Quantitative Impact Assessment of Child Grants in Papua and Aceh Provinces

Baseline Report

Virginia Barberis, Martina Garcia, Bilal Hakeem, Paul Jasper, and Revita Wahyudi

Final Version – July 2019

Oxford Policy Management

About Oxford Policy Management

Oxford Policy Management Limited (OPML) is committed to helping low- and middle-income countries achieve growth and reduce poverty and disadvantage through public policy reform.

We seek to bring about lasting positive change using analytical and practical policy expertise. Through our global network of offices, we work in partnership with national decision makers to research, design, implement, and evaluate impactful public policy.

We work in all areas of social and economic policy and governance, including health, finance, education, climate change, and public sector management. We draw on our local and international sector experts to provide the very best evidence-based support.

Oxford Policy Management Limited
Registered in England: 3122495

Level 3, Clarendon House
52 Cornmarket Street
Oxford, OX1 3HJ
United Kingdom

Tel: +44 (0) 1865 207 300
Fax: +44 (0) 1865 207 301
Email: admin@opml.co.uk
Website: www.opml.co.uk
Twitter: @OPMglobal
Facebook: @OPMglobal
YouTube: @OPMglobal
LinkedIn: @OPMglobal

Acknowledgements

This report was produced by Paul Jasper (team leader), Virginia Barberis (data analyst), Martina Garcia (data analyst), Bilal Hakeem (survey expert), and Revita Wahyudi (project manager). Many organisations and individuals were involved in activities that led to the production of this report. In particular, we would like to thank:

- Our research partner Myriad, all of its staff members, and our fieldwork teams in Papua, without whom the implementation of the baseline survey in Papua would not have been possible.
- Badan Pusat Statistik (BPS) staff for their insightful feedback on previous drafts of the report and their help in gaining access to data required for this analysis.
- Empatika for generous inputs in discussions of preliminary findings with OPML and other stakeholders to this evaluation.
- MAKHOTA and KOMPAK for their inputs in validation workshops in Jakarta in May 2019.
- BANGGA Papua staff and all participants in the validation workshop for the Papua findings in Jayapura in May 2019.
- Participants in the validation workshop for the Aceh findings in June 2019.
- United Nations Children's Fund (UNICEF) colleagues for all their support and technical inputs at various stages of this evaluation.

Executive summary

Introduction

UNICEF has contracted OPML to undertake an impact assessment of the Child Grant (CG) programmes being implemented in the Indonesian provinces of Aceh and Papua between December 2018 and December 2020. This impact assessment has two objectives:

1. To identify the causal effects that the CG programmes will have on the lives of beneficiary households and children, as measured by a set of key outcome indicators. This primary objective will be completed once endline data is available after December 2020.

2. To describe the situations of CG beneficiary households and children at baseline and endline and to compare these to a group of non-beneficiaries.

This report presents the results from baseline analyses undertaken as part of Objective 2.

Methods

In Papua, this impact assessment will be undertaken using primary data collected via household surveys. The surveys target households that fulfil eligibility criteria for the CG throughout the period of this evaluation. This means that, at baseline, households needed to include at least one child under the age of two with an indigenous Papuan parent or caregiver. The baseline survey was carried out prior to any grant payments. At the end of 2018, approximately 1,400 households were surveyed across three treatment districts (Asmat, Lanny Jaya, and Paniai) and three comparison districts (Boven Digoel, Keerom, and Waropen). At endline, the aim is to collect data from the same households and children, when these will be between two and four years old. Econometric modelling will then be used to compare households in treatment and comparison areas and to estimate the effects of the CG.[1]

In Aceh, the impact assessment will be undertaken using secondary data collected by BPS via its national socioeconomic household (SUSENAS) and health surveys (RISKESDAS). These surveys cover the general population across all districts in Aceh. For the baseline analysis, this evaluation uses SUSENAS data from March 2018. To date, RISKESDAS data has not been made available. At endline, the aim is to include data from a second round of surveys in order to compare eligible households, with children under the age of seven, in treatment and comparison areas over time.[2] Currently, only one treatment district has been identified in Aceh: Sabang district. Other districts are likely to start disbursing grants between the baseline and endline phases of this evaluation.

[1] This will be a Propensity Score Matching analysis; see the main body of this report for details.
[2] This will be a Difference-in-Differences analysis; see the main body of this report for details.

For this report, simple descriptive analyses of the baseline data from Papua and Aceh are presented separately. In both cases, households and children eligible for the CG in treatment areas are compared to similar households and children in comparison areas within the same province.

Impact assessment in Papua

BANGGA Papua

The CG in Papua – called BANGGA Papua – started disbursements in three treatment districts: Asmat, Lanny Jaya, and Paniai. Papuan children under the age of four are eligible for the grant of IDR 200,000 per month, which is given to the mother or recognised caregiver. These payments will lead to an increase in the income of beneficiary households that, in turn, is expected to lead to an increase in household consumption and expenditure. This increase is then assumed to lead to positive changes along a variety of different dimensions: better diets, higher educational attainment of the children, better childcare, increased birth registration, and healthier children. The CG payments are also expected to boost the local economy. In the baseline survey, indicators that relate to most of the outcomes mentioned above, in addition to the background characteristics of households and children, were measured.

Key baseline results

The focus of this summary is on a set of baseline comparisons where indicators vary significantly between the treatment and control groups in Papua.

First, many more households in treatment areas (over 40%) were recipients of some form of financial support at baseline than in the comparison areas (10%). This implies that households in the treatment group are likely to already be a recipient of support and therefore to receive the BANGGA Papua CG *on top of* existing payments.

Second, average household consumption is slightly, but significantly, lower in treatment areas compared to control areas. Despite this, the proportion of the population estimated to lie below the Papuan poverty line does not vary by treatment status. Overall, 17% of the surveyed population in the treatment and control areas falls below the Papuan poverty line, as set by BPS in September 2018.

Third, estimates along a variety of other dimensions indicate that households and children in the treatment areas are worse off than households in the control areas. For example, school attendance among children (5–17 years old) in surveyed households is 32% in the treatment areas, compared to 60% in the control areas. The proportion of children (0–17 years old) with birth certificates is significantly lower in households in the treatment group (15%) than in the control group (20%). Surveyed children in the control group are more likely to have access to playthings (33%) compared to those in the treatment areas (26%). Children in treatment areas are also more likely to have been left alone for more than one hour at least once in the week prior to the survey (90%) than children in comparison areas (62%). Finally, the proportion of children with

minimum dietary diversity is lower in the treatment group (20%) than in the comparison group (27%).

Fourth, while child malnutrition prevalence is worryingly high overall, more surveyed children in treatment areas are stunted (28%) than in the comparison areas (15%). Child wasting is similar across the two groups, with an estimated overall prevalence of 20%, which is very high. In Asmat and Waropen, more than a quarter of all children surveyed are estimated to be wasted, which is extremely high.

From an evaluation design perspective, these differences are not surprising: the treatment areas were selected purposefully by BANGGA Papua because they were considered to be the poorest areas in the province. In order to be able to compare like with like at endline, and in order to be able to attribute changes over time to BANGGA Papua rather than these underlying differences across households, econometric modelling will be used to estimate programme impact. For instance, an illustrative analysis in this report shows that, using baseline data, differences across treatment and comparison areas do disappear once Propensity Score Matching (PSM) is employed.[3]

Methodological caveats and implications

There are other methodological caveats, however, that need to be emphasised here. As referenced above, the imbalanced implementation of financial support programmes in treatment areas means that – at endline – it will be important but difficult to differentiate between the effects of those programmes and that of BANGGA Papua. The evaluation team aims to address this by collecting comprehensive data on the other financial support programmes at the household level in order to control for them in analyses, and to implement sensitivity checks.

Additionally, the very high levels of wasting among children surveyed at baseline sadly imply that many children might pass away before the end of 2020. Thus, surveying these same children at endline might not be possible. The evaluation team suggests implementing a tracking exercise, prior to the endline survey, to assess the risk of this happening with the potential to adapt the evaluation approach.

Furthermore, a mixture of operational difficulties for BANGGA Papua and strong assumptions underlying the programme implementation might mean that this evaluation, contrary to expectations, will not be able to identify significant effects on many of the indicators measured in our surveys. Careful messaging of findings and managing expectations in dissemination events is therefore crucial for this evaluation.

Finally, discussions around consumption estimates derived from the baseline data have indicated the need to adapt the survey instrument for the next round of data collection. Stakeholders suggested trying to measure household wellbeing via additional indicators, such as a multidimensional poverty index or an asset index. The

[3] Please refer to the main body of this report for more details on the econometric approach that will form the basis of the impact analysis.

evaluation team will review the survey instrument at endline in response to these suggestions.

Impact assessment in Aceh

The CG programme in Aceh

In contrast to the CG programme in Papua, grants have yet to be disbursed in Aceh. One district in Aceh, Sabang, has allocated funding for the implementation of the CG and is planning to start payments at the end of July 2019. It should be noted that, at the beginning of this evaluation, more districts in Aceh were expected to be part of this programme earlier in the evaluation period.

The CG in Sabang will unconditionally benefit all children aged between zero and six years to the amount of IDR 150,000 per child per month. Payments will be made once a month into an individual bank account in the name of the child's main caregiver. As in Papua, the expectation is that these payments will increase household income and – in further steps – will positively affect household consumption and a variety of different dimensions of household and child wellbeing. The CG in Aceh will, however, be part of a larger integrated set of interventions, which is not the case in Papua. Therefore, any effects that this impact assessment might identify will possibly be, at least partly, due to the concurrent implementation of other interventions.

Key baseline results

The focus of this summary is on key differences between households and children in the treatment area, Sabang, and comparison areas, i.e. other districts in Aceh. The analysis looked at households that included children under the age of seven, as these were households fulfilling the eligibility criteria for the CG in Sabang.

First, households in Sabang were much more likely to already be the recipients of some form of financial support compared to the rest of Aceh. In particular, 72% of households in Sabang received support from the local government compared to 8% in the comparison districts.

Second, consumption levels (estimated in IDR) were similar in Sabang compared to other districts. However, the consumption composition was slightly different, with treatment households in Sabang consuming proportionally less food than households in the comparison districts.

Finally, when looking at a variety of other indicators, there were either no or only small differences between households eligible for the CG in Sabang compared to households in other districts in Aceh. For example, around 80% of all children aged 5–17 were attending school both in eligible households in Sabang and in other districts in Aceh. Similarly, between 60% and 70% of children in Sabang and the other districts had a valid birth certificate, a much higher estimate than in Papua. The one area where differences could be observed was childcare. For example, children were less likely to have been left alone in Sabang (6%) than in the comparison districts (15%).

Methodological caveats and implications

A key methodological caveat to this baseline report is that, unfortunately, no analysis of RISKESDAS data could be conducted. This means that this report does not include any analyses on health-, diet-, or nutrition-related indicators for children in Aceh. The evaluation team is in discussions with UNICEF in order to assess whether some analysis of this data can happen at the endline phase of this evaluation.

In addition, as noted above, the impact assessment in Aceh was designed under the assumption that several districts would be implementing the CG rather than a single district. As a result, the treatment sample for the analysis is currently limited to Sabang, which means a small sample size of only 120 households. If Sabang remains the only district to implement the CG, this will limit the evaluation's ability to be able to identify CG effects at endline.

Similar to Papua, implementation difficulties and strong assumptions underlying the CG programme logic mean that there is a possibility that – contrary to expectations – programme effects at endline might not be visible. One particular issue is that treatment households will be able to receive a maximum of nine months' worth of payments before March 2020, when the next round of SUSENAS will be implemented in Indonesia. The evaluation team is discussing with UNICEF whether using a different round of SUSENAS data (March 2021) would be preferable for this evaluation in Aceh, thus enabling a longer period of programme effects to materialise.

Finally, attributing changes between baseline and endline to the CG implementation in Aceh alone will be difficult. As noted earlier, this is due to the fact that the programme is being implemented concurrently with other interventions and, in addition, that many households in Sabang are already recipients of some form of local financial support. At endline, the evaluation team will attempt to assess how much bundling of different components actually took place and how long this local financial support continued. Findings will need to be interpreted in light of these parallel interventions.

Table of contents

List of tables, figures, and boxes

List of abbreviations

ATT	Average Treatment Effect on the Treated
BLSM	*Bantuan Langsung Sementara Masyarakat* (Direct Community Cash Transfer – unconditional cash transfer)
BPS	*Badan Pusat Statistik* (Indonesia Statistic Office)
CAPI	Computer Assisted Personal Interviewing
CG	Child Grant
DID	Difference-in-Differences
FANTA	Food and Nutrition Technical Assistance Project
GPS	Global Positioning System
HH	Household
IDR	Indonesian Rupiah
IYCF	Infant and Young Child Feeding
LAZ	Length-for-Age
MAM	Moderate Acute Malnutrition
MICS	Multiple Indicator Cluster Survey
OECD	Organisation for Economic Co-operation and Development
OPML	Oxford Policy Management Limited
PIP	*Program Indonesia Pintar* (Smart Indonesia Programme – education cash assistance programme)
PKH	*Program Keluarga Harapan* (Family Hope Programme – conditional cash transfer programme)
PODES	*Potensi Desa* (Village Potential Statistics)
PSM	Propensity Score Matching
PSU	Primary Sampling Unit
QA	Quality Assurance
RISKESDAS	*Riset Kesehatan Dasar* (Basic Health Research)
SAM	Severe Acute Malnutrition

SD	Standard Deviation
SUSENAS	*Survei Sosial Ekonomi Nasional* (National Socio-Economic Survey)
ToC	Theory of Change
UNICEF	United Nations Children's Fund
WAZ	Weight-for-Age Z-score
WHO	World Health Organization
WLZ	Weight-for-Length Z-score

1 Introduction

1.1 Background to the evaluation

UNICEF has contracted OPML to undertake an impact assessment of the CG programmes being implemented in the provinces of Aceh and Papua in Indonesia between December 2018 and December 2020. UNICEF defined the objectives of this project as follows:

> *UNICEF is seeking an agency/firm to design and implement a quantitative impact assessment of Child Grants in Aceh and Papua, including a baseline and endline assessment after 24 months.* (Terms of Reference, p.3)

The term 'quantitative impact assessment' implies that the primary objective of this project is the identification of the causal effects that the CG programmes will have on key outcomes of interest. This primary objective will only be achieved once endline data is available and can be analysed, i.e. after December 2020.

A secondary objective of this project is, however, the production of descriptive statistics on outcomes of interest, irrespective of their use for the impact assessment itself, both at baseline and endline in the areas in Papua and Aceh where the CG will be implemented. This report, which presents the results from descriptive baseline analyses conducted in the context of this evaluation, follows from this secondary objective.

1.2 Objectives and structure of this report

Resultantly, the key objective of this report is to present statistical analyses on indicators that describe the situation of households – and their members – eligible for CGs in specific areas in both Aceh and Papua.

Importantly, these analyses will focus on comparing eligible households in areas where the CG will be disbursed (treatment areas) to eligible households in areas where it will not (comparison areas or controls). These comparisons will show how – prior to the disbursement of any grants – the situation of eligible households compares to the situation of similar households in areas in which no CG will be disbursed.

In addition, this report will present analyses that compare indicators across different subgroups of households. For example, some outcome indicators will be disaggregated by gender. The objective of these additional comparisons is to provide more detail on the situation of households and their members in study areas, as these can shed more light on particularly vulnerable subgroups of households that might particularly benefit from the grants.

Finally, this report will present methodological details, results from robustness checks, and data quality analyses, which will be aimed at a more technical audience. The objective of these is to provide our client and any interested party with a comprehensive assessment of the robustness with which this baseline study was implemented. For ease of presentation, however, most of this technical information is presented in methodological annexes.

The remainder of this report is structured as follows: section 2 provides a brief description of how results are presented in this report. The rest of the report is divided into two large sections that separately present results for Papua (section 3) and Aceh (section 4). Within each of these sections, we first present our understanding of the respective CG programmes (sections 3.1 and 4.1), the specific areas of interest that the evaluation will focus on in each of the provinces (sections 3.2 and 4.2), and the methods used (sections 3.3 and 4.3). We then present results along key dimensions and outcomes of interest. We end each section with a discussion of province-specific methodological caveats – that relate to our impact assessment – and the implications of our findings for the next phase of this study (sections 3.10 and 4.8). We end the main body of the report with a general conclusion and a summary of next steps for this evaluation (section 5). Annex A then presents additional analyses and a detailed description of methods used with respect to the Papua baseline survey. Annex B presents full statistical tables for a selection of key indicators used in this report.

1.3 Audience for this report

The specific target audiences for this report are:

- Stakeholders of the CG implementation in Aceh and Papua (such as implementing organisations);
- UNICEF Indonesia; and
- The Government of Indonesia, including the Ministry of Planning (BAPPENAS), BPS, the Ministry of Finance, and local governments in Aceh and Papua.

2 How we present results in this report

Mostly, results in this report are presented in graphs that look similar to Figure 1. Wherever possible, we follow this format to present any findings. The graphs can be interpreted as follows:

- The title at the top of the graph describes the indicator that is being plotted in the graph. In the case of the example below, this is the age of children measured in months. The square bracket gives an indication for the group of individuals for which this indicator is defined – in this case all selected eligible children in our survey, for which there is one in each household, which in graphs is abbreviated as HH.

- Estimates are mainly disaggregated by treatment status. This means that the blue dot on the left-hand side shows the estimate for the control (or comparison) group, while the dot on the right-hand side shows the estimate for the treatment group.

- In some graphs, this disaggregation will change and we will show estimates for different subgroups. The labels at the bottom of the graph will always clearly indicate which groups are being compared in a graph.

  It is important to emphasise that we do not generally present results disaggregated by small administrative areas in this report. This is because the survey was designed and powered to produce estimates at the level of the treatment or comparison groups and not lower levels. Estimates at lower administrative levels have large sampling uncertainty associated with them, which means that presenting them is not generally advisable.

- Each of the dots – and the numbers to the right of them – show the estimate for each of the subgroups. In this case, the average age of children in months.

- The whiskers above and below the dots indicate the 95% confidence interval for these estimates. The 95% confidence interval represents the area within which the true value of the indicator – in this case the true average age for children in study areas – will lie with 95% probability. If – when comparing estimates across groups – these intervals overlap, we generally say that differences in estimates between the two groups are not statistically significant. If they do not overlap, differences are statistically significant.[4]

- The left-hand y-axis will indicate the scale on which indicators are being measured. In this case, these are months. In many other cases, the indicators are proportions, which are measured in percentages, and hence the scale will show ‘Proportion (%)’.

