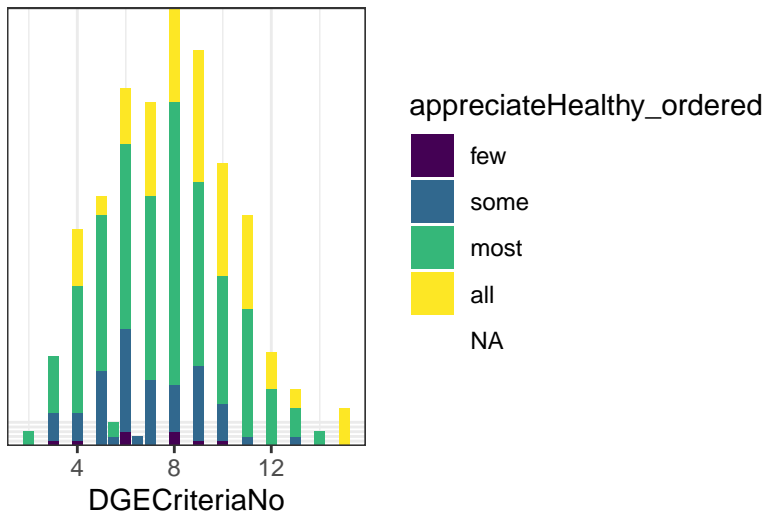
# Variable „seltener krank“: Zusammenhang mit DGE-Kriterium



# Variable „erweitertes Ernährungswissen“: Zusammenhang mit DGE-Kriterium



# Variable „Wertschätzung gesunder Ernährung“: Zusammenhang mit DGE-Kriterium



# Variable „seltener krank“: Zusammenhang mit DGE-Kriterium

Tabelle: Zusammenhang zwischen DGE-Kriterium und Variable 'seltener krank'

	(1)	(2)	(3)	(4)	(5)
(Intercept)	0.02 (0.08)	0.46** (0.16)	0.09 (0.07)	0.39*** (0.12)	0.05 (0.07)
DGECriteriaNoScaled	0.33*** (0.08)	0.35* (0.16)	0.25*** (0.07)	0.24 (0.14)	0.18* (0.07)
ML1					0.12* (0.06)
ML2					0.27*** (0.06)
R <sup>2</sup>	0.12	0.29	0.07	0.16	0.19
Adj. R <sup>2</sup>	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 \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$



# Variable „erweitertes Ernährungswissen“: Zusammenhang mit DGE-Kriterium

**Tabelle:** Zusammenhang zwischen DGE-Kriterium und Variable 'erweitertes Ernährungswissen'

	(1)	(2)	(3)	(4)	(5)
(Intercept)	0.02 (0.07)	0.08 (0.19)	0.02 (0.06)	0.21 (0.18)	0.02 (0.07)
DGECriteriaNoScaled	0.11 (0.06)	-0.02 (0.12)	0.12* (0.05)	0.10 (0.14)	-0.00 (0.06)
ML1					0.26*** (0.06)
ML2					0.24*** (0.06)
ML3					0.37*** (0.06)
R <sup>2</sup>	0.01	0.00	0.02	0.01	0.31
Adj. R <sup>2</sup>	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 \*\*\* $p < 0.001$ , \*\* $p < 0.01$ , \* $p < 0.05$

# Variable „Wertschätzung gesunder Ernährung“: Zusammenhang mit DGE-Kriterium

**Tabelle:** Zusammenhang zwischen DGE-Kriterium und Variable 'Wertschätzung gesunder Ernährung'

	(1)	(2)	(3)	(4)	(5)
(Intercept)	-0.03 (0.07)	0.26 (0.18)	0.02 (0.06)	0.37* (0.17)	0.05 (0.07)
DGECriteriaNoScaled	0.27*** (0.07)	-0.02 (0.15)	0.25*** (0.06)	0.01 (0.13)	0.03 (0.06)
ML1					0.03 (0.07)
ML2					0.47*** (0.05)
ML3					0.24*** (0.05)
R <sup>2</sup>	0.06	0.00	0.06	0.00	0.37
Adj. R <sup>2</sup>	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 \*\*\* $p < 0.001$ , \*\* $p < 0.01$ , \* $p < 0.05$ .

