Guide to Dataset: Colombia ECD Project

Contents

Project context

This project covered 96 small towns in central Colombia, representing a large number of small communities across a relatively big geographical area. It exploited structures in place from the government's Conditional Cash Transfer Programme, Familias en Accion (FeA), which targets the poorest 20% of households in the country.

There are currently two waves of data, a baseline, pre-intervention wave collected between February and June 2010, and a follow-up wave 18 months later between September and December 2011, at the end of the intervention period.

The beneficiaries of FeA periodically elect a female representative, called the Madre Lider (ML). We randomly selected three from each town (municipality), and then from the families represented by the ML we randomly selected 5 children aged 12 to 24 months to be eligible for the intervention. Within each municipality, eligible households were randomly allocated (at the municipality level) to each of the following treatment arms:

1. **Control**
2. **Stimulation + Supplementation**
3. **Stimulation**
4. **Supplementation**

These MLs were then asked to take on the role of home visitor to implement the intervention of the different treatment arms. However, a significant proportion (nearly 40%) of the MLs either opted out or were not deemed suitable for the role, in which case they were replaced following their recommendations. In order to maintain the randomisation of the design, the families attached to the original ML were maintained, even if in the end she was not the actual home visitor.

The interventions

The stimulation intervention consisted of weekly visits to the homes of the target children, each visit lasting around one hour. The home visitors received a three week training programme in activities designed to stimulate children at different ages. They also received a weekly curriculum as a guide, and a set of locally produced materials (homemade toys from recycling material, picture books, puzzles, etc.).

The supplementation arm consisted of providing daily sachets of multiple micronutrient powder to mothers, via the home visitors, to add to the target child's food. Sachets were designed to provide iron (12.5mg), zinc (5mg), Vitamin A (300 μg retinol equivalent), Vitamin C (30mg) and folic acid (160 μg) for the children targeted.

The dataset

Sources of data

There are four main sources of data:

- Household survey
- Child Assessment -- Bayley test
- Intervention Data
- Community Survey

Household Survey

This collected information on the household and on household members in a set of modules (questionnaires) administered by interviewers. The mapping from the questionnaires to the raw datasets is shown in tables 2 and 3. Note that some additional questions were asked in the follow-up round, so it is worth checking the wording of questions in the original questionnaires.

Separate questionnaires were asked at household and individual levels. At the individual level there are several standalone modules for the mother, children under six years old, the target child of the intervention, the main care giver, and the home visitor.

The modules contain a rich variety of data: the home environment and composition of the household; education and health; employment and time use etc. There are also a series of modules on child development, such as the MacArthur Communicative Development Inventories (for vocabulary), the Bates Infant Characteristics Questionnaire (for temperament), and the Rothbart Infant Behaviour Questionnaires.

Table 1: Identifiers and sample sizes

	Identifier at Baseline	Number at Baseline	Identifier at Follow-up	Number at Follow-up
Cluster	dane	96	dane	96
Households	llave	1,429	nim	1,331
Target children	llaveper	1,429	llaveper	1,332
Biological Mothers	llaveper	1,395	llaveper	1,259
All household members	llaveper	7,499	llaveper	7,698
Children more than 6 months and less than 6 years	llaveper	2,147	llaveper	1,829
All children less than 6 years	llaveper	2,210	llaveper	1,889
All persons over age 5	llaveper	5,288	llaveper	4,665
Primary caregiver	llaveper	151	llaveper	122
Home Visitor	hv_0_id_bl	340	hv_0_id_fu	326

Table 2: Mapping of datasets to questionnaires at Baseline

Dataset	Questionnaire	Questions	Questionnaire Name	Sections	Unique identifer	N
Mapping of questionnaires - Baseline
modulo1a:	modulo 1	1 – 211, 321 - 1103	Housing and Home	Location and identification, Housing Characteristics, Distances, Property and Ownership of Property, Home savings and debts, Expenses, Other revenues and transfers, Incidents in the past 12 months, Participation in Familias en Accion	llave	1,429
modulo1b	modulo 1	301 - 320	Housing and Home	Household structure	llaveper	7,499
modulo1c	modulo1	1002 – 1003	Housing and Home	Food consumption of children in the household (for children between 6 months and 6 years)	llaveper	2,147
modulo2:	modulo 2	1 - 315	Education, work and health (for persons over age 5)	Education, Health, Work	llaveper	5,288
modulo3:	modulo 3	200 – 215, 310 - 507	Children less than 6 years	Institutional and non-institutional care, Nutritional status – growth and development, Home characteristics, Morbidity	llaveper	2,210
modulo3a:	modulo 3	216 - 218	Children less than 6 years	Historical care arrangements	llaveper + linea_2[a]	6,210
modulo4:	modulo 3 (annex)	101 – 208, 311 - 508	Annex for module 3 – all questions refer to target child	Interaction with biological parents (if don't live at home), Bates temperament scale, Quality of home environment	llaveper	1,429
modulo4a	modulo 3 (annex)	301	Annex for module 3 – all questions refer to target child	Time use	llaveper + time interval (ma_301a1 – ma_301a4)[a]	3,313
modulo4b	modulo 3 (annex)	302-310	Annex for module 3 – all questions refer to target child	Time use	llaveper +orden_302[a]	2,038
modulo4c	modulo 3 (annex)	VI. A	Annex for module 3 – all questions refer to target child	MacArthur_Bates results (12-18 months)	llaveper	798