- We also present information on the numbers of observations on which estimates for each graph are based. First – each of the group labels will be

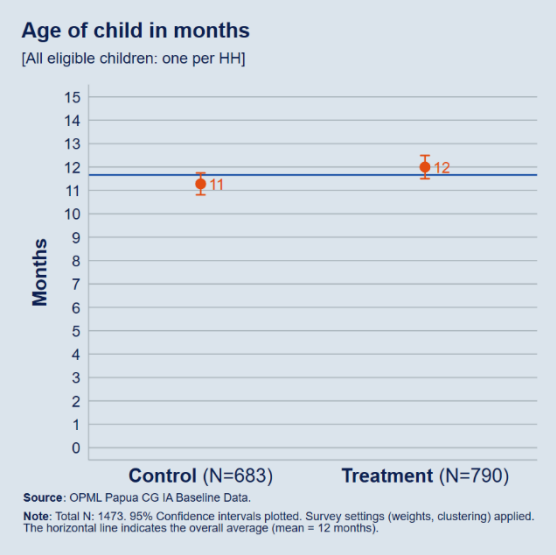
[4] Note that – strictly speaking – we base statements of statistical significance on t-tests. Results for these are presented in Annex B. For the purposes of accessibility of findings in this report, we limit ourselves to confidence intervals in the main parts of this report, using them as proxies to show whether differences in estimates are significant or not.

accompanied by statements of numbers of observations per group in parentheses behind the name of the group: 'Control (N=683)' below means that there are 683 observations in the Control group. The overall sample size for the graph is also presented in the note at the bottom of the graph as 'Total N'.

- The note at the bottom of the graph provides some additional background information, including on the source of the data used to produce this graph. In this case we are using data from our baseline survey in Papua: 'OPML Papua CG IA Baseline Data'. For the Aceh analysis, we will use national socioeconomic survey (SUSENAS) data.

- Finally, the blue horizontal line in this graph indicates the overall estimate – in this case the overall average age – which results from pooling all observations used for this graph (in this case pooling the comparison and treatment groups together). The value for this estimate is also presented in the note at the bottom of the graph.

Figure 1: Example graph – average age of children in months



In the following sections, we present results following a simple structure. For each of Papua and Aceh, we present background information on households and individuals in our sample first, then present information on whether households were receiving any cash support prior to the baseline, then focus on indicators that could be considered to be directly affected by the CG (first-order outcomes), and then move on to indicators that are more likely to be indirectly affected or very difficult to be affected (second- and third-order outcomes). In practice, this means that we present results as follows:

- First – background characteristics of households and children.

- Second – exposure to any other support programmes.

- Third – consumption indicators. Given that the CG is a cash grant, the assumption is that these will be indicators that will be directly affected by the grant.

- Fourth – indicators related to education access, birth registration, childcare, and child feeding.

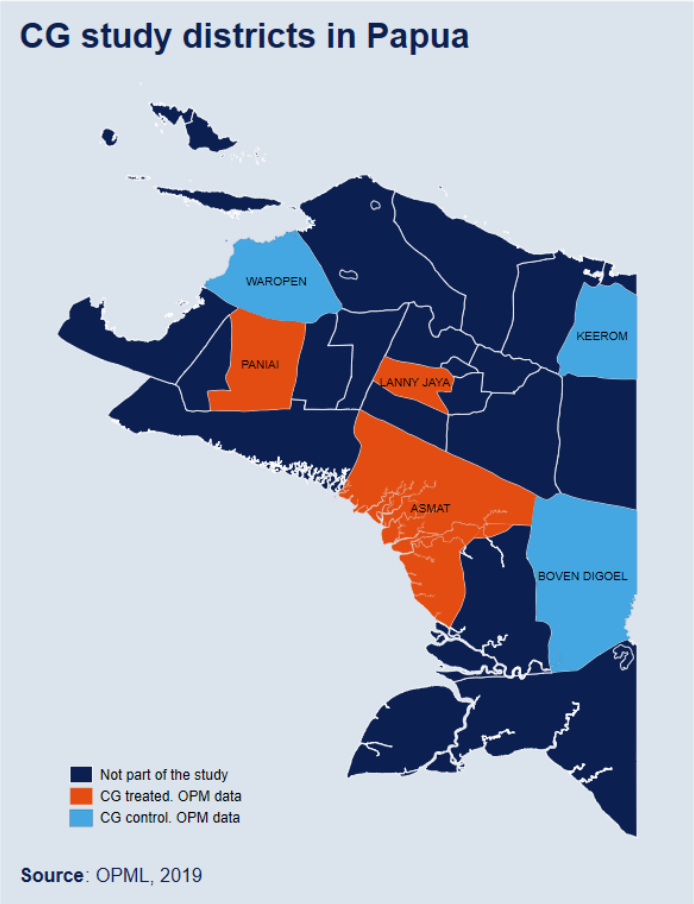
- Fifth – indicators related to child malnutrition (for the case of Papua only).

The number of exact indicators presented in the two main sections of this report (sections 3 and 4) depends – of course – on the data available to the research team, which we discuss in detail in the sections below.

3 Papua

3.1 A summary of our understanding of the intervention

Figure 2: Map of Papua Province



The CG programme in Papua – called BANGGA Papua – commenced operations in December 2018, with first payments being made to eligible households before Christmas 2018. The programme is being implemented via a partnership between the Papuan Provincial Government and participating districts. The programme started in the three districts of Lanny Jaya, Paniai, and Asmat (highlighted in orange in the map above). Pending funding, the programme is planned to progressively cover all remaining districts in the province in the future.

The CG unconditionally benefits all Papuan children (i.e. children with a Papuan parent or caregiver) under four. The benefit level is IDR 200,000 (around US$ 13.16)[5] per eligible child per month, which is given to the mother or recognised guardian of the eligible infant. This initial amount may be reviewed annually, however, in order to ensure continued funding to meet programme objectives. The following paragraphs

[5] This report uses XE exchange rates as at 19 October 2018.

describe our understanding of how the programme is supposed to be implemented in the province.

Payments of the CG will be handled through Bank of Papua, for which most recipients will be assisted in opening individual bank accounts at the institution. Bank of Papua is expected to collaboratively work with the local governments to organise payment points in order to facilitate account registration and cash withdrawal.

During a validation visit in Papua in May 2019, OPML received information that, at the time of writing, there were a total of six payment points in operation in Asmat, eight in Paniai, and one in Lanny Jaya. Eligible households have been registered and – via socialisation activities – have been made aware of the location and functioning of payment points.

BANGGA Papua aims to be appropriately coordinated with other programmes at the local level, and its beneficiaries will not be excluded from benefiting from other social protection schemes such as the conditional cash transfers for poor families, known as *Program Keluarga Harapan* (PKH), or *Rastra* (formerly known as *Raskin*), if they are deemed to be eligible for those as well.

It should be noted that BANGGA Papua has also informed OPML that it intends to implement some sensitisation and complementary activities around health and nutrition in targeted communities, with the objective of raising awareness round those issues in beneficiary households. This evaluation will not focus on those activities, although they will need to be taken into account when interpreting findings at the endline phase of the project.

A light-touch Theory of Change (ToC)

OPML has not been able to implement a comprehensive ToC workshop with the implementing organisations of BANGGA Papua. This section therefore presents a short, simplified summary of what the evaluation team considers to be the key causal pathways and assumptions underlying BANGGA Papua's ToC, which informs the analyses implemented as part of this evaluation. Figure 3 below is a graphical representation of this simplified ToC. The outcomes that this evaluation focuses on can be found within the red rectangle.

We consider outcomes on the left in the causal pathway to be first-order outcomes, whereas the outcomes further to the right, which are less directly affected by programme outputs, are second- or third-order outcomes.

First-order outcomes for the BANGGA Papua programme are related to an increase in household income, directly derived from receiving a cash transfer per eligible child. The programme expects beneficiary households to use the increase in income in ways that have immediate effects on household expenditure and consumption patterns related to education, food, health, and other additional services. Via this increased expenditure, the CG is expected to result in improvements in short- to medium-term second- and third-order outcomes, including improved nutritional, health, and educational outcomes for children. The programme is also expected to have longer-term effects on households' human capital, asset accumulation and livelihoods strategies, reducing

poverty, and increasing overall quality of life and welfare. There is also a potential role of cash transfers in affecting productivity and growth within local economies.

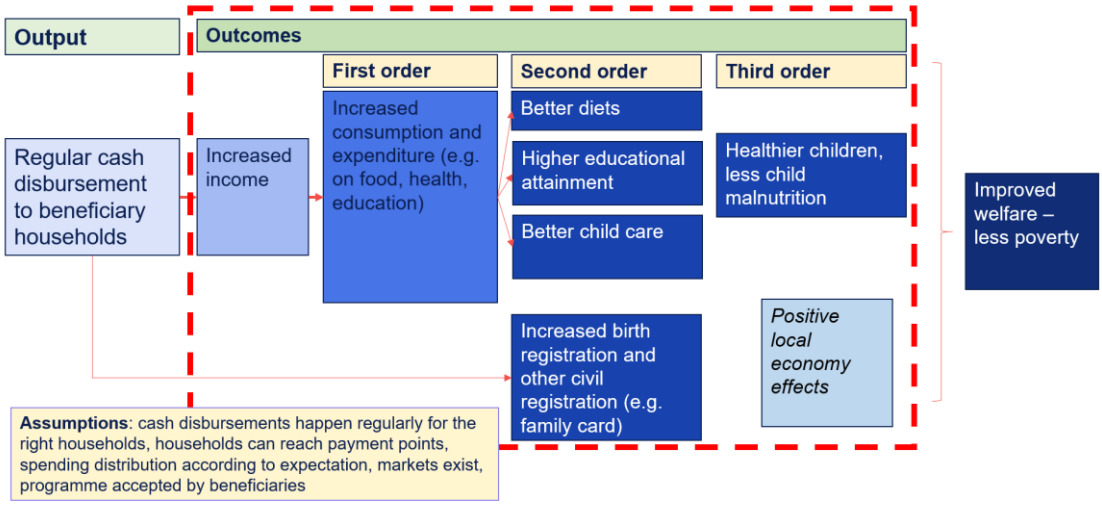
From our perspective, there are a few key areas that should be highlighted when assessing this ToC:

- First, Papua is a challenging environment to operate in, in particular when taking into account the remoteness and difficult accessibility of some communities. It is our understanding that BANGGA Papua intends to roll out the programme to all communities in treatment districts. However, from OPML's perspective, there is a risk that remote communities will also be hard to reach for BANGGA Papua payments, which means that dilution effects could potentially limit the effects that the programme could have in these districts. Note, for example, that at the time of writing there is only one payment point in Lanny Jaya, which means that communities in that district are potentially difficult to reach. In addition, during a validation workshop with BANGGA Papua in Papua in May 2019, we learned that that at that point in time only about 40% of all target children in Paniai had been registered.

- Second, in terms of causal links between outcomes, the ToC depicted below includes some very strong assumptions. For example, it is not entirely clear how causal links between first- and second-order outcomes (e.g. from increased income and expenditure to higher educational attainment) will materialise. Similarly, it is not fully clear how the CG will directly lead to a reduction in child malnutrition (e.g. stunting) in the relatively short period of time in which this evaluation is implemented, given the complex nature in which stunting is related to many different aspects of households' lives.[6] While BANGGA Papua CG payment will be accompanied with some complementary activities targeting health and nutrition issues, it is our understanding that these mainly amount to awareness raising and information campaigns. Tracking these events and their implementation will be relevant for this evaluation, but it is important to mention that **these strong assumptions mean that there is a real risk that within the time period of this impact evaluation no effects on some outcome indicators will be measured.**

[6] See, for example: www.who.int/nutrition/healthygrowthproj/en/index1.html.

Figure 3: Simplified version of the BANGGA Papua ToC



3.2 The focus of the evaluation in Papua

This impact assessment focuses on a subset of the outcome areas presented in Figure 3 above. The impact assessment methodology chosen – described in further detail in section 3.3 – implied that OPML has implemented a baseline survey in December 2018 and will implement an endline survey in 2020. This means that we will measure outcome indicators directly via two surveys at those two points in time. OPML collaborated with Myriad, an Indonesian research firm, for the baseline survey in 2018.

Table 1 below presents a summary of outcomes that are to be measured in Papua using the surveys implemented by OPML. For each of those indicators, the objective of the impact assessment will be to understand the extent to which changes between baseline and endline, in the treatment districts, can be attributed to the implementation of the CG.

It should be emphasised here that all indicators listed in the table below are to be measured in the households of children that we consider to be eligible for the CG at baseline and at endline. Because the objective is to measure effects of the CG on a population of children in treatment areas that are eligible for the grant at baseline and will be eligible throughout the evaluation, this means that the baseline study focused on Papuan children aged 0–23 months at baseline and the endline survey will focus on those same children who will be 24–47 months old at endline.

Table 1: Key outcome areas and examples of related indicators covered by the impact assessment in Papua

Outcome level	Outcome area	Selected key indicators (examples) – measured in survey target population[7]	Notes
First-order outcomes	Household consumption and expenditure	Average monthly household consumption per adult equivalent	Based on adapted consumption module from SUSENAS
Second-order outcomes	Access to education	Proportion of children attending school	
	Food consumption	Food share of monthly total consumption	Based on adapted consumption module from SUSENAS
	Child nutrition and child feeding	Exclusive breastfeeding, dietary diversity, minimum acceptable diet	Based on Food and Nutrition Technical Assistance Project (FANTA) indicators for assessing infant and young child feeding (IYCF) practices
	Birth registration	Percentage of children who have a birth certificate	
	Early childhood development and childcare	Proportion of children who have three or more children's books	Module adapted from UNICEF surveys
Third-order outcomes	Child malnutrition	Wasting and stunting prevalence	Measured using WHO standards

3.3 Baseline data collection and analysis methodology

Annex A.2 provides a detailed methodological description of the baseline survey implemented in Papua. Annex A.3 describes how the survey was implemented. This section provides a short summary of the key characteristics of this data collection exercise and the related analysis.

3.3.1 Baseline data collection

In the context of this evaluation, the baseline and endline surveys have three main objectives:

- First, to collect data from individual children and their households who are eligible to receive the CG in the treatment districts and are estimated to be

[7] Note that the target population is Papuan children aged 0–23 months at baseline and their households. These same children will be aged 24–47 months at endline.

eligible until the endline data collection. This implies collecting data from Papuan children aged 0–23 months, and their households, at baseline.

- Second, to collect data in a way in which estimates are representative for this population of children and their households in the areas in which the survey is implemented.
- Third, to collect data in comparison areas in a way that allows appropriate comparison to data in the treatment areas for the purposes of the impact assessment.

In order to achieve these objectives, OPML and Myriad implemented a baseline survey in November and December 2018 that covered eligible children aged 0–23 months and their households in the three treatment districts and three comparison districts (see Figure 2 above).

Sampling

The sampling for this survey was implemented in four steps:

- First, three appropriate comparison districts to the three treatment districts were selected using secondary data and matching methods: Keerom, Boven Digoel, and Waropen (see the light blue districts highlighted in Figure 2).[8]
- Second, sub-districts that could be included in the survey were then selected within each of the six districts. In each district, the objective was to select accessible and safe sub-districts that were spread across the district's area and to include the district capital in the survey. It is important to emphasise here that – given the security situation in Papua – the list of accessible sub-districts in some districts was limited.
- Third, OPML then implemented a random sampling approach among a selected list of accessible and safe villages in the selected sub-districts and districts. These were so-called primary sampling units (PSUs). The list of villages to sample from was derived from the publicly available Village Potential Statistics (*Potensi Desa/PODES*) 2011 dataset. Again – as with sub-districts – inaccessible villages (e.g. due to tribal conflicts) were removed from these lists as soon as such security information was available. In total, per district, the objective was to sample and visit a total of 20 PSUs.
- Fourth, within villages, a listing exercise was conducted to identify eligible children for the survey and their households. **As described above, eligibility for the purposes of our survey was defined as being 0–23 months old and having a Papuan caregiver or parent.** In a second step, a random 'main' sample of a maximum of 15 children and a sample of 10 'replacement' children were drawn for the purposes of data collection.

The table below summarises the composition of our planned sample in treatment and comparison areas, the maximum sample size that would have been possible, the target

[8] Note that this matching also had to take into account the fact that other districts were to be covered by the CG programme in 2019 and 2020, i.e. prior to endline. Those districts were not eligible to be part of this study.

sample size, and the actual sample size achieved at baseline and included in the analysis for this report.

The overall maximum sample size at baseline would have been 1,800 children and households. The actual sample size in our survey was planned to be lower than this, as we kept in mind the fact that reaching 15 children and households in each of the sampled villages would not have been possible. On average, i.e. across all villages, we expected that this survey would collect data from 12 children and their households per village. Please see Annex A.2 for more detail on how sampling was implemented for this survey, including a discussion of methodological caveats.

Table 2: Baseline sample size target and achieved

Unit	Treatment districts	Control districts
Districts	3	3
PSUs per district	20	20
Maximum households and children (aged 0–23 months) per PSU	15	15
Total maximum sample size of households and children	**900**	**900**
Target average household and children (aged 0–23 months) per PSU	12	12
Total expected sample size of households and children	**720**	**720**
PSUs included in the baseline analysis	59	56
Achieved average household and children (aged 0–23 months) per PSU included in baseline analysis	13.3	12.2
Total achieved sample size of households and children included in baseline analysis	**790**	**683**

Data collection

Data collection itself was implemented in Papua between 1 November and 11 December 2018. OPML worked together with Myriad to implement this survey. Per district, two teams of one supervisor and up to four enumerators were deployed to collect data. Each team also included an additional team member who was specifically trained on collecting anthropometric data from children. Annex A.3 provides a detailed description of the data collection implementation, the fieldwork model, how enumerators were trained, and the data quality assurance system set up during fieldwork.

Data collection instruments used in the baseline survey were designed so as to cover the outcome areas listed in Table 1. In addition, these instruments were used to collect important background characteristics of households, household members, and children living in the household. Annex A.4 provides a description of the instruments used in the baseline survey and the sources on which they were based. Most of the modules used in the questionnaires were based on modules used by BPS in Indonesia or followed international guidance, e.g. from the World Health Organization (WHO). In summary, two main instruments were used:

- **One household questionnaire** – to which the household head was the respondent.