# Zusammenfassung der Ergebnisse

- Einfaches lineares Modell, geschätzt mit OLS: Großer, teilweise signifikanter Zusammenhang zwischen DGE-Kriterium und gesundheitsrelevanten Variablen
- Hinzunahme von extrahierten Faktoren als Kontrollen:
  - Variable „seltener krank“: nach wie vor großer, statistisch signifikanter Koeffizient
  - Variablen „erweitertes Ernährungswissen“ und „Wertschätzung gesunder Ernährung“: kein Zusammenhang mehr erkennbar

# Fragestellung

- Welchen Effekt besitzt die Förderung einer sozialen Einrichtung durch den CHILDREN Entdeckerfonds auf die teilnehmenden Kinder und Jugendlichen?
- Herausforderung: Identifizieren einer geeigneten empirischen Methode, um die Wirkungseffekte des CHILDREN Entdeckerfonds zu bestimmen
- Hypothese: Die Teilnahme einer sozialen Einrichtung am CHILDREN Entdeckerfonds besitzt einen positiven Effekt auf die Alltagskompetenzen und das Selbstwertgefühl der Kinder und Jugendlichen

# Hintergrund

- Alle geförderten Einrichtungen erhalten finanzielle Mittel für die Bereitstellung des CHILDREN Mittagstischs
  - Aber: Nicht jede soziale Einrichtung nimmt am CHILDREN Entdeckerfonds teil, um den Kindern und Jugendlichen Ausflüge und Aktivitäten anzubieten
- ⇒ Der Unterschied zwischen den Einrichtungen hinsichtlich der Teilnahme am CHILDREN Entdeckerfonds wird dazu verwendet, um die Wirkung des Programms zu messen

# Einteilung in Treatment- und Kontrollgruppe

- Treatmentgruppe: Alle Einrichtungen, die sowohl am CHILDREN Mittagstisch als auch am CHILDREN Entdeckerfonds teilnehmen
- Kontrollgruppe: Alle Einrichtungen, die nicht am CHILDREN Entdeckerfonds teilnehmen, sondern nur durch den CHILDREN Mittagstisch gefördert werden
- Um die Einrichtungen in Treatment- und Kontrollgruppe einzuteilen, wurde analysiert, ob bei den Survey-Fragen zum Entdeckerfonds in einem bestimmten Jahr Angaben gemacht wurden

# Treatment-Variable

- Um die Einrichtungen in Treatment- und Kontrollgruppe einzuteilen, wird eine Dummy-Variable konstruiert
- $TreatEF_{it} = 1$ , wenn Einrichtung  $i$  im Jahr  $t$  am Entdeckerfonds teilgenommen hat und sich somit in der Treatmentgruppe befindet
- $TreatEF_{it} = 0$ , wenn Einrichtung  $i$  im Jahr  $t$  nicht am Entdeckerfonds teilgenommen hat und sich somit in der Kontrollgruppe befindet
- Die Kontrollgruppe ist wesentlich kleiner als die Treatmentgruppe

# Variante 1

- „Einmal Treatment, immer Treatment“
  - Sobald eine Einrichtung am Entdeckerfonds teilgenommen hat, gilt  $TreatEF_{it} = 1$  für das Jahr der ersten Förderung durch den Entdeckerfonds und alle darauffolgenden Jahre
- ⇒ Kein Wechsel von der Treatmentgruppe in die Kontrollgruppe möglich
- Solange eine Einrichtung keine Förderung vom CHILDREN Entdeckerfonds erhält, befindet sich diese in der Kontrollgruppe



# Variante 2

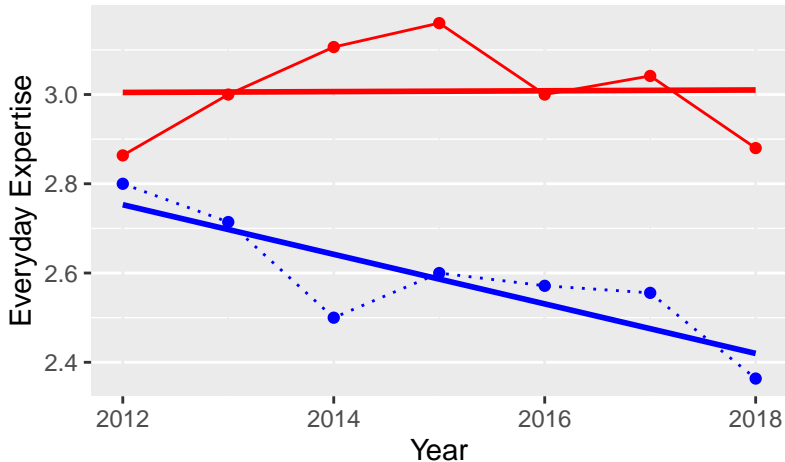
- Zeit-flexibler Treatment-Dummy
  - Eine Einrichtung befindet sich im Jahr  $t$  nur dann in der Treatmentgruppe, wenn diese tatsächlich Fördergelder vom CHILDREN Entdeckerfonds erhalten hat
- ⇒ Wechsel von der Treatmentgruppe in die Kontrollgruppe möglich