Dataset	Questionnaire	Questions	Questionnaire Name	Sections	Unique identifer	N
Mapping of questionnaires – Baseline (continued)
modulo4d	modulo 3 (annex)	VI. B	Annex for module 3 – all questions refer to target child	MacArthur_Bates results (19-24 months)	llaveper	631
modulo5a	modulo 4	101 – 306, 601 – 902	Biological Mother	Time use, Empowerment, Family information, Contraception and reproductive history, Risk aversion, Inequality aversion, Knowledge of child development, Depression	llaveper	1,395
modulo5b	modulo 4	402-406	Biological Mother	Birth and breastfeeding of biological children	llaveper + llave_hijo[a]	3,237
modulo5c	modulo 4	502 – 504	Biological Mother	Physical development of biological children	llaveper + llave_hijo[a]	3,237
modulo6:	modulo 5	101 – 702	Primary Caregiver	ID and location, Education, Health, Work, Time use, Knowledge of child development, Depression	llaveper	151
modulo7:	modulo 3 (annex)	701 – 708	Annex for module 3 – all questions refer to target child	Anthropometrics (height, weight etc.) and haemoglobin	llaveper	1,429
modulo8:	modulo 4	100 – 1008	Biological Mother	Anthropometrics (height, weight etc.) and haemoglobin	llaveper	1,395
modulo10:	madre lider / home visitor	101 – 601	Madre Lider	ID, Household info, Education, Participation in Familias en Accion, Work, Knowledge of child development	hv_0_id_bl[b]	340

[a]In these datasets there exists more than one line of data per individual (e.g. representing different periods of time, different biological children etc) and so data entries are uniquely identified using a combination of identifiers. [b]Merge with data on target children using hv_0_id_bl id in modulo4 – N.B hv_0_id_bl is recorded as missing for all children in the control group.

Table 3: Mapping of questionnaires to datasets at Follow-up

Dataset	Questionnaire	Questions	Questionnaire Name	Sections	Unique identifier	N
Mapping of questionnaires – Follow-up						
modulo1a:	modulo 1	101 – 211, 401 – 403, 701 – 710, 901 - 1103	Housing and Home	Location and identification, Housing Characteristics, Distances, Expenses, Other revenues and transfers, Incidents in the past 18 months, Participation in Familias en Accion	nim	1,331
modulo1b	modulo 1	301 - 320	Housing and Home	Household structure	llaveper	7,698
modulo1c	modulo1	1002 - 1003	Housing and Home	Food consumption of children in the household (for children between 6 months and 6 years)	llaveper	1,829
modulo2:	modulo 2	101 - 313	Education, work and health (for persons over age 5)	Education, Health, Work	llaveper	4,665
modulo3:	modulo 3	101 – 215, 219 – 507	Children less than 6 years	Institutional and non-institutional care, Nutritional status – growth and development, Home, Morbidity	llaveper	1,889
modulo3a:	modulo 3	216 - 218	Children less than 6 years	Historical care arrangements	llaveper + orden	6,529
modulo4:	modulo 3 (annex)	101 – 208, 311 – 313, 401 – 417, 501 - 508	Annex for module 3 – Target Child	Interaction with biological parents (if don't live at home), Bates temperament scale, Quality of home environment	llaveper	1,332
modulo4a	modulo 3 (annex)	301	Annex for module 3 – Target Child	Time use	llaveper + time interval (ma_301a1 – ma_301a4)	3,217
modulo4b	modulo 3 (annex)	302 – 310	Annex for module 3 – Target Child	Time use		1,845
modulo4c	modulo 3 (annex)	601 – 610	Annex for module 3 – Target Child	MacArthur Bates results	llaveper	1,332