- It included background information on household members such as, for example, age, education, marital status, and gender.
- It also included the consumption and expenditure module used to collect data on household consumption.
- It also included information on whether the household was benefiting from other programmes that were being implemented in the survey area.

- **One child and caregiver questionnaire** – targeted to capture information on the child sampled who is aged 0–23 months at baseline. The respondent was the main caregiver of the child.
  - This questionnaire included questions on child feeding and child health.
  - It also included the section on anthropometrics used to measure children's height and weight.
  - Finally, it also included the childcare and early childhood development module.

All data collection was implemented using Computer Assisted Personal Interviewing (CAPI), i.e. using tablets. The software used to collect data was Survey Solutions (http://support.mysurvey.solutions/). All questionnaires were translated into Bahasa. It should be noted, however, that Papua is characterised by a very high diversity of local indigenous languages, which means that individual enumerators had to rely on local translators to conduct interviews in some villages.

3.3.2 Data analysis

As described in section 2, the present baseline report focuses on a descriptive analysis of results from the baseline survey. This means that the results presented here do not require sophisticated modelling. Rather, we mainly present estimated averages and proportions of indicators, i.e. summary statistics. All analyses take into account that we are operating with survey data and hence have been weighted using survey weights, while standard errors have been clustered appropriately given the multi-stage sampling set-up.[9] All analyses were implemented using the statistical programming software Stata 15.

At endline, we will combine baseline and endline data to estimate the impact of the CG in Papua. To do that, we will employ PSM. We provide a short description of PSM and an initial analysis that refers to it in section 3.9.

The following sections present key results from our baseline survey. We start by presenting the background characteristics of households and children (section 3.4). The next section presents results relating to the exposure of households to cash support programmes (section 3.5). Section 3.6 then covers results with respect to the household level of consumption and consumption patterns. Section 3.7 focuses on access to education, birth registration, early childhood development and childcare, and child nutrition and diet indicators. Section 3.8 presents results on child malnutrition and section 3.9 provides a short description of PSM and an initial analysis that refers to it.

[9] We present the unweighted estimates for key indicators in Annex B.

We end this section with a discussion of methodological caveats and some implications of our findings (section 3.10).

3.4 Demographics and background characteristics

Box 1: Key findings on demographics and background characteristics

- **Households in the treatment and control groups differ slightly with respect to key demographic and background characteristics.** Treatment households are slightly smaller, with more working age members. Household heads in treatment districts are less likely to have ever attended school. Overall, these differences are small.
- **Generally, there are no significant differences in selected characteristics of eligible children in the treatment and comparison groups.** On average, children are about 12 months old. About 50% are girls. Almost all caregivers of children are female and, according to this data, the biological mother of the child.

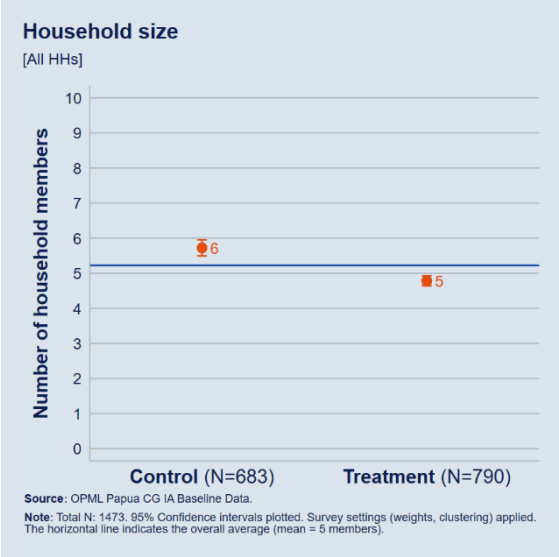
3.4.1 Households

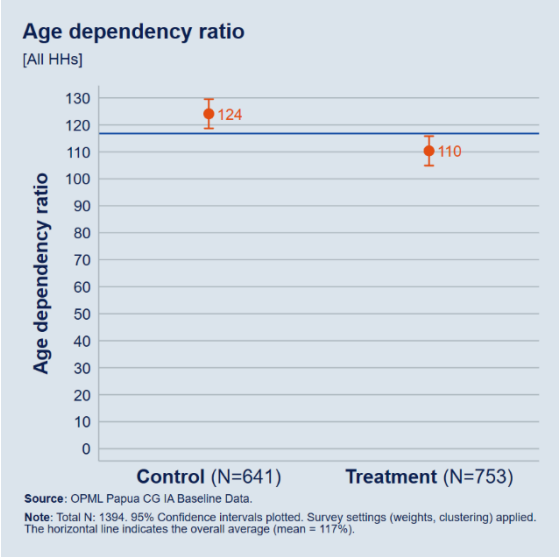
For the purposes of this study, we are defining households as a group of individuals who share a common cooking place (i.e. kitchen and 'budget' for food) and who identify one common person as the household head. In the context of Papua, applying this definition in the survey was sometimes difficult, as shared cooking arrangements across full villages were sometimes observed. In such situations, identifying the common household head – as the person who makes key decisions in a household – was important in defining a household. On average, the households surveyed in this study have about five members, with households in the control areas being slightly but significantly larger than household in the treatment districts (Figure 4 left). Almost all (over 95%) of these households were identified as male-headed (data not shown here).

One measure of household composition is the age dependency ratio. This is the ratio between the number of children (below 15 years old) and old people (above 64 years old) in the household, referred to as dependants, and all other household members, referred to as non-dependants.

In our survey, the average dependency ratio is about 117%, which means that there are 1.17 dependants per non-dependant in each household. Households in the control group have a significantly higher number of dependant members than households in the treatment group (Figure 4 right). This difference – together with the fact that households in the control group are slightly larger – indicates that these households have more working age members.

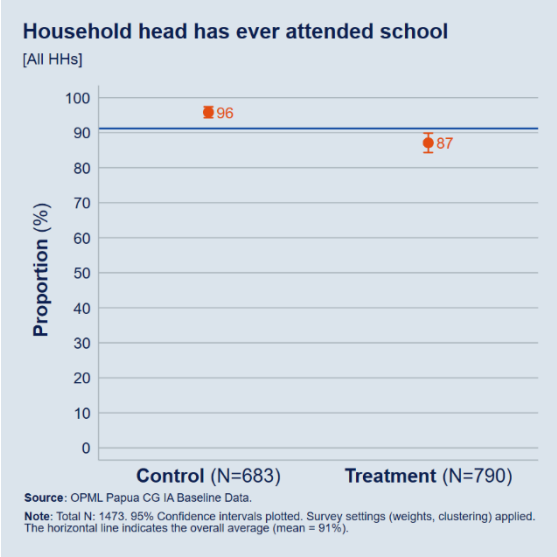
Figure 4: Average household size (left) and age dependency ratio (right)





Most (91%) of the household heads attended school. However, the level of education of the household head was significantly lower in the treatment group, where 87% of household heads had been to school, compared to 96% in the control group (Figure 5).

Figure 5: Proportion of household heads who have ever attended school



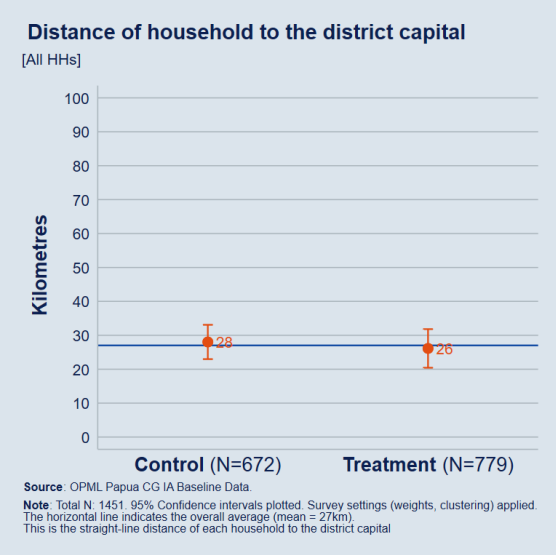
Given that the baseline survey was implemented using electronic tablets, we were also able to record the Global Positioning System (GPS) coordinate of where households were located. Together with the geo-location of district capitals, we were then able to calculate the straight-line distance of each household from the district capital.[10] It is important to note that this straight-line distance is only a proxy for the travel distance of households from their home to the district capital. This is because of the geographical

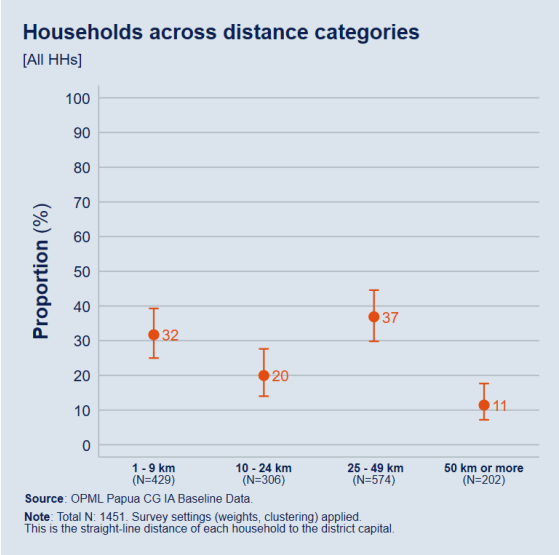
[10] The Stata command *geodist* was used for this purpose. The *geodist* command computes geodetic distances, i.e. the length of the shortest curve between two points along the surface of a mathematical model of the earth.

characteristics of Papua, where often roads being in poor condition, river flows, or mountain ranges might mean that people need to take routes that are not straight at all. Still, our hypothesis was that important outcome indicators would be correlated with this proxy in some cases, which we show further below.

Figure 6 shows the average distance of households to the district capital, in kilometres (km). On average, households are slightly less than 30 km away from the district capital, and this does not vary by treatment status (Figure 6 left). This average hides some variation, however. Figure 6 (right) shows that about 11% of households live over 50 km away from the district capital, with the maximum distance reaching over 100 km (not plotted here).

Figure 6: Average distance of household to the district capital (left) and proportion of households across different distance categories (right)

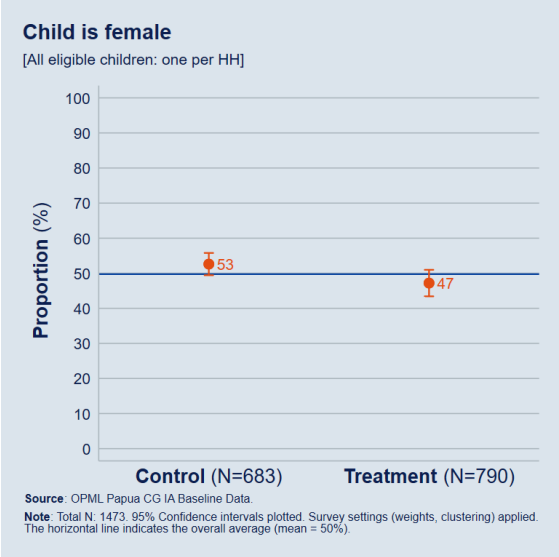


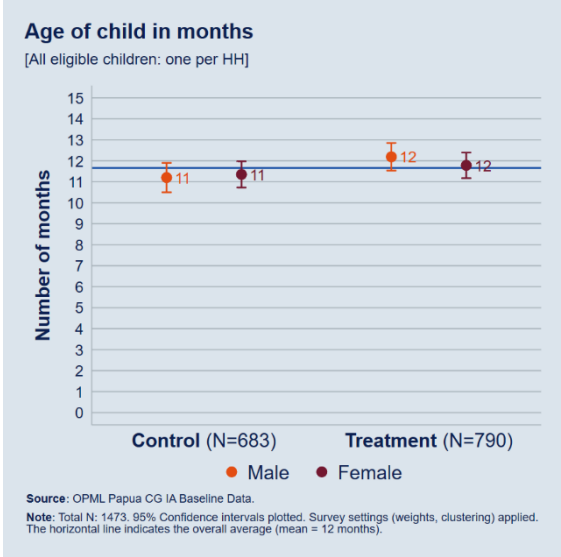


3.4.2 Eligible children

This section focuses on describing the sample of eligible children in our survey. Eligible children were defined as children aged 0–23 months who had a Papuan caregiver or parent. About 50% of the eligible children surveyed are girls, with no significant difference between the treatment and control groups (Figure 7 left). The average age of eligible children is 12 months in the treatment group and 11 months in the control group, regardless of the child's gender (Figure 7 right).

Figure 7: Proportion of girls (left) and average age of children (right)

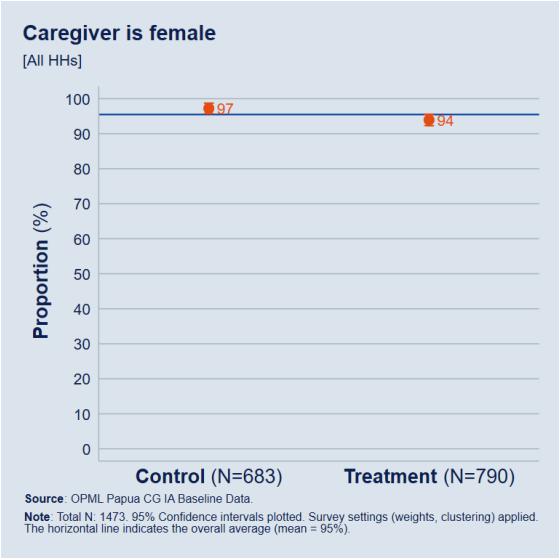


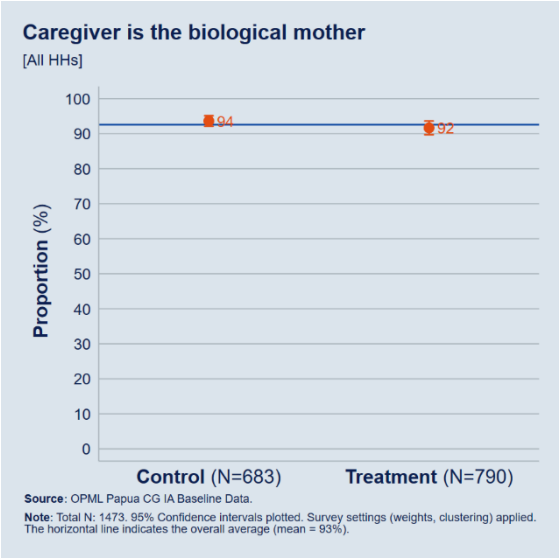


The large majority of caregivers of eligible children are female (Figure 8 left). Most of them indicate that they are the biological mother of the child. Percentages are similar across evaluation groups (Figure 8 right).

It should be highlighted here that this finding varies from what was presented by the Empatika study (Jupp *et al*., 2018), which emphasised that adoption and the sharing of care duties were very common in the villages visited as part of their qualitative research. It is difficult to assess where this discrepancy in observation originates from. Qualitative evidence shared by Empatika does indicate, however, that it was often only after a long probing process with caregivers that it became clear that they were not the biological parents to a child. This was because even step-parents had internalised the view that they were the father or mother to a child – irrespective of whether this was by birth or adoption. Although the OPML survey did include instructions for probing – and enumerators were trained to do so – it is possible that in the short time-span of the questionnaire caregivers did not instantly indicate that they were not the biological parent to a child.

Figure 8: Proportion of children whose main caregiver is female (left) and whose main caregiver is the biological mother (right)





3.5 Exposure of households to financial support from other programmes

Box 2: Key findings on exposure of households to financial support from other programmes

- **More than 40% of households in treatment areas received some financial support in the 12 months prior to the survey. In contrast, only 10% received such support in the control areas.** This is problematic for the impact evaluation. If these cash transfer interventions continue and have an impact on the indicators of interest for the evaluation, the measure of impact attributed to CG will be affected.

This section explores whether there were any other programmes providing financial support to households in the survey areas. Specifically, it investigates whether households have received cash support from the government or other organisation up to one year prior to the survey. The survey asked about the main national cash transfer programmes, as well as district-specific interventions.[11]

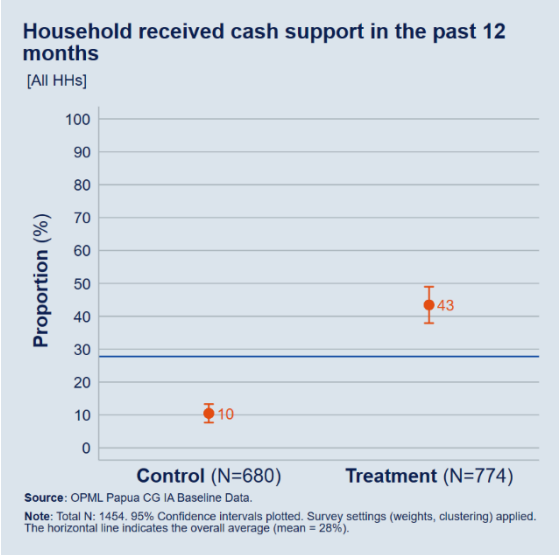
It is important to emphasise that these findings have methodological implications for the quantitative impact assessment analysis at endline. Programmes that are being implemented in parallel to the CG in Papua have the potential to contaminate or confound our impact assessment findings. Including questions about these programmes in our survey is therefore important in order to be able to control for

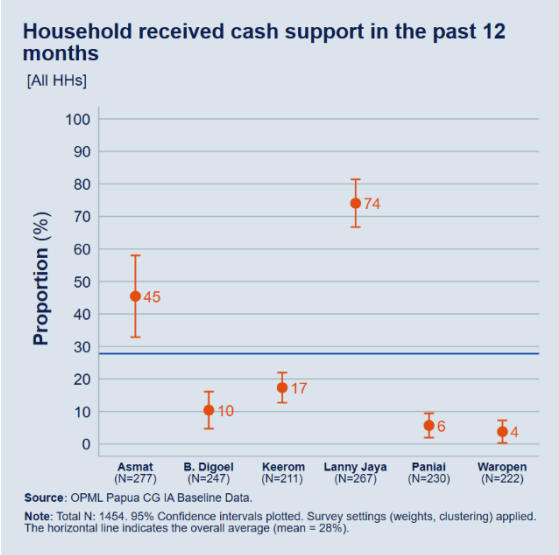
[11] These are *Program Indonesia Pintar* (PIP), *Program Keluarga Harapan* (PKH), *Raskin/Rastra* (Rice subsidy), *Bantuan Siswa Miskin* (BSM), *Bantuan Langsung Sementara Masyarakat* (BLSM), as well as Bangga Papua CG and cash from NGOs or other organisations. Anecdotal evidence indicates that Rastra possibly includes a cash component in some communities in Papua, along with the usual in-kind support, in the form of a package of basic food or vouchers to be used for rice and other staple goods.

alternative sources of support that households might be receiving. We are also aiming to collect comprehensive secondary information about these programmes in order to be able to take this into account during endline analysis. Further detail on this methodological caveat is presented in section 3.10.