# Zielvariable

- Problem: Keine Variablen zum Entdeckerfonds für Einrichtungen, die nicht am Entdeckerfonds teilgenommen haben (= Kontrollgruppe)
  - Verwendete Zielvariablen vom Mittagstisch: Alltagskompetenzen und Selbstwertgefühl
- ⇒ Anwendbar auf den CHILDREN Mittagstisch und den CHILDREN Entdeckerfonds
- ⇒ Über den gesamten Beobachtungszeitraum verfügbar
- ⇒ Die Alltagskompetenzen und das Selbstwertgefühl der Kinder und Jugendlichen könnten dadurch beeinflusst werden, dass eine Einrichtung am Entdeckerfonds teilnimmt

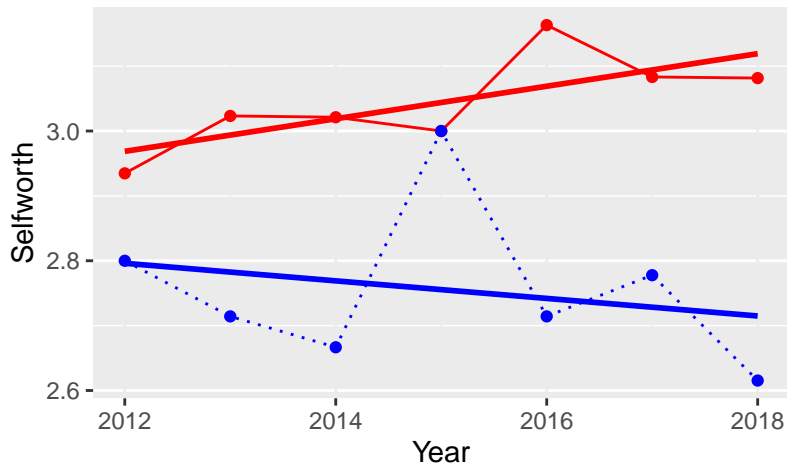
# Graphische Darstellung: Alltagskompetenzen

## Trend Of Everyday Expertise



# Graphische Darstellung: Selbstwertgefühl

## Trend Of Selfworth



# DID - Schätzung

- Empirische Methode: Differences-in-Differences (DID)
- Der DID-Schätzer misst den Effekt des Entdeckerfonds, indem die Veränderung der abhängigen Variable über die Zeit in der Treatmentgruppe mit der Veränderung in der Kontrollgruppe verglichen wird
- Regressionsgleichung:

$$Y_{it} = \alpha + \beta \cdot \text{TreatEF}_{it} + \gamma_i + \delta_t + \epsilon_{it} \quad (2)$$

- $\gamma_i$  = Einrichtungs-Fixed Effects,  $\delta_t$  = Jahr-Fixed Effects
- Der Regressionskoeffizient  $\beta$  entspricht dem DID-Schätzer

# Annahmen und Probleme

Zentrale Annahme des DID-Ansatzes:

- Common Trend Assumption: Ohne den Entdeckerfonds würden sich die Zielvariablen in der Treatment- und Kontrollgruppe mit dem gleichen Trend entwickeln

Potentielle Probleme:

- Verletzung der Common Trend Assumption
  - Selection bias / Endogenität: Nicht zufällig, welche Einrichtungen am Entdeckerfonds teilnehmen
- ⇒ Implementierung von Kontrollvariablen, die sich auf die Eigenschaften der geförderten Einrichtungen beziehen