Mapping of questionnaires – Follow-up (continued)						
Dataset	Questionnaire	Questions	Questionnaire Name	Sections	Unique identifier	N
modulo4f	modulo 3 (annex)	801 – 832	Annex for module 3 – Target Child	Rothbart Infant behaviour questionnaire	llaveper	1,332
modulo5a	modulo 4	101 – 309, 801, 902	Biological Mother	Time use, Empowerment, Family information, Contraception and reproductive history, Risk aversion, Inequality aversion, Knowledge of child development, Depression	llaveper	1,259
modulo5b	modulo 4	401 – 406	Biological Mother	Birth and breastfeeding of biological children	llaveper + m4_iv leaves one duplicate	1,451
modulo5c	modulo 4	502 – 504	Biological Mother	Physical development of biological children	llaveper + m4_iv leaves one duplicate	1,451
modulo5d	modulo 4	XI	Biological Mother	Matching words to pictures	llaveper	1,259
modulo6:	modulo 5	100 - 702	Primary Caregiver	ID and location, Education, Health, Work, Time use, Knowledge of child development, Depression	llaveper	122
modulo6a:	modulo 5	IX	Primary Caregiver	Matching words to pictures	llaveper	122
modulo7:	modulo 3 (annex)	706 - 708	Annex for module 3 – Target Child	Haemoglobin	llaveper	1,332
modulo8:	modulo 4	1001-1008	Biological Mother	Haemoglobin	llaveper	1,259
modulo10:	madre lider / home visitor	101 – 601	Madre Lider	ID, Household info, Education, Participation in Familias en Accion, Work, Knowledge of child development	hv_0_id_fu	326
Modulo10a	madre lider / home visitor	VIII	Madre Lider	Matching words to pictures	hv_0_id_fu	326

[a]In these datasets there exists more than one line of data per individual (e.g. representing different periods of time, different biological children etc) and so data entries are uniquely identified using a combination of identifiers. [b]Merge with data on target children using hv_0_id_fu id in modulo4 – N.B hv_0_id_fu is recorded as missing for all children in the control group.

Assessment of outcomes of the target child

In additional to the information collected in the questionnaires administered by interviewers at the household level, the target children also attended a centre where her developmental level was assessed by a trained psychologist using the "Bayley Scales of Infant Development". The variables in this dataset (the child's cognitive, language and motor development) can be matched to the child's characteristics in the household questionnaire using the unique individual identifier "llaveper". The key variables of interest are the aggregated outcomes for cognition, receptive language, expressive language, fine motor skills and gross motor skills (b_tot_cog b_tot_lr b_tot_le b_tot_mf b_tot_mg).

Intervention data

This contains administrative information related to the manner in which the intervention was carried out. The most important is the "group" variable which identifies in which treatment arm each household was placed. Since this was randomized at the municipality level, the appropriate level for merging is using the municipality identifier "dane" in this dataset. There are also identifiers of the Bayley assessor (eva*) and the household interviewer (ent*).

Community Survey

This collects the information at the municipality level included in the questionnaire "community.pdf". The data are included in the datasets "a334_community_a" and "a334_community_b" and the key identifying variable for merging to other datasets is "dane".

Key to Outcome Variables

Table 4: Main Outcome Variables

Outcome	Measuring	Method of collection	Baseline File name	Follow-up File name
Target Child				
Bayley	Cognitive language and motor development	Assessment by trained psychologist (centre based)	Bayley_baseline	Bayley_followup
MacArthur-Bates	Communicative development	Household interview	modulo4c, modulo4d	modulo4c
Wolke Behaviour Rating	Emotional tone, activity, co-cooperativeness, approach	Assessment by trained psychologist	Bayley_baseline	Bayley_followup
Time Use	Time use of target child	Household interview	modulo4a, modulo4b	modulo4a, modulo4b
Anthropometrics	Height, Weight	Household interview (baseline), Bayley tester (follow-up)	modulo7	Anthro_followup
Haemoglobin	Haemoglobin concentration (g/dL)	Test (HemoCue) by household interviewer	modulo7	modulo7

Outcome	Measuring	Method of collection	Baseline File name	Follow-up File name
Target Child (continued)				
Rothbart infant behaviour	Observed (by mother) behaviours in past 15 days	Household interview	n/a	modulo4f
Food consumption	Food consumption in past 7 days	Household interview	modulo1c	modulo1c
Morbidity	Disability, Morbidity in past 15 days	Household interview	modulo3	modulo3
Bates' Infant Characteristics Questionnaire	Scenario based questions measuring temperament	Household interview	modulo4	modulo4
Mothers				
Maternal Depression	Measures of symptoms of depression	Household interview	modulo5a	modulo5a
Knowledge of Child Development	Series of questions testing knowledge of child development	Household interview	modulo5a	modulo5a
Time use	Time use in the last working day	Household interview	modulo5a	modulo5a
Primary Caregiver				
Knowledge of Child Development	Series of questions testing knowledge of child development	Household interview	modulo6	modulo6
Time use	Time use during the last working day	Household interview	modulo6	modulo6

Additional sources of information

For more detailed information on the intervention and the data, please refer to ongoing work which will be updated on Evaluation of Development Policy department's website, at the Institute for Fiscal Studies: http://www.ifs.org.uk/centres/EDePo/about

If there are any further questions regarding use of the data, please contact the researchers involved in this project at the IFS:

- Emla Fitzsimons: emla_f@ifs.org.uk
- Marta Rubio-Codina: marta_r@ifs.org.uk
- Alison Andrew: alison_a@ifs.org.uk
- Orazio Attanasio: o.attanasio@ucl.ac.uk