Figure 9 below shows that, firstly, over 40% of households in treatment areas and about 10% of households in control areas report that they received cash support. In addition, it shows that this estimate not only varies by treatment group but also by district. In Lanny Jaya a large majority of households (74%) receive some sort of financial support. Almost half of households benefit from cash support in Asmat (45%). The percentage is below 20% in the other districts (Figure 9 right). Overall, this reflects the fact that in many areas in Papua some sort of social support system is in place, with a particular focus on poorer districts that are also being targeted by the CG.

Figure 9: Proportion of households receiving cash support in the 12 months prior to the survey by treatment group (left) and by district (right)





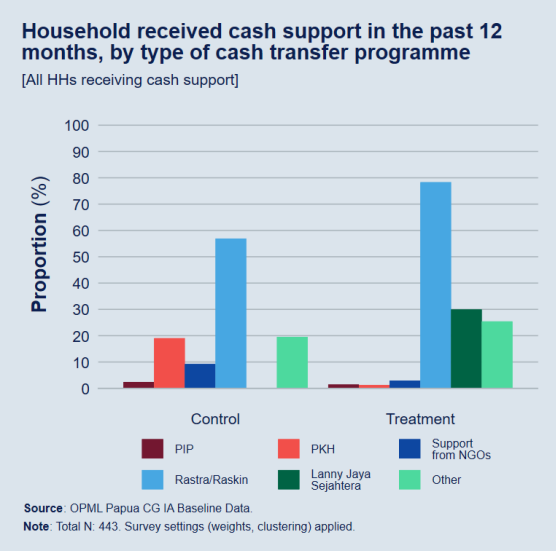
Our data also reveal that, of the households that do receive financial support, most are part of the rice subsidy scheme, Rastra (formerly known as Raskin) (Figure 10 left). This is especially the case in Asmat and Lanny Jaya, where Rastra covers 89% and 74% of households receiving some form of cash support, respectively (Figure 10 right).

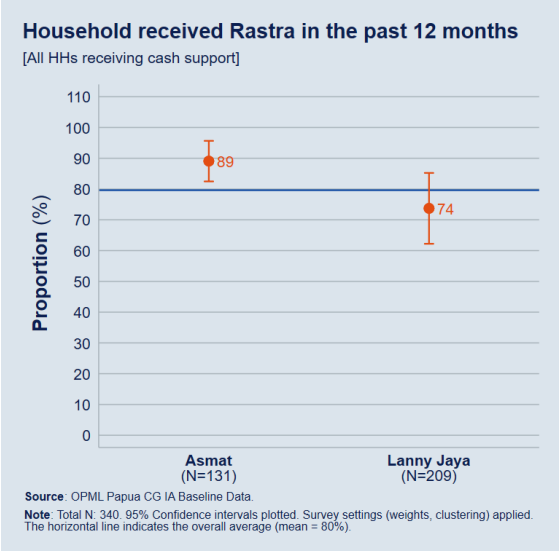
In general, very few households are enrolled in Indonesia's education cash assistance program, *Program Indonesia Pintar* (PIP), and PKH. With regard to local cash distribution schemes, the household survey also found that cash handouts from local initiatives and village funds are widespread, especially in Lanny Jaya (Lanny Jaya Sejahtera; Figure 10 left). Anecdotal evidence collected during validation visits in Papua and qualitative findings from Empatika (Jupp *et al*., 2018) confirm the low coverage of the major national social assistance programmes and the existence of many local social security and cash support programmes in Lanny Jaya and Asmat.

As explained above, the differential existence of these support programmes – across treatment and control groups – poses a risk to our impact identification strategy. Because cash support beyond BANGGA Papua is much more common in the treatment areas compared to the control areas, it is possible that differences in outcomes that we might potentially observe at endline could be due to both the CG itself and these other support programmes. We are aiming to deal with this issue by

collecting comprehensive information on the support that households receive at endline again and to use this information to control for programmes other than the CG and to conduct sensitivity analyses. For example, we could compare households that did not receive any support other than the CG in treatment areas to households that did not receive any support in the control areas to see whether any observed differences persist. Of course, implementing such analyses relies on households being able to differentiate between BANGGA Papua payments and other payments in our survey.

Figure 10: Cash transfer coverage by type of programme (left) and proportion of households receiving Rastra in Asmat and Lanny Jaya (right)





3.6 First-order outcomes: household consumption levels, consumption patterns, and location of consumption

Box 3: Key findings on household consumption levels, patterns, and location

- **Households in the treatment group reported significantly lower monthly consumption than households in the control group.** Despite this, we estimate that the proportions of the population below the Papuan poverty line do not vary by treatment status. Overall, we estimate that 17% of the surveyed population in the treatment and control districts falls below the Papuan poverty line, as set by BPS in September 2018.
- **Households mainly consume food (71% of total monthly consumption on average)**, followed by household maintenance, electricity costs, and durable goods, clothing and toiletries, and tobacco. Education expenses and health expenses are relatively small.
- **A large majority of household usually buy their consumption goods within their sub-districts.**

This section presents estimates on household consumption and expenditure as captured by our survey. The questionnaire used in this survey was based on the BPS

SUSENAS questionnaire from March 2018. Our consumption and expenditure module was designed to cover items that together comprised 95% of the total household consumption as measured by SUSENAS in Papua in 2017.

Before describing results, it is important to emphasise a few important features of our data related to the consumption and expenditure of households:

- First, it is essential to reiterate that estimates presented below refer only to the population targeted by our survey, i.e. Papuan households with children under two years old in a set of accessible areas in treatment and control districts. Hence, these estimates and are not representative of the overall population in Papua.
- Second, as part of the analysis presented below, we used the Papua poverty line, as set by BPS in September 2018, as a threshold to provide an indication of households' material level of wellbeing.[12] It is important to emphasise that, while this helps to describe the welfare of households, it fails to capture other dimensions of wellbeing, providing a narrow understanding of consumption-related poverty. Other dimensions of wellbeing (e.g. those related to the nutritional status of children in households captured by this survey) are covered in other sections of this report.
- Third, the binary welfare indicator measuring the percentage of the population who fall below the poverty line is inevitably sensitive to the choice of this poverty line. Small changes in this poverty line can have large effects on the percentage estimate. We present estimates for this sensitivity in Figure 15 below.

Please note that Annex A.1 provides a technical description of how consumption data was used in this analysis and presents results for a set of related robustness checks.

3.6.1 Consumption levels and patterns

Figure 11 presents consumption estimates and the consumption patterns for households in the treatment and control groups. Total monthly household consumption is estimated to be at around IDR 5,700,000 on average (Figure 11 left). This corresponds to monthly consumption expenditure per adult equivalent of around IDR 1,700,000 (US$ 114.95), on average (

Figure 12 left).[13] Households in the treatment group consume around IDR 4,700,000, which is more than IDR 2,000,000 less on average than households in the control group (IDR 6,900,000). Note that the transfer from BANGGA Papua (IDR 200,000 per

[12] The food poverty line can be found at www.bps.go.id/dynamictable/2016/01/18/1123/garis-kemiskinan-makanan-gkm-menurut-provinsi-2015---2018.html and the non-food poverty line at www.bps.go.id/dynamictable/2016/01/18/1124/garis-kemiskinan-non-makanan-gknm-menurut-provinsi-2015---2018.html.

[13] Equivalence scales help to assign households consumption values that are in proportion to the household type needs. The factors commonly taken into account to assign these values are the size of the household and the age of its members (whether they are adults or children). A wide range of equivalence scales exist. For the purpose of this evaluation, we used the Organisation for Economic Co-operation and Development (OECD) equivalence scale. This assigns a value of 1 to the first household member, of 0.7 to each additional adult, and of 0.5 to each child. See www.oecd.org/eco/growth/OECD-Note-EquivalenceScales.pdf

month per child) therefore roughly corresponds to less than 5% of the average estimated household consumption per month (assuming households have one eligible child) in treatment households.

Food makes up 70% of household monthly consumption in the control districts and 73% in the treatment districts (Figure 11 right). Other large shares of consumption are household needs (13%) – which includes household maintenance, electricity, and durable goods – clothing and toiletries (7%), and tobacco (6%) (Figure 11 right).

Figure 11: Average monthly household consumption (left) and distribution of monthly consumption by item (right)

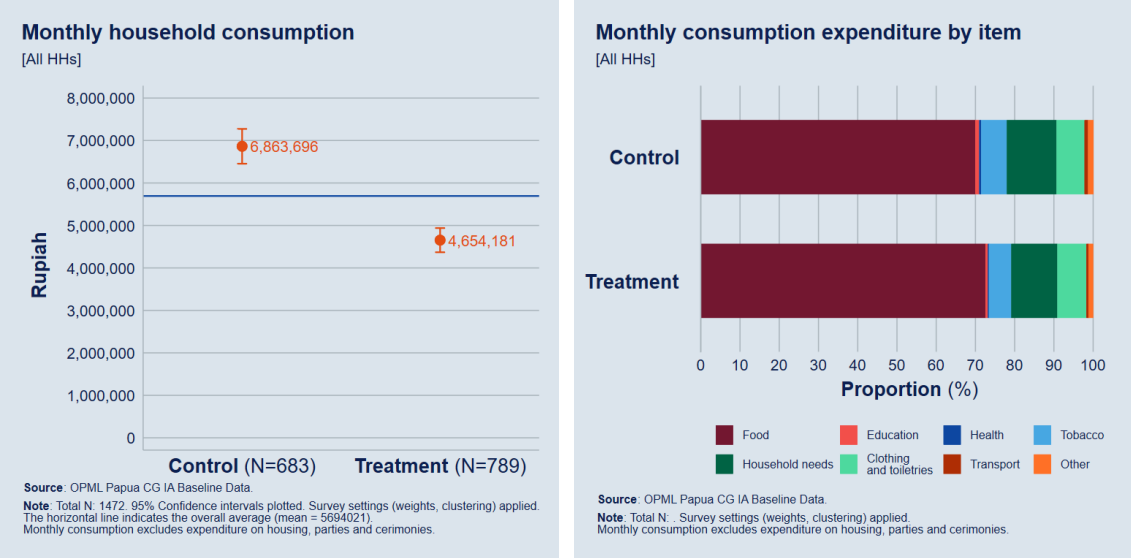
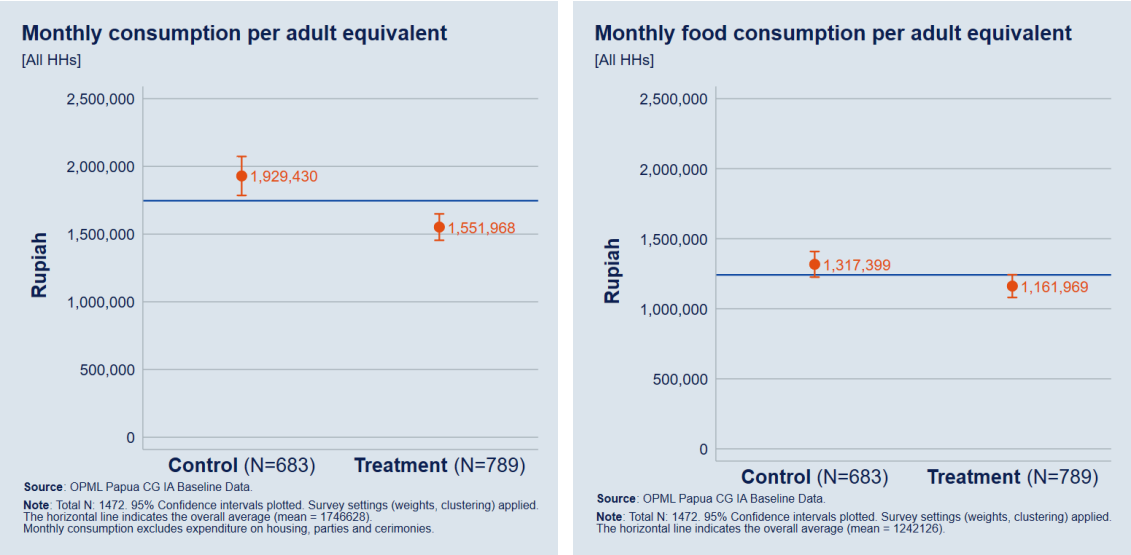


Figure 12: Average monthly consumption per adult equivalent (left) and average monthly *food* consumption per adult equivalent (right)



When looking at these consumption patterns in a bit more detail, we find that household food consumption varies across district, reflecting local diets and food availability (Figure 13). Households primarily consume rice, potatoes, roots and other tubers, followed by vegetables and meat (mainly fish and chicken). Tea, coffee, and

sugar make up 8% of total food consumption on average (Figure 13). As expected, education expenses and health expenses are relatively small, given that they are subsided by the government (Figure 14).

Figure 13: Food consumption share by district

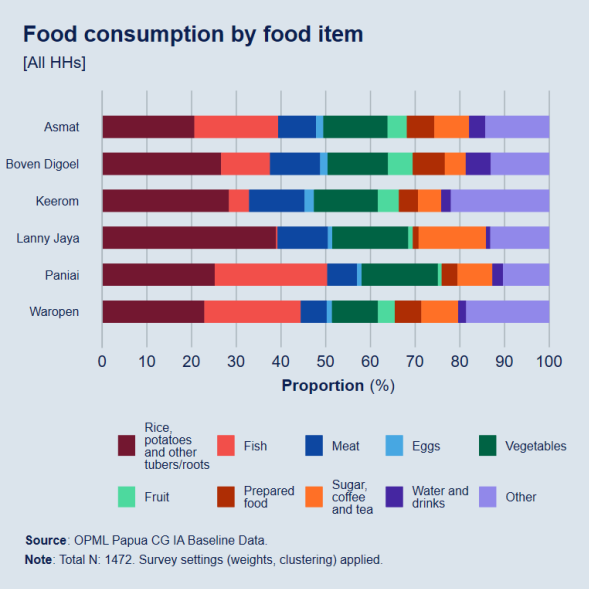
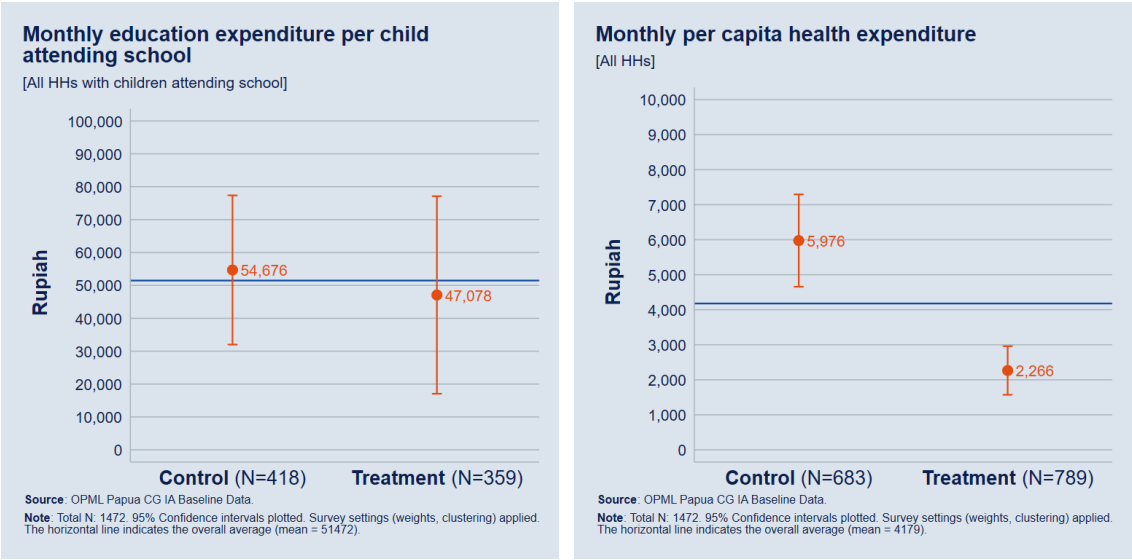


Figure 14: Average monthly education expenditure per child attending school (left) and average monthly health expenditure *per capita* (right)

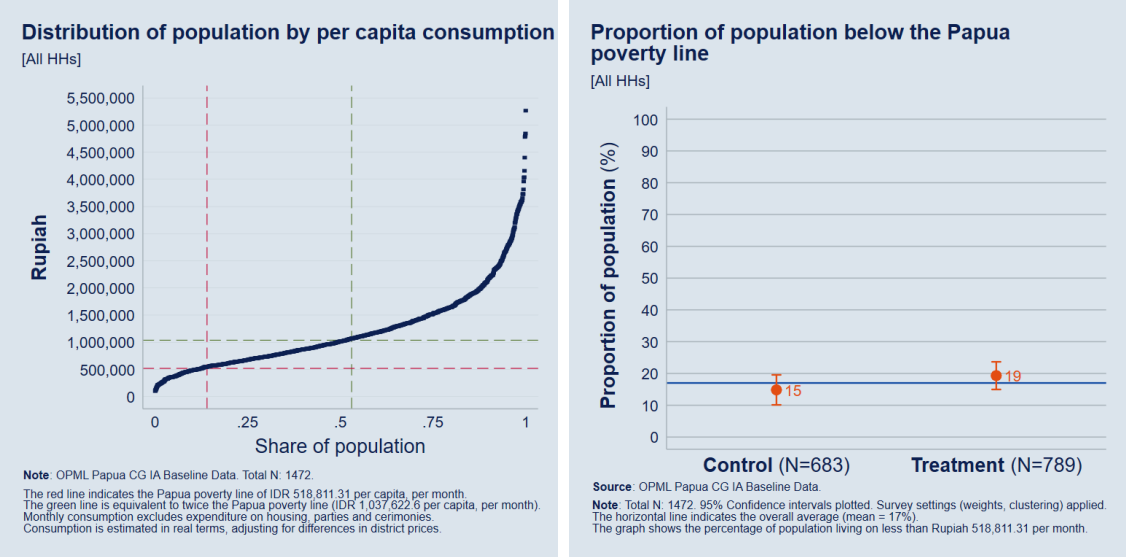


The left-hand panel of Figure 15 presents the distribution of households by consumption *per capita*. In this graph, households have been sorted by *per capita* consumption: the households with lowest *per capita* consumption are on the left, while households with the highest *per capita* consumption are on the right of the horizontal scale (the x-axis). The vertical axis (the y-axis) shows the level of *per capita* consumption for each household. The Papua poverty line that corresponds to a *per capita* consumption of IDR 518,811.31, as set by BPS in September 2018, has been

indicated with a horizontal red dashed line.[14] Similarly, double that poverty line has been indicated with a horizontal green dashed line.