# Alltagskompetenzen

	<i>Abhängige Variable:</i>			
	Alltagskompetenzen			
	(1)	(2)	(3)	(4)
treatEF	-0.143 (0.402)	-0.166 (0.405)	0.247 (0.299)	0.255 (0.310)
subsidy		0.019 (0.014)		0.016 (0.014)
totalCost		0.001** (0.000)		0.001* (0.000)
weeklyCooks		0.166** (0.072)		0.162** (0.073)
ID fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Number of observations	428	410	428	410
R <sup>2</sup>	0.475	0.490	0.476	0.491

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Tabelle: DID-Schätzung: Ergebnisse für Alltagskompetenzen

# Alltagskompetenzen

- Das Vorzeichen des Effekts hängt von der Definition der Treatment-Variable ab
- Hauptresultat: Die Teilnahme einer Einrichtung am Entdeckerfonds besitzt keinen statistisch signifikanten Effekt auf die Alltagskompetenzen der Kinder und Jugendlichen
- Aber: Die Anzahl der Beobachtungseinheiten in der Kontrollgruppe ist sehr gering
- Wenn der Stichprobenumfang steigt, dann könnte der Effekt des Entdeckerfonds gegebenenfalls positiv und statistisch signifikant werden



# Selbstwertgefühl

	<i>Abhängige Variable:</i>			
	Selbstwertgefühl			
	(1)	(2)	(3)	(4)
treatEF	-0.474 (0.309)	-0.481 (0.312)	-0.328 (0.247)	-0.442* (0.256)
subsidy		0.011 (0.018)		0.014 (0.017)
totalCost		0.000 (0.001)		0.000 (0.001)
weeklyCooks		0.036 (0.069)		0.037 (0.069)
ID fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Number of observations	428	410	428	410
R <sup>2</sup>	0.475	0.484	0.474	0.485

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

**Tabelle:** DID-Schätzung: Ergebnisse für Selbstwertgefühl

# Selbstwertgefühl

- Hauptresultat: Der Effekt des Entdeckerfonds auf das Selbstwertgefühl der Kinder und Jugendlichen ist negativ und teilweise statistisch signifikant

Mögliche Gründe:

- Die Anzahl der Beobachtungseinheiten in der Kontrollgruppe ist gering
- Die Fragebögen werden nicht direkt von den Kindern und Jugendlichen beantwortet, sondern von den Betreuern der geförderten Einrichtungen
- Die Skalierung der Zielvariable „Selbstwertgefühl“ führt zu geringer Variation

⇒ Daher sollte dieses Ergebnis nicht überinterpretiert werden

# Statistisches Lob für CHILDREN

- Erfassung vieler Variablen über die ganze Zeit

# Statistisches Lob für CHILDREN

- Erfassung vieler Variablen über die ganze Zeit
- Unveränderte Datenskala (alle bis keine)

# Statistisches Lob für CHILDREN

- Erfassung vieler Variablen über die ganze Zeit
- Unveränderte Datenskala (alle bis keine)
- Erfassung von zwei Variablen sowohl bei Mittagstisch als auch bei Entdeckerfonds

# Statistisches Lob für CHILDREN

- Erfassung vieler Variablen über die ganze Zeit
- Unveränderte Datenskala (alle bis keine)
- Erfassung von zwei Variablen sowohl bei Mittagstisch als auch bei Entdeckerfonds
- Fragebogen wird automatisch ausgefüllt  $\Rightarrow$  wenige fehlende Werte

# Empfehlungen für CHILDREN

- a) Datenstruktur und -erhebung:
  - Kinder und Jugendliche direkt befragen
  - Einheitliche Variablen, die jedes Jahr erfragt werden
  - Elektronische Fragebögen und Datenerhebung
  - Ordinale Variablen an sich kein Problem, Ersetzung durch metrische (in Prozent) nur sinnvoll, wenn dadurch nicht häufiger fehlende Werte
- b) Auswahl der Variablen: Ähnliche Variablen durch zugrundeliegende ersetzen, anhand von Faktoranalyse oder Partition
- c) Interventionen: Noch stärkeres Gewicht auf gesunder Ernährung in Einrichtungen

# References I

Millstein, J., Battaglin, F., Barrett, M., Cao, S., Zhang, W., Stintzing, S., Heinemann, V. & Lenz, H.-J. (2020). Partition: a surjective mapping approach for dimensionality reduction [Journal Article]. *Bioinformatics (Oxford, England)*, 36(3), 31504178, 676–681. <https://doi.org/10.1093/bioinformatics/btz661>



Danke für die Zusammenarbeit!