The graph shows that 17% of the population lies to the left of the point where the Papua poverty line meets the distribution of *per capita* household consumption (see the intersection between the two red lines in Figure 15). Doubling this poverty line would increase the proportion of the population living below the poverty line to more than 50% of the population (see the intersection between the two green lines in Figure 15). This suggests two things. First, the estimate of the population living below the poverty line is highly sensitive to the exact location of this poverty line and it does not change proportionally with it, as doubling the poverty line more than doubles the estimate. Second, the wealth distribution is highly concentrated around the poverty line, i.e. many households have a slightly higher *per capita* consumption than the poverty line but not by large margins. The right-hand panel in Figure 15 shows that the percentage of the population living below the poverty line is 19% in the treatment districts and 15% in the control districts, although the difference is not statistically significant.[15]

Figure 15: Distribution of population by *per capita* consumption (left) and percentage of population below the Papua poverty line (right)



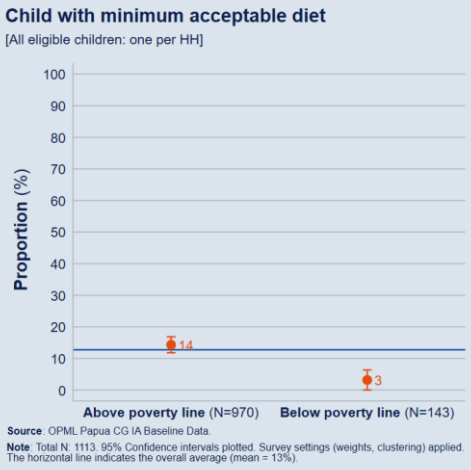
[14] Food poverty line: www.bps.go.id/dynamictable/2016/01/18/1123/garis-kemiskinan-makanan-gkm-menurut-provinsi-2015---2018.html; non-food poverty line: www.bps.go.id/dynamictable/2016/01/18/1124/garis-kemiskinan-non-makanan-gknm-menurut-provinsi-2015---2018.html

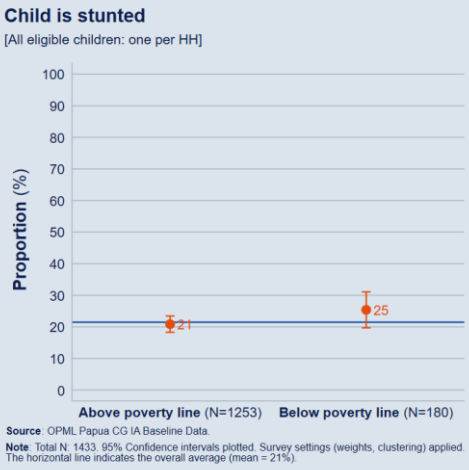
[15] As part of the consumption analysis we also computed the proportion of the population whose monthly *per capita* consumption is below the Papua food poverty line, as well as the poverty gap index measuring the depth of poverty in the treatment and control groups. Results are presented in Annex A.1. In addition, we ran several sensitivity checks to assess how our findings are robust to changes to the methodology. Annex A.1 provides more details.

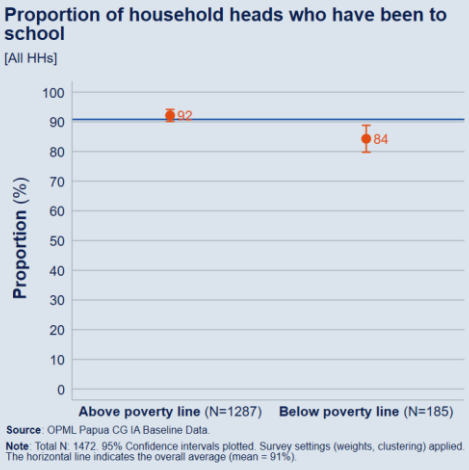
Box 4: Correlation between household consumption and other dimensions of wellbeing

The graphs below show how certain key characteristics of households and children, which capture their wellbeing and welfare beyond household consumption, vary by whether households lie below or above the poverty line. In each graph, we plot the proportion of children or households with a certain characteristic disaggregated by whether household consumption is estimated to be below or above the Papuan poverty line.

The top left panel shows that, within households below the poverty line, children are much less likely to have a minimum acceptable diet. The top right panel shows that children in households below the poverty line are also more likely to be malnourished, as measured by child stunting. The third graph in the bottom panel shows that, among households below the poverty line, household heads are less likely to have been educated in school.







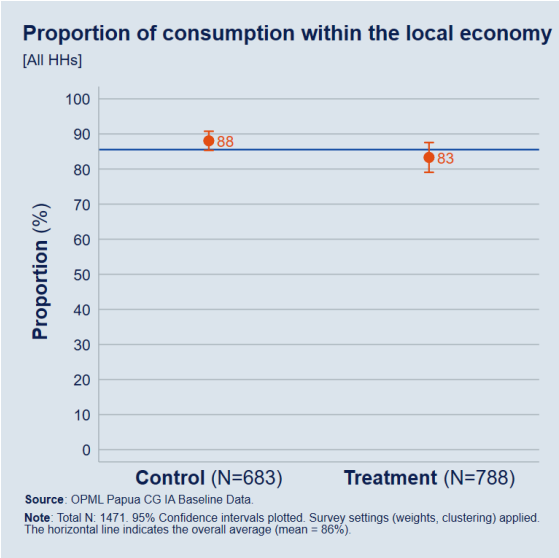
Overall, this analysis shows that low levels of consumption are an indicator for general household poverty, as measured by other indicators. Sections 3.7.4 and 3.8 provide a full description of baseline findings with respect to child feeding and child malnutrition.

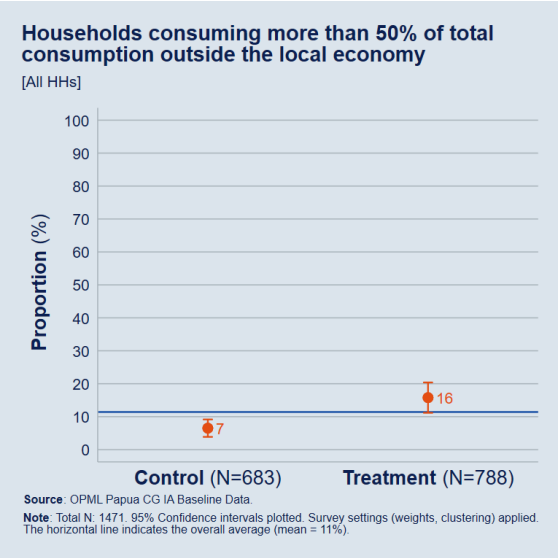
3.6.2 Location of consumption

In our survey, we also asked households about where they obtained (self-produced or purchased) items that they reported consumption on. More specifically, we asked about whether the items were obtained within or outside the sub-district boundaries in which households were located. We considered items within the sub-district to be consumed 'within the local economy'. This analysis of the location of consumption sheds light on the characteristics of the local economy, as captured by the sub-district boundaries. It also provides us with an insight into whether we can assume that increased income from the CG will be spent locally or not.

The graphs below show that, firstly, household consumption lies mostly within the local economy (Figure 16 left). Second, the proportion of households whose monthly consumption is mostly purchased outside the local economy is significantly higher in the treatment group (16%) compared to control areas (7%) (Figure 16 right).

Figure 16: Location of consumption





3.7 Second-order outcomes: education, birth registration, early childhood development, childcare, and child feeding

Box 5: Key findings on children education, birth registration, early childhood development, childcare, and child feeding

- **Overall, access to education is low, especially in the treatment group.** School attendance among children is at 32% in treatment areas and 60% in the control areas. Higher pre-school but lower school attendance in the treatment group indicates that children here tend to stay in pre-school longer than in the control group. Drop-out rates are low (6% on average), but significantly higher for girls.
- **Birth registration is low**, at 20% on average. The proportion of children with a birth certificate is significantly lower in the treatment group (15%).
- **Availability of learning materials, defined as children's books and playthings, is low.** Children in the control group are more likely to have playthings, compared to those in the treatment areas, although the overall average is below 30%.
- **The majority of boys and girls have been left alone for more than one hour at least once in the week prior to survey.** This is more common in the treatment areas, compared to the control areas.
- **On average 53% of children experienced some sort of violence**, with the percentage being higher in the control group, compared to the treatment group.
- **Diets and feeding practices of children in the study area are inadequate.** Most children do not achieve the minimum requirements in terms of dietary diversity and only slightly above 50% of children meet the recommended minimum meal frequency. Findings are similar across evaluation groups.

This section presents baseline findings for second-order outcome indicators in this evaluation. Second-order outcome indicators are those outcomes presented in the ToC (sections 3.1 and 3.2) that might be indirectly affected by the increased income derived from the CG.

Table 3 presents the key second-order outcome indicators used in this impact assessment.

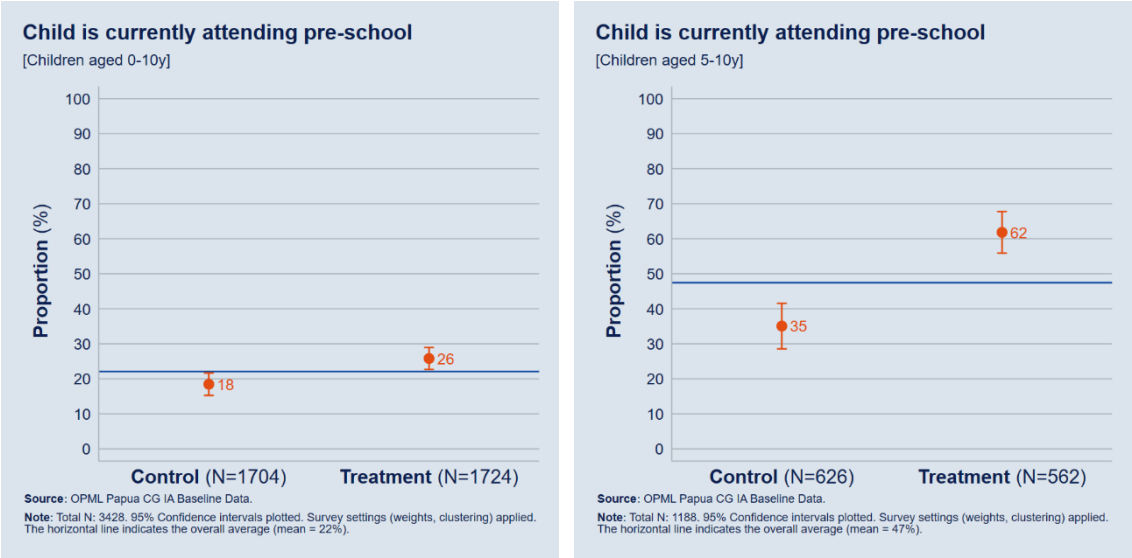
Table 3: Second-order outcome indicators

Dimension	Indicator
Access to education	• Percentage of children currently attending / ever been to pre-school • Percentage of children attending school
Birth registration	• Percentage of children under age 17 who have a birth certificate from the civil registry office
Early childhood development and childcare	• Availability of children's books: Percentage of children under two years who have three or more children's books available in their household • Availability of playthings: Percentage of children under two years who have two or more types of plaything available in their household • Inadequate supervision: Percentage of children under two years left alone or under the supervision of another child younger than 10 years for more than one hour at least once in the last week • Violent discipline: Percentage of children aged 1–2 years who experienced any physical punishment and/or psychological aggression by caregivers in the past one month
Child feeding and diet	• Minimum acceptable diet: percentage of children aged 6–23 months who had at least the minimum dietary diversity and the minimum meal frequency during the previous day • Percentage of children aged 6–23 months who had at least the minimum dietary diversity during the previous day • Percentage of children aged 6–23 months who had at least the minimum meal frequency during the previous day • Exclusive breastfeeding: percentage of infants aged under six months who are exclusively breastfed

3.7.1 Access to education

Figure 17 shows that pre-school attendance among children aged 0–10 years among our target household population is at 18% in the control group and 26% in the treatment group. Further disaggregation by age shows that these proportions vary, depending on the age group looked at. The percentage is significantly higher among children aged 5–10 and, in particular, in treatment areas (Figure 17 right). This suggests that children tend to stay longer in pre-school in treatment districts compared to control areas. There is anecdotal and qualitative evidence that this might be due to the fact that, in treatment areas, households are encouraged to send their children to pre-school by free distribution of food at these facilities, as part of government initiatives to improve children's nutritional status.

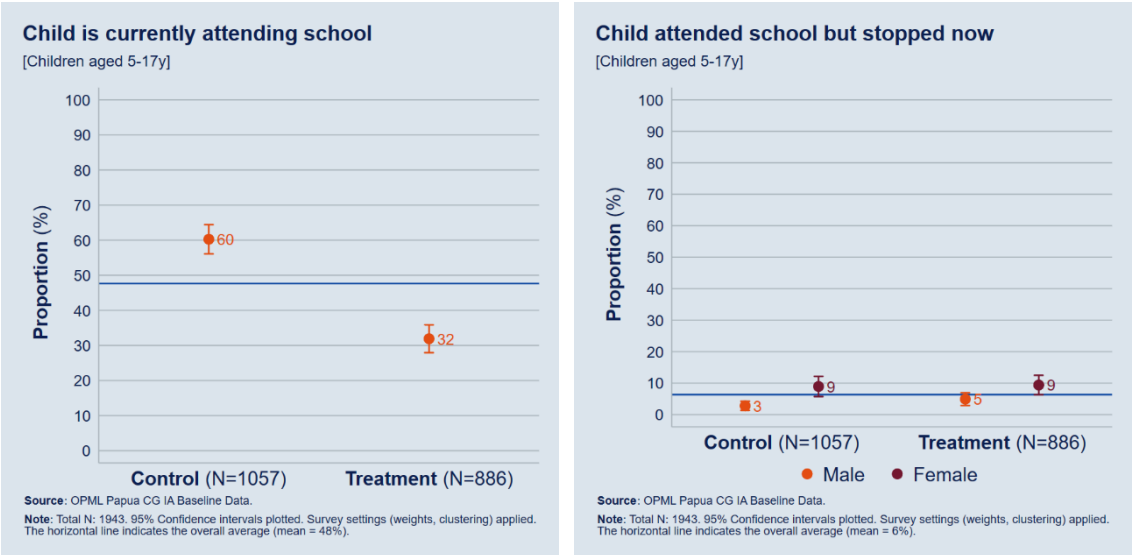
Figure 17: Proportion of children aged 0–10 years (left) and 5–10 years (right) currently attending pre-school



The left-hand panel in Figure 18 shows that school attendance is around 48% among children aged 5–17 years old, on average. The percentage is significantly lower in the treatment areas (32%), where children are about 30 percentage points less likely to attend school than in the control areas (60%). Together with the findings on pre-school attendance presented above, this indicates that in treatment areas children are more likely to stay longer in pre-school than in control areas.

Dropout, as measured by the percentage of children who stopped attending school but were at school in the year prior to the survey, is at 6% on average, with no difference between treatment and control groups. Dropout is significantly higher for girls than boys (Figure 18 right).

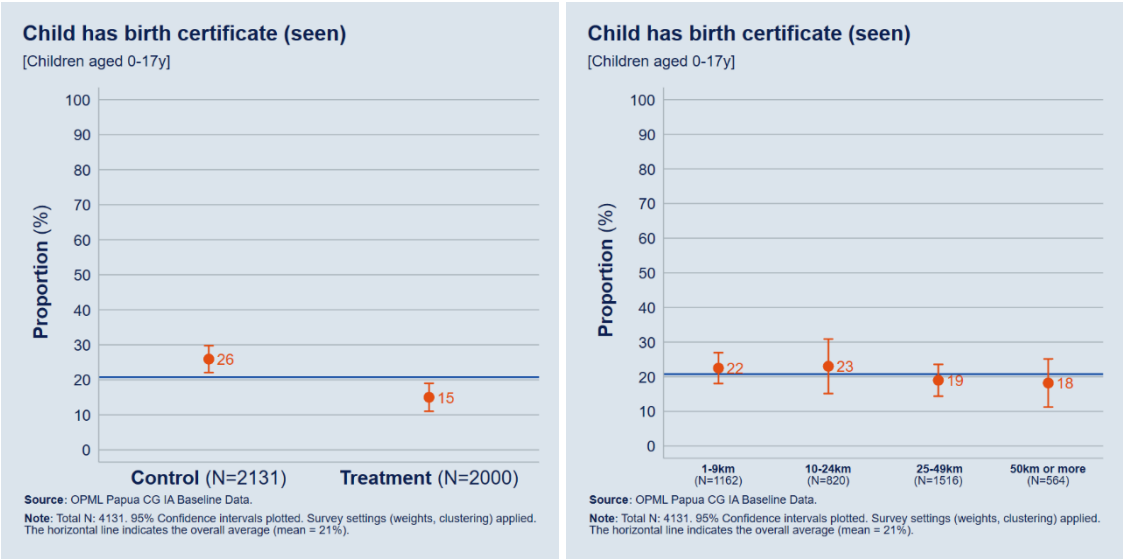
Figure 18: Proportion of children aged 5–17 years currently attending school (left) and who dropped out from school (right)



3.7.2 Birth registration

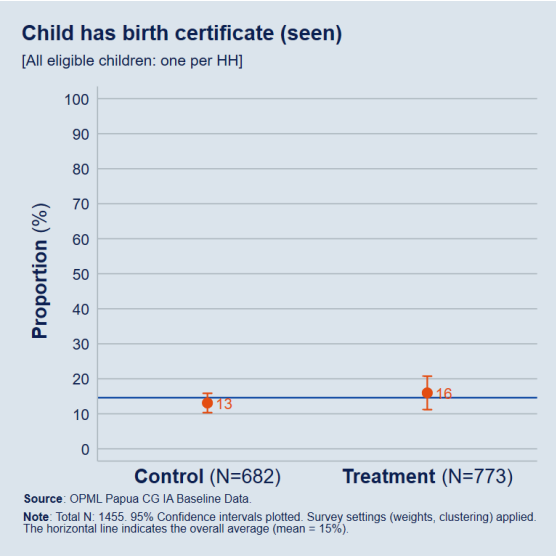
In our survey, we also ask respondent households whether they can provide birth certificates for their children. These are documents produced by a civil registration office. In our survey, enumerators asked to actually see the certificates to prove their existence. Figure 19 shows that, overall, the percentage of children with valid birth certificate is 21% across treatment and control groups. The percentage is significantly lower in the treatment group. In treatment areas, 15% of children aged 0–17 years old have a birth certificate, compared to 26% in the control group (Figure 19 left). The likelihood of children being registered with the civil authorities decreases slightly moving away from the capital, although the correlation is not significant in statistical terms (Figure 19 right).

Figure 19: Proportion of children aged 0–17 years who have a birth certificate (left), by distance from the district capital (right)



Focusing on the sample of eligible children, birth registration is lower than the average rate measured for the entire sample of children, with 15% of children aged 0–23 months being registered with the civil authorities on average (Figure 20). This indicates that households take time to register their children, despite the fact that birth registration should happen soon after birth.

Figure 20: Proportion of eligible children who have a birth certificate (seen by the enumerator)



3.7.3 Early childhood development and childcare

Our survey looked at four key indicators related to early childhood development and childcare: availability of children books in households, availability of playthings, adequate supervision of children by adults, and the use of violence in childcare.

We measure these indicators using a subset of questions from the early childhood development section in UNICEF's Multiple Indicator Cluster Survey (MICS) questionnaire for children under five (UNICEF, 2019).The percentage of children with three or more books is very low (2% on average), with no differences between evaluation groups and children's gender (Figure 21). When it comes to playthings, children in the control group are more likely to have two or more playthings, compared to children in the treatment group (33% in the control group and 26% in the treatment).[16] The difference is even bigger when focusing on manufactured toys (45% in the control group compared to 31% in the treatment group). Girls seem to have less access to playthings compared to boys (Figure 22).

[16] Playthings include: Homemade toys, such as dolls, cars, or other toys made at home; Toys from a shop or manufactured toys; Household objects, such as bowls or pots, or objects found outside, such as sticks, rocks, animal shells, or leaves (UNICEF, 2019).

Figure 21: Proportion of children with three or more books

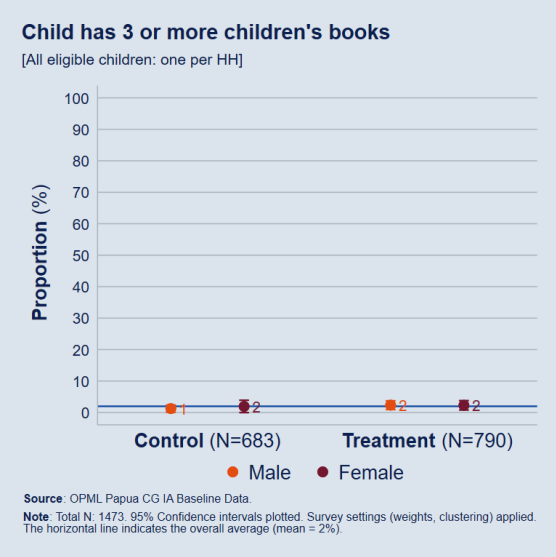
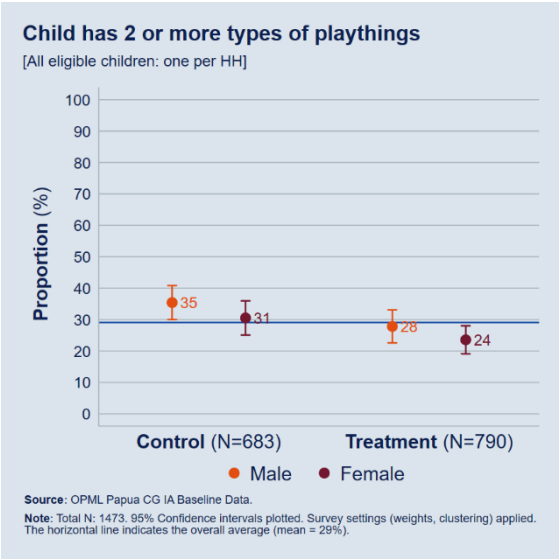
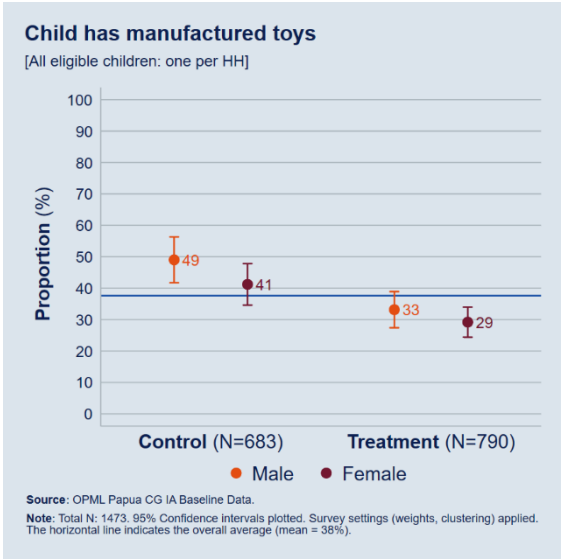


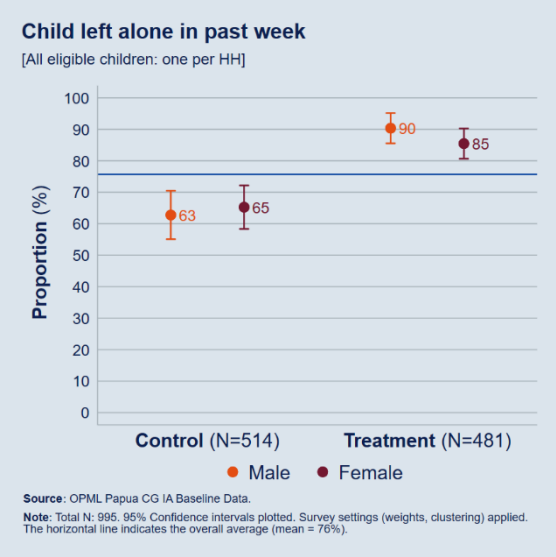
Figure 22: Proportion of children with two or more playthings: any type of toys (left) and manufactured toys (right)





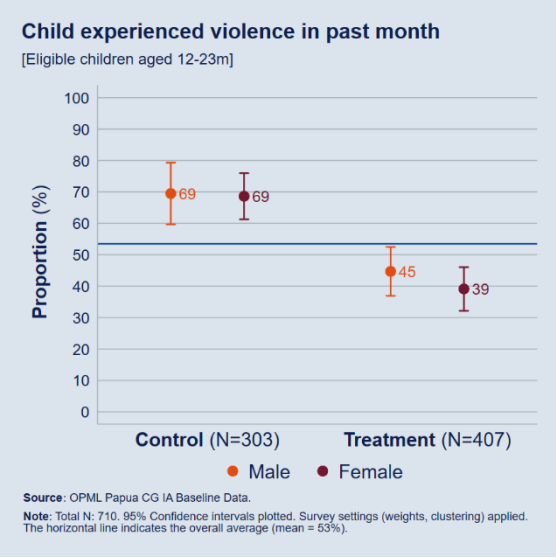
When it comes to childcare, our survey reveals that the majority of boys and girls have been left alone, or in the care of another child younger than 10 years, for more than one hour at least once in the week prior to survey. The percentage is significantly higher in the treatment group, where it is around 90% (Figure 23). This indicates that children in the treatment group are more likely to care for themselves than in the control group, even at a young age, and correlates well with qualitative evidence produced by Empatika (Jupp *et al*., 2018), where children have often been observed roaming villages on their own.

Figure 23: Proportion of children left alone in the past week



Across Papua, physical and psychological violence is commonly used in the contexts of bringing up, raising, or educating a child. On average, and according to their caregiver, one of two children (53%) between 12 and 23 months old experienced violence in the month prior to the survey.[17] The percentage is significantly higher in the control group (69%) compared to the treatment group (42%) (Figure 24).

Figure 24: Proportion of children who suffered some sort of violence in the past month



3.7.4 Child feeding and children's diets

As per BANGGA Papua's ToC (section 3.1), an important area of expected benefits of the CG is an improvement in child feeding and the diets of children in households that

[17] Violence is defined as any physical punishment and/or psychological aggression in the month prior to the survey (UNICEF, 2019).

receive the grant. Our survey therefore also included a dietary recall module. This module asked the respondent, i.e. the child's caregiver, to list all meals and ingredients that a child consumed during the day prior to the survey. We followed international guidance on how to design and administer this module (WHO, 2008).

Our estimates show that, on average, only 23% of children aged 6–23 months in the target population have a diet with minimum dietary diversity, with the percentage in the treatment group (20%) being below the overall average (Figure 25 left). According to the FANTA indicators for assessing IYCF practices (WHO, 2008), a child whose diet conforms to a minimum of diversity has a high likelihood of consuming at least one animal-source food and at least one fruit or vegetable a day, in addition to a staple food (grain, root, or tuber).[18] These results mean that under a quarter of all children in eligible households achieve such minimum dietary diversity and that this proportion is even lower in the treatment areas.

The right-hand panel in Figure 25 shows the percentage of children (6–23 months old) who received solid, semi-solid, or soft foods the minimum number of times or more the day prior to the survey, referred to as minimum meal frequency. Minimum meal frequency is at 53% in the treatment group and 54% in the control group (see Figure 25 right).[19] Given the complexity of measuring actual caloric intake in a household survey, this indicator is a proxy for energy intake from foods other than breast milk.[20] Note that this indicator is significantly higher than the indicator on minimum meal frequency, which indicates that many children might eat often enough, but not from a very varied selection of foods.

[18] The proportion of children with minimum dietary diversity is the proportion of children aged 6–23 months who receive foods from four or more food groups among: grains, roots, and tubers; legumes and nuts; dairy products (milk, yogurt, and cheese); flesh foods (meat, fish, poultry, and liver/organ meats); eggs; vitamin-A rich fruits and vegetables; and other fruits and vegetables, on the previous day. See www.fantaproject.org/monitoring-and-evaluation/iycf-indicators

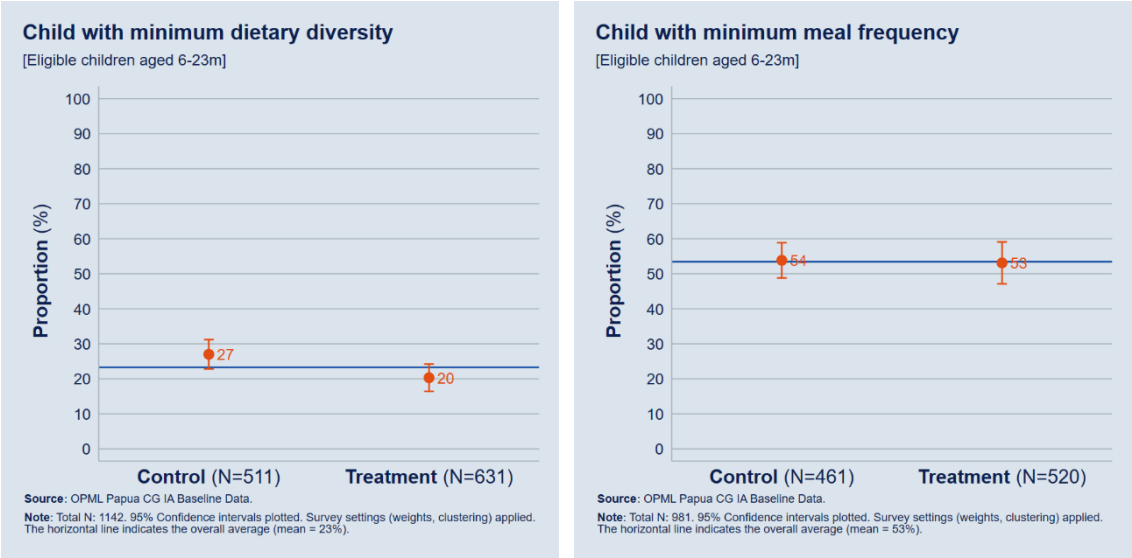
[19] Minimum is defined as:

- Two times for breastfed infants aged 6–8 months
- Three times for breastfed children aged 9–23 months
- Four times for non-breastfed children aged 6–23 months

'Meals' include both meals and snacks (other than trivial amounts), and frequency is based on caregiver report. See www.fantaproject.org/monitoring-and-evaluation/iycf-indicators

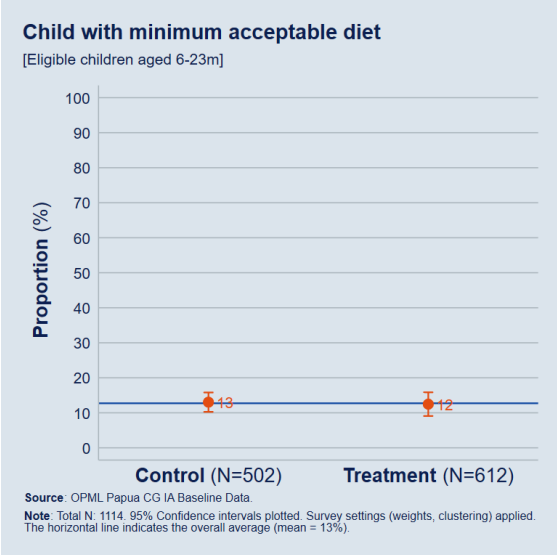
[20] See www.fantaproject.org/monitoring-and-evaluation/iycf-indicators

Figure 25: Proportion of children with minimum dietary diversity (left) and with minimum meal frequency (right)



Estimates for the summary indicator for minimum acceptable diet, which takes into account both meal frequency and dietary diversity, confirm that below 15% of children have a minimum acceptable diet in the study areas, on average.[21] The percentage is 13% in the control group and 12% in the treatment group (Figure 26). Overall, these results indicate that many children in both treatment and control areas do not have an acceptable diet.

Figure 26: Proportion of children with minimum acceptable diet



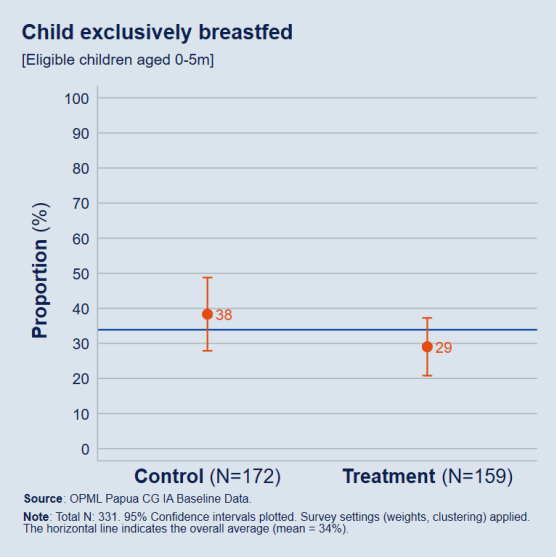
Our survey also assessed whether young infants aged 0–5 months were being exclusively breastfed or not. Exclusive breastfeeding means that infants did not

[21] The minimum acceptable diet indicator measures the proportion of children aged 6–23 months who receive a minimum acceptable diet (apart from breast milk). This is a combination of minimum dietary diversity and minimum meal frequency, with specific adjustments for breastfed and non-breastfed children. See www.fantaproject.org/monitoring-and-evaluation/iycf-indicators

consume anything else other than breastmilk from their mother during the day prior to the survey. In particular, this excludes water and small snacks. It does allow, however, for medicine – e.g. oral rehydration solution, drops, and syrups (vitamins, minerals, and medicines) – to be fed to children (WHO, 2007).

Overall, 34% of mothers of children aged 0–5 months report that they exclusively breastfeed their child. The percentage is around 10 percentage points higher in the control group (38%), compared to the treatment group (29%). However, this difference is not statistically significant (Figure 27).

Figure 27: Proportion of children aged 0–5 months who are exclusively breastfed



3.8 Third-order outcomes: child malnutrition

Box 6: Key findings on child malnutrition

- **Overall, about 20% of children are acutely malnourished, as measured by the prevalence of wasting.** In two districts, children are significantly more likely to be acutely malnourished. In both Asmat and Waropen, more than a quarter of all children are estimated to be wasted.
- **These findings are worrying. The WHO considers prevalence of wasting above 15% to be critical.** A high incidence of wasting usually points to insufficient food intake among children or to a high incidence of infectious diseases, like diarrhoea. In particular, UNICEF and WHO estimate that children who are severely acutely malnourished (SAM) have a significantly increased risk of death compared to children who are healthy (nine-fold compared to children with WLZ above -1 SD).
- Stunting can be interpreted as an indication of poor environmental conditions or a long-term restriction to a child's growth potential**. Stunting prevalence in the study area (21%) is classified as medium by WHO standards.** However, this prevalence varies across evaluation groups and districts. Stunting is higher in the treatment group (28%), especially in Paniai. In Paniai, there is a very high prevalence of stunting according to WHO guidelines, with 40% of children aged 0–23 months being stunted.
- **The prevalence of underweight (24%) is high according to the WHO standards for public health significance.** The percentage of children underweight is significantly higher in treatment districts, where it is around 28%, compared to 20% in the control areas.
- Child feeding practices in the study area are inadequate. Most of children do not achieve the minimum requirements in terms of dietary diversity and only slightly above 50% of children meet the recommended minimum meal frequency. Findings are similar across evaluation groups.

In our ToC framework, third-order outcomes refer to those that are usually mediated by external factors and relate to longer-term impacts of the CG. At this level, we measure child malnutrition prevalence, as captured by the prevalence of wasting, stunting, and underweight (see section 3.1).

In the baseline survey, we collected anthropometric data of the surveyed children aged 0–23 months, i.e. their length and their weight. In order to identify malnutrition among children, together with a child's age, these measurements were then transformed into Z-scores that compare them to a healthy reference population following WHO standards.[22] Three types of Z-scores were constructed:[23]

[22] See www.who.int/nutrition/publications/anthropometry-data-quality-report/en/
[23] See www.who.int/nutgrowthdb/about/introduction/en/index2.html

- **Weight-for-Length Z-scores (WLZ)**, which can be used to assess the prevalence of wasting or acute malnutrition. In most cases, wasting in children indicates a recent and severe process of weight loss, which is often associated with acute starvation and/or severe disease of children. However, wasting may also be the result of a chronic unfavourable condition, such as high incidence of infectious diseases, like diarrhoea. Children are considered to be wasted if they have WLZ scores below -2 Standard Deviations (SD). Children are considered to be severely wasted, or to suffer from severe acute malnutrition (SAM) if they have WLZ scores below -3 SD. If they have WLZ scores between -3 and -2 SD they are considered to suffer from moderate acute malnutrition (MAM). Wasting includes both categories of acute malnutrition.

- **Length-for-Age Z-scores (LAZ),** which can be used to assess the prevalence of stunting. Stunted growth in children reflects a process of failure to reach linear growth potential as a result of suboptimal health and/or nutritional conditions, and hence reflects long-term undernutrition. Children are considered to be stunted if they have LAZ scores below -2 SD.

- **Weight-for-Age Z-scores (WAZ),** which can be used to assess the prevalence of underweight children. Underweight reflects a mixture between stunted growth and wasting, and is hence difficult to interpret. Children are considered to be underweight if they have WAZ below -2 SD.

We present a rapid data quality assessment on our child malnutrition data in Annex A.5.

3.8.1 Prevalence of wasting

Figure 28 and Table 4 below present results on estimates of wasting prevalence by treatment group and gender. Overall, 19% of children are estimated to be wasted, with no significant difference between the two treatment groups. However, severe wasting, i.e. the prevalence of SAM, is significantly higher in the treatment group (8%) than in the control group (5%) (Table 4). Overall, the prevalence of wasting does not vary significantly by gender of the child (Figure 28 right).

Figure 28: Prevalence of wasting (left), by gender (right)

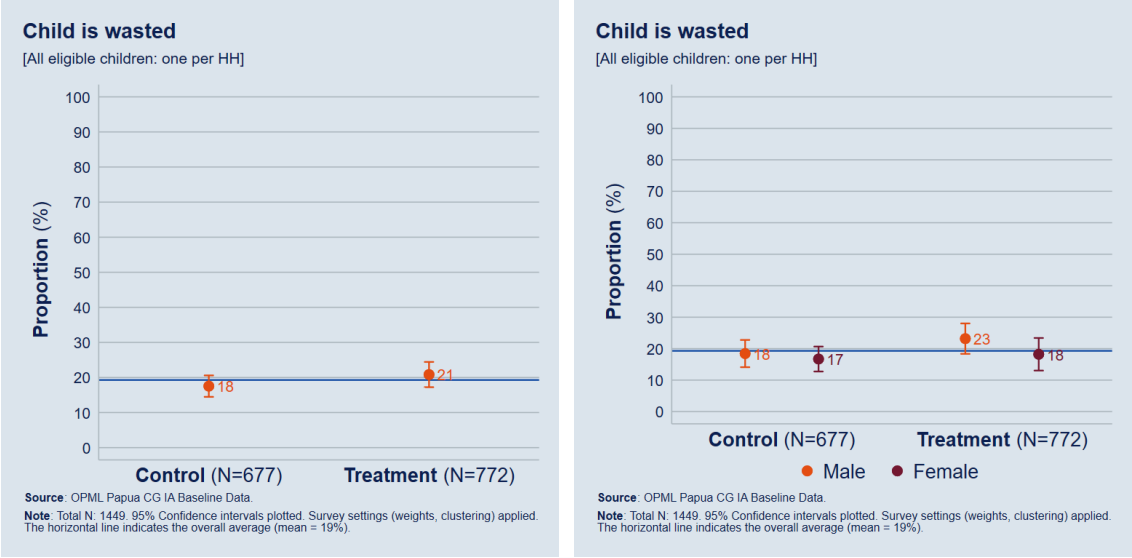


Table 4: Prevalence of SAM and MAM

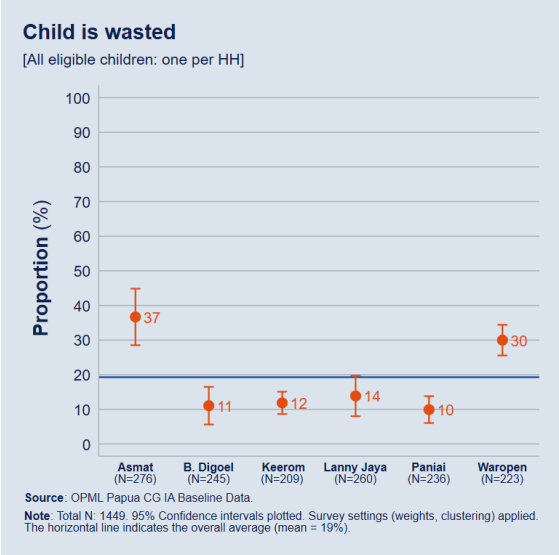
	Overall			Treatment			Control			Difference (T – C)
Proportion of children (%) who are …	N	%	95% CI	N	%	95% CI	N	%	95% CI	
MAM (moderate wasting)	1449	**12.7**	*[10.9-14.5]*	772	**12.6**	*[10.4-14.8]*	677	**12.8**	*[10.0-15.7]*	-0.2
SAM (severe wasting)	1449	**6.6**	*[5.2-7.9]*	772	**8.2**	*[6.1-10.4]*	677	**4.7**	*[3.3-6.2]*	3.5***
MAM or SAM (overall)	1449	**19.3**	*[16.9-21.7]*	772	**20.8**	*[17.2-24.4]*	677	**17.5**	*[14.5-20.6]*	3.3

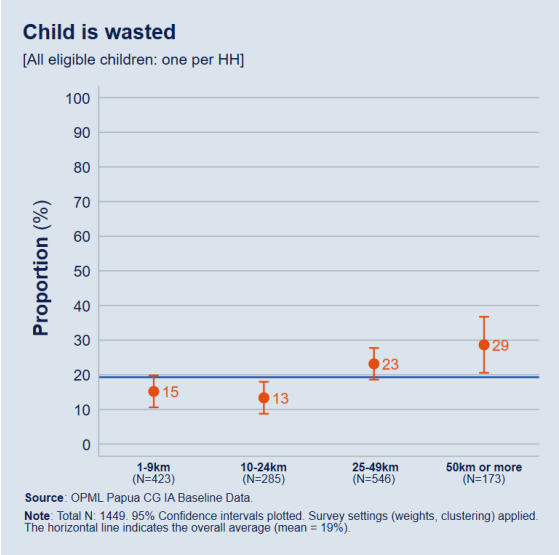
Source: OPML Papua CG Impact Analysis Baseline data

Note: SAM is measured as WLZ below -3 SD, while MAM is measured as WLZ below -2 SD but above -3 SD. 95% CI refers to the 95% confidence interval around the main estimate. Significance stars refer to p-values of * p<=0.1, ** p<=0.05 *** p<= 0.01 in a t-test that compares the estimates in the treatment with the control group.

Variability for wasting seems to be partially explained by geographic factors. We estimate that child wasting is more prevalent in Asmat and Waropen compared to other districts (Figure 29 left). In addition, being further away from the district capital increases the risk of wasting. Wasting prevalence in remote areas (50 km or more from the district capital) is twice as high as the prevalence observed within a radius of 10 km from the capital (Figure 29 right).

Figure 29: Prevalence of wasting by district (left) and by distance from the district capital (right)



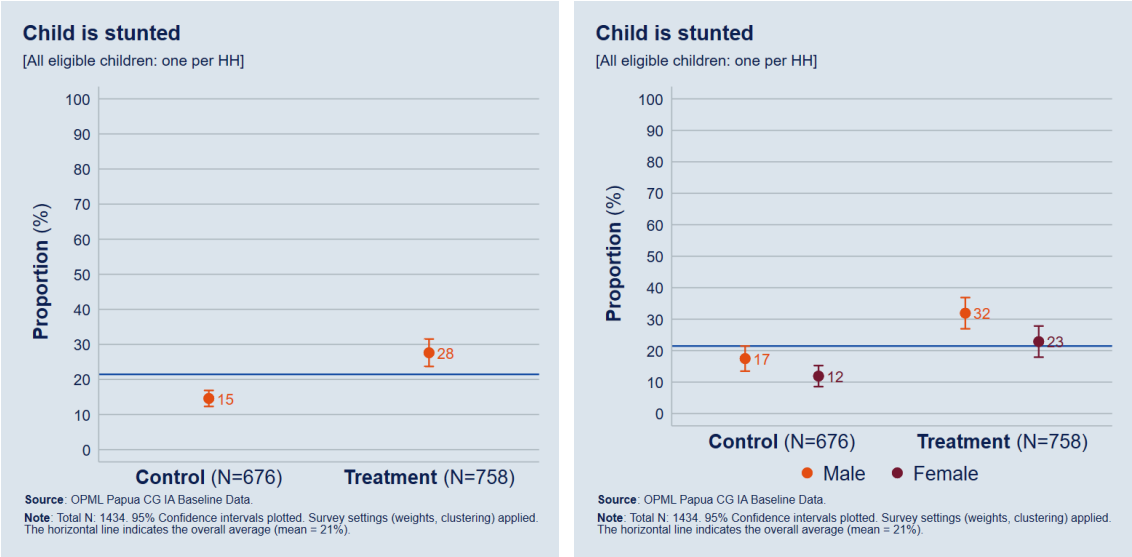


These estimates of wasting are very high and worrying. As described in the summary box at the beginning of this section, UNICEF and WHO estimate that children who are SAM have a significantly increased risk of death compared to children who are healthy (nine-fold compared to children with WLZ above -1 SD). In addition, as mentioned above, a prevalence of wasting above 15% (SAM or MAM) is considered to be critical and might require an emergency response. It is important to mention here that such high wasting prevalence might be seasonal, e.g. influenced by heavy rains that increase the risk of water-borne diseases. It is our objective to collect data at endline at the same time of the year in order to control for such seasonality as far as possible.

3.8.2 Prevalence of stunting

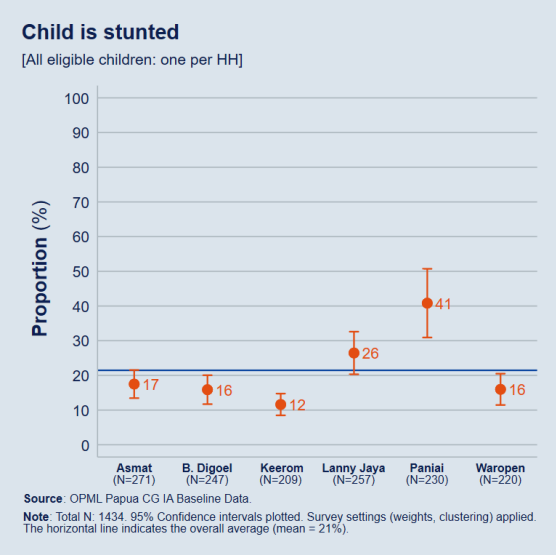
The graphs below present estimates of stunting prevalence in treatment and control areas. Overall, 21% of children are estimated to be stunted across those areas. However, stunting is significantly higher in the treatment group, where 28% of eligible children are stunted compared to 15% in the control group (Figure 30 left). On average, stunting prevalence is lower among girls in both groups and significantly so in the treatment group (Figure 30 right).

Figure 30: Prevalence of stunting (left), by gender (right)



Looking at disaggregation by district, the high stunting estimates seem to be driven primarily by Paniai and Lanny Jaya. The proportion of children stunted is around 40% in Paniai and 26% in Lanny Jaya (Figure 31). Given that Paniai and Lanny Jaya are both highland districts, this indicates that stunting might be significantly higher in the Papuan highlands compared to lower-lying districts. Note that this is the inverse of the relationship found with respect to wasting, where the estimated prevalence was particularly high in lowland districts (Waropen and Asmat; Figure 29 left).

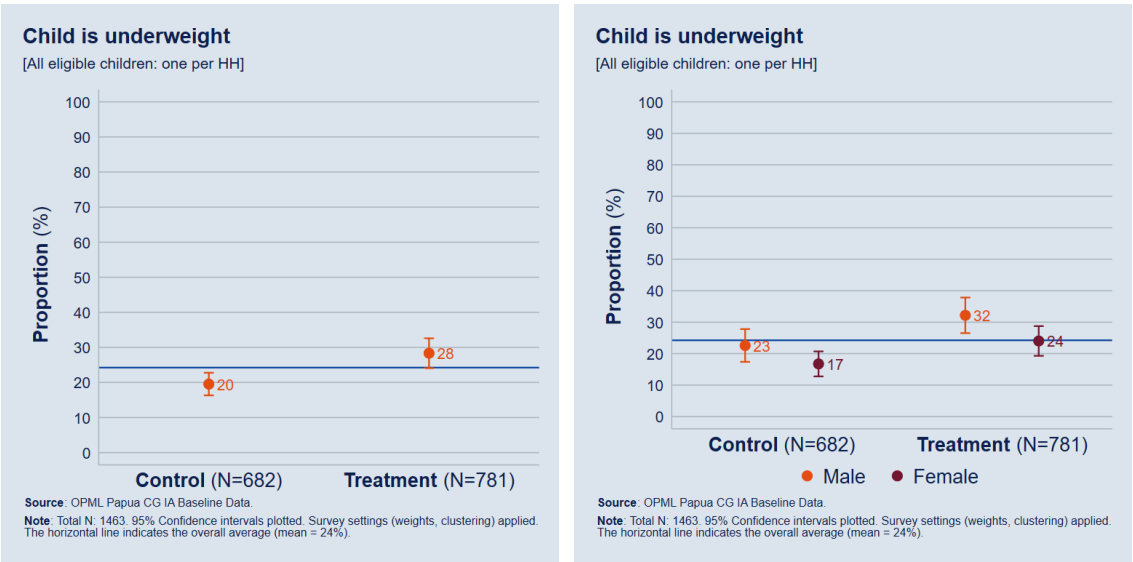
Figure 31: Prevalence of stunting by district



3.8.3 Underweight

As with stunting and wasting, the estimated prevalence of underweight among children in treatment and control areas is high according to the WHO prevalence cut-off values for public health significance.[24] About one-third (28%) of all children are underweight in the treatment group, compared to 20% in the control group (Figure 32 left). The risk of being underweight seems to be lower for girls, although differences by gender are not statistically significant (see Figure 32 right).

Figure 32: Prevalence of underweight (left), by gender (right)



[24] See www.who.int/nutrition/nlis_interpretation_guide.pdf

3.9 Preview of our PSM approach

3.9.1 Methodological summary

As indicated in section 3.3.2, at endline we will use PSM to estimate the impact of the CG in Papua. This means that full PSM modelling will only be implemented after endline data collection. This section serves as a preview of how this modelling will work.

The decision to use PSM for impact estimation purposes is mainly driven by the fact that PSM helps overcome selection bias by only comparing eligible households (children) to non-eligible households (children) who are observationally similar. Given that the CG can affect the characteristics of households and children at many different levels, this similarity must be established before the CG start, i.e. at baseline when no treatment effects could have materialised yet. Assessing the similarity of households and children involves assessing the 'covariate balance' at baseline. This means that, after applying PSM, households and children in the treatment and control groups are compared along a set of variables, employing statistical tests to assess whether they are statistically similar to each other or not. If they are, the conclusion is that matching worked well. Hence, assessing the balance of covariates at baseline between treatment and control groups after matching is a key step for PSM modelling.

PSM is a two-stage analytical approach that employs a propensity score as a 'comparator metric' that summarises the information provided by a series of characteristics for each unit of observation. The first stage of PSM is used to compute the propensity score for each unit of observation (households or children) using a set of covariates that represent the characteristics of interest. In the second stage, outcome indicators are compared across matched treatment and comparison groups to estimate treatment effects.

Although there are a variety of algorithms available to implement the second stage of PSM (i.e. to match comparison and treatment units to each other based on the propensity score estimated in the first stage), the results presented in section 3.9.2 correspond to using a Kernel matching approach with appropriate trimming and enforcement of common support. Enforcing common support means that observations (both in the treatment and control group) who are not deemed to be 'comparable enough' are dropped from the analysis.

Section 3.9.2 below provides an example of assessing the balance of covariates, after matching, for one particular outcome indicator: wasting of children aged 0–23 months.

3.9.2 A balancing example using PSM: prevalence of wasting

This section presents the results obtained from applying PSM to the Papua baseline survey data. In particular, PSM was applied to eligible child-level (0–23 months) data for the outcome indicator of wasting (see section 3.8 for a detailed description of this indicator).

Table 5 below displays the averages of covariates (means) estimated in the treatment and control groups before (unmatched) and after matching (matched). The table also shows whether estimates are significantly different from each other or not, using simple

significance stars that indicate significance thresholds after a t-test (see note below table).

The table demonstrates that, after matching, differences in the covariate means between the treatment and control groups disappear: there are no significance stars after matching. This means that after matching our sample is balanced across all the characteristics listed in Table 5 in the treatment and comparison groups. For example, in the unmatched sample, the average proportion of households where the head has attended school is 86% in the treatment group, compared to 96% in the control group. This difference is statistically significant at 1%. Once matching occurs, the proportion of households with a household head who has attended school is 90% in the treatment group while it is 91% in the comparison group, and the slight difference is not statistically significant.

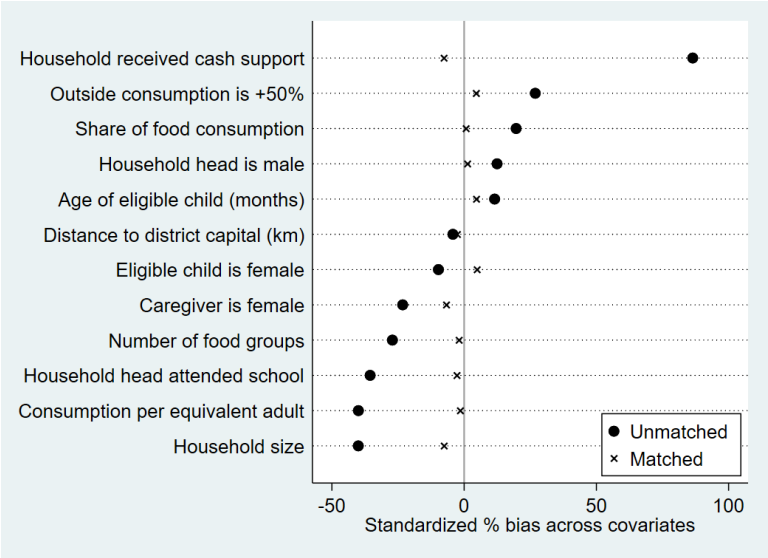
Table 5: Covariate balance in unmatched and matched samples

Covariate	Unmatched sample (mean)		Matched sample (mean)		N
	Treatment	Control	Treatment	Control	
Proportion of female children	48.0*	53.0	48.0	45.0	1,332
Age of children in months	12.1**	11.35	11.7	11.4	1,332
Proportion of children's caregivers who are female	93.0***	98.0	95.0	96.0	1,332
Number of food groups consumed by child in the past day	2.5***	3.3	2.6	2.7	1,332
Household size	4.8***	5.7	4.9	5.1	1,332
Proportion of households where head has attended school	86.0***	96.0	90.0	91.0	1,332
Proportion of households where head is male	97.0**	95.0	97.0	97.0	1,332
Distance of household to district capital	28.0	29.0	27.4	28.0	1,332
Household total consumption per equivalent adult	1,680,547***	2,128,113	1,722,076	1,738,429	1,332
Household's food share consumption	72.7***	70.2	72.6	72.5	1,332
Proportion of households that consume more than 50% of total consumption outside of the local economy	16.0***	7.0	13.0	12.0	1,332
Proportion of households that received cash support in the past 12 months	46.0***	10.0	40.0	43.0	1,332
Note: Significance stars correspond to p-values from a t-test: * = p<0.1, ** = p<0.05, *** = p<0.01.					

The same conclusion can be drawn from Figure 33 below, which visually shows how the household- and child-level covariates balance before and after matching. The x-

axis displays the standardised bias[25] for both the unmatched and matched samples. As can be seen, the unmatched sample displays large imbalances with standardised bias being present across many of the covariates. However, once matching takes place, the standardised imbalances considerably decrease, and the points are tightly distributed around the vertical line at 0% standardised bias.

Figure 33: Covariate bias in unmatched and matched samples



Our findings are corroborated by the statistical results shown in Table 6 below. While Table 5 and Figure 33 above demonstrated how each individual covariate balanced across treatment and control groups after matching, the indicators displayed in Table 6 provide information about the overall balancing properties of the unmatched and matched samples.

For interpretation purposes, it is important to highlight that a Rubin's B score under 25 and a Rubin's R score between 1 and 1.25 are the preferred values after matching, as they would indicate good matching performance overall (Rubin, 2001). As can be seen, our unmatched sample is unbalanced, but both indicators move to the ideal ranges after matching (Rubin's B goes from 128.26 to 15.97 and Rubin's R from 1.52 to 1.08). These values confirm how matching removes the previous imbalances in the set of covariates of interest.

Finally, Table 6 also presents the estimated Average Treatment Effect on the Treated (ATT) derived from this matching exercise. In this case, the estimate is close to zero, with a corresponding p-value of 0.607. Such a high p-value indicates, in line with the previous results, that there is no statistically significant difference in the proportion of wasted children between the treatment and control groups at baseline, once the matching process has taken place.

[25] Percentage difference of the sample means in the treated and non-treated (unmatched or matched) subsamples as a percentage of the square root of the average of the sample variances in the treated and non-treated group (Rosenbaum and Rubin, 1985).

Table 6: PSM overall balance indicators

	Before Matching	After matching
Rubin's B	128.26	15.97
Rubin's R	1.52	1.08
ATT		0.02
P-value		0.607
N on common support		1,332

In conclusion, the matching procedure based on PSM has achieved positive results. The analysis indicates that, after matching, the sample is balanced across the characteristics of interest in the treatment and comparison groups for the wasting impact indicator, which means that PSM will be a viable impact estimation strategy at endline.

3.10 Methodological caveats and implications

3.10.1 Methodological caveats

This section of the report has presented the key results of the baseline household survey conducted in Papua in the context of the quantitative impact assessment of BANGGA Papua. These results have been discussed and validated with UNICEF, BANGGA Papua, and other stakeholders during a validation visit that OPML conducted in Indonesia in May 2019. In light of these discussions, and the analyses presented above, some key caveats need to be taken into account when interpreting this study's findings:

- **Selection bias and contamination:** First, many of the analyses presented above show that the characteristics of households and children included in this study vary across treatment and control groups. Even if these differences might not be very large by magnitude, they are often statistically significant. For example, the proportion of children who are malnourished – a key outcome area in this evaluation – is significantly different between the two groups (section 3.8). These differences were to a large extent expected given that BANGGA Papua targeted the poorest districts in Papua.

  We have shown in section 3.9 that, for the purposes of the impact analysis at endline, we will be able to appropriately account for such differences, which might be driving selection bias, using PSM, assuming that we will be able to interview the same households and children again at endline (see the next bullet point for a discussion of this).

  It is important to note, however, that this does not solve the problem of contamination posed by the parallel and unbalanced implementation of other cash transfer programmes in treatment areas (see section 3.5). To account for this, detailed data collection about these programmes at endline (both in our survey and from secondary sources) will be required to ensure that we include up-to-date information on whether households did receive payments other than BANGGA Papua in our estimation models and to allow us to conduct sensitivity

analyses to assess how findings vary by whether we take this information into account or not. Ideally, households will be able to identify BANGGA Papua payments as separate from other support that they continued to receive between baseline and endline or that they unexpectedly received between those two points in time. We would then be able to conduct analyses that match households across treatment and control areas based on the level of support that they receive (other than BANGGA Papua) or identify households that did not receive any support other than the CG payments and compare those to households in the control areas that did not receive any payments. These analyses will allow us to make a qualitative statement about how likely it will be that findings at endline might be affected by this parallel implementation of other support programmes in treatment areas.

- **High levels of child attrition:** Second, our data on child malnutrition revealed a very high prevalence of wasting and stunting in some of the areas we surveyed (see section 3.8). In particular, the findings around wasting are very worrying, as they indicate that a large proportion of the children included in the survey are very sick and have a high likelihood of passing away soon. The evaluation team has informed all relevant stakeholders in Indonesia of this situation and its public health implications.

  In addition to these implications, it is important to note that these findings also have methodological consequences for our study. As mentioned in section 3.9, our impact assessment strategy currently assumes that we will interview the same households and children at endline again. Our findings on child malnutrition imply, however, that it is likely that a significant number of children interviewed at baseline might not be alive at endline (in 2020) any more, which could lead to high levels of attrition and – in particular – selective attrition in treatment areas where wasting is particularly high.

  We will use two strategies during the endline phase of this study to deal with this situation: first, together with UNICEF, the study team has discussed the implementation of a small tracking test in 2020 in order to try to assess the severity of this attrition. Second, if this tracking test reveals that attrition of children is indeed very high, we will need to switch to an evaluation strategy that involves collecting data from two cross-sections of children (i.e. not a longitudinal survey). Possible identification strategies in that case could involve a difference-in-differences (DID) approach or using PSM on a constructed pseudo panel, where children are matched over time as well as across treatment and control groups.

- **Difficulty seeing effects at endline:** Third, as mentioned in section 3.2, it is important to emphasise that BANGGA Papua's ToC includes some strong assumptions relating to how causal effects will materialise along first-, second-, and third-order outcomes. While we consider it possible that the cash transfer will affect household consumption and expenditure to some extent, it is unclear whether this will translate into changes in second- or third-order outcomes, given that such changes are often mediated by external factors and require changes in household behaviour that might not directly follow from the receipt of CGs.

  This issue is exacerbated by the logistical and operational difficulties that BANGGA Papua faces in Papua. These difficulties were emphasised by stakeholders during discussions with OPML in Papua and they mean that

reaching all potentially eligible households with cash payments over the period of this evaluation might be difficult. All these factors combined mean that it is possible that it will be difficult for treatment effects to materialise in a way that our evaluation can pick them up at endline.

- **Differences in consumption estimates compared to other sources:** Our results presented in section 3.6, in particular around the proportion of the population living below the Papua poverty line, differ from estimates found in other sources, such as SUSENAS 2018 reports. It is important to emphasise here that these differences can be explained, at least partly, by differences in our approach to measuring consumption levels, including with respect to the target population, surveyed areas, and differences in survey instruments. Nonetheless, our analyses and robustness checks conducted with this data indicate that our estimates are internally consistent and provide a measure of household consumption that will serve as a key indicator to assess impact on at endline (see Annex A.1 for details). Note that, in order to corroborate these findings further, we suggest collecting additional information on household welfare at endline below.

3.10.2 Implications of findings and discussions with stakeholders

Throughout this section, we present key findings related to each of the areas investigated in blue boxes at the beginning of each subsection. These key findings are the main conclusions and takeaways for this baseline study and will thus not be repeated here. Section 3.10.1 also presents methodological caveats that need to be taken into account when interpreting findings and their methodological implications for this study.

However, there are two key implications of our findings and discussions with stakeholders for the next phase of this evaluation that we would like to emphasise:

- First, stakeholders have highlighted the need to reassess the survey instrument used to collect data at baseline in order to include information on a few key additional dimensions of household welfare or to measure some existing dimensions in a different way.

  In particular, this derives from the fact that the consumption and poverty headcount estimates presented in section 3.6 are different from estimates on Papua obtained from other sources. Stakeholders therefore suggested trying to measure household wellbeing via other indicators, such as via a multidimensional poverty index or an asset index. In addition, suggestions were made with respect to measuring gender relationships within the household (e.g. relating to decision making) and assessing registration of family members not only via birth certificates but also other registration documents (e.g. family cards). OPML will review the survey instrument at endline to assess whether these changes are feasible and can be included in the endline survey.

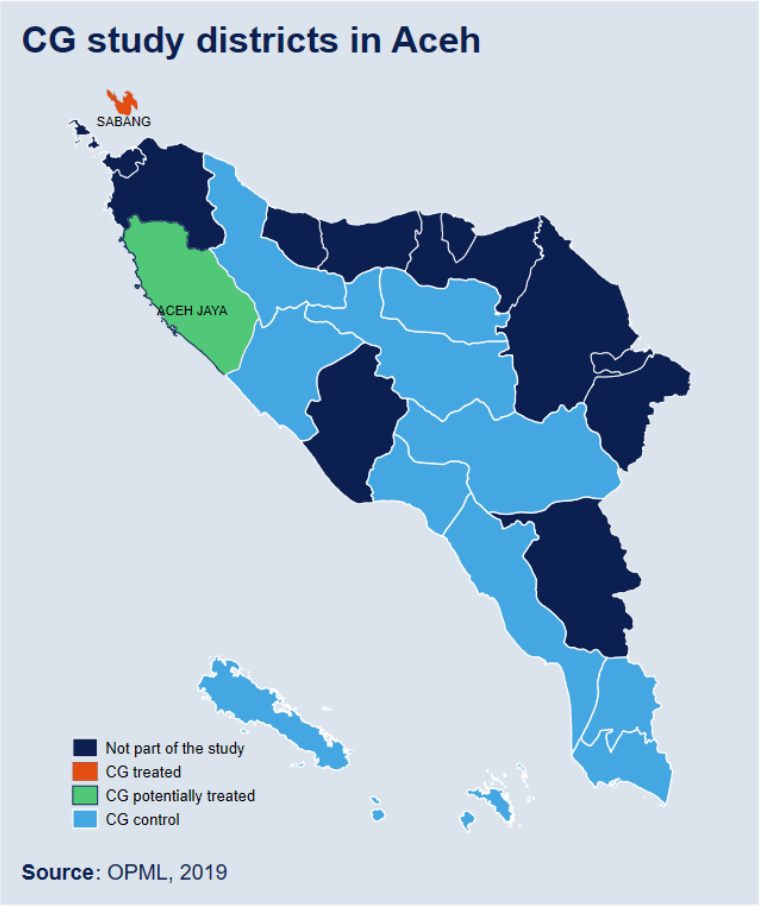
- Second, throughout the implementation of the baseline phase of this evaluation, OPML has communicated regularly with BPS in Jakarta. This communication was very helpful in both designing the survey and interpreting results. At endline, it will be important to continue collaborating closely with BPS in order to

ensure that this evaluation produces findings that are credible and useful for stakeholders in Indonesia and, in particular, Papua.

4 Aceh

4.1 A summary of our understanding of the intervention

Figure 34: Map of Aceh Province



In contrast to the CG programme in Papua, at the time of writing this baseline report (June 2019) grants have not yet been disbursed in Aceh. So far, the Government of Sabang – one district within Aceh - has allocated funding for the implementation of the CG in this district from July 2019 onwards. In Figure 34 above, Sabang is highlighted in orange. It is an island off the coast in North-West Aceh. It is likely that further districts will roll out the CG in the future, such as for example Aceh Jaya, which has been highlighted in green.

The CG in Aceh will unconditionally benefit all children aged 0–6 years inclusive, and the benefit level will be lower than in Papua. Sabang district expects to begin with IDR 150,000 (US$ 9.87) per child per month. The evaluation team's current understanding

is that payments will be made once a month into an individual bank account in the name of the child's main caregiver, who is normally the biological mother.[26]

It is important to highlight that the CG in Aceh will be part of a larger integrated model wherein the CG will be combined with other bundled interventions in Child Protection (positive parenting) and Child Survival (Health, Nutrition and WASH) at the *Posyandu*, *Puskesmas*, village, district, and province levels.[27] In the same manner, Sabang district already has in place an education grant programme, which provides cash annually for every school-aged child in the household. The fact that these interventions are being implemented concurrently with the CG has implications for how – at endline – we will be able to interpret findings from our impact analysis. In essence, if parallel implementation does indeed happen, it will be difficult to identify whether any effects might materialise due to the CG itself or to the bundle of interventions.

4.1.1 A light-touch ToC

As in Papua, no comprehensive ToC workshop with the implementing agency in Aceh has been conducted by OPML. This section very briefly summarises key aspects of our understanding of the implementation of the programme in Aceh, taken from the secondary information and documentation OPML has received so far. It is important to emphasise again that the CG in Aceh is planned to be implemented in an integrated manner with other programmes.

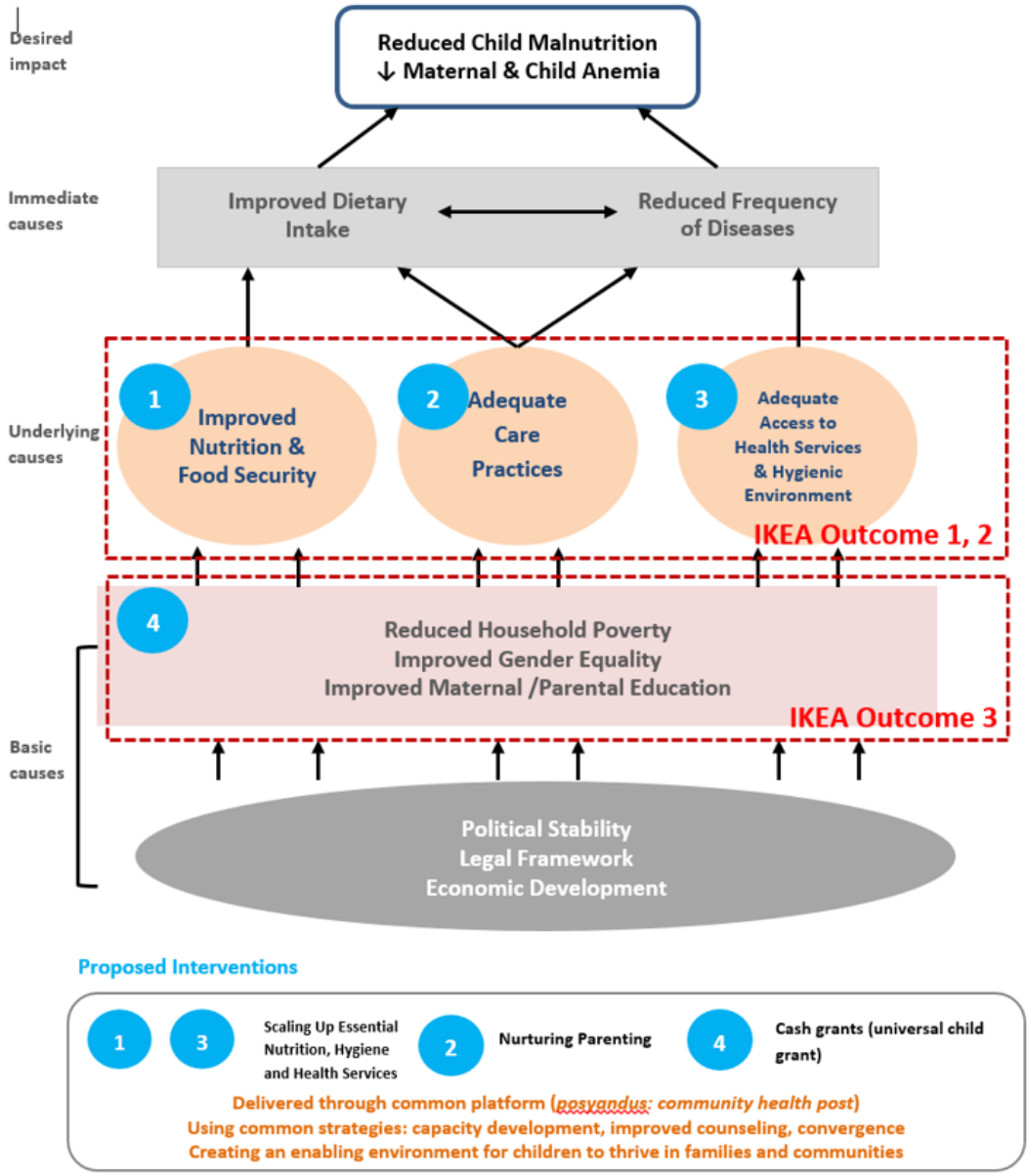
As depicted in Figure 35 below, the ultimate goal of the CG (together with the other integrated services in Aceh) is to tackle child malnutrition and maternal and child anaemia. In contrast to the Papua programme, in Aceh the CG payments are supposed to be made in conjunction with the implementation of other integrated services around child protection and child survival (health, nutrition, WASH). It is unclear to OPML how exactly this coordination will be implemented in practice. Importantly, however, the ToC below states that the final outcome of this programme can only be reached with these different interventions being properly integrated.

The ToC below indicates that the assumption is that by providing beneficiary households with a regular cash transfer, together with the integrated services, the increased household income is expected to trigger behavioural changes that lead them to increase their use of educational and health-related services, as well as to improve their food consumption habits. These changes will translate into an improved dietary food intake, which will eventually contribute to a reduced incidence of malnutrition and anaemia among children and mothers.

For the purposes of this impact evaluation, OPML will focus on the causal pathway mainly driven by the CG, which means assessing a more limited set of indicators. As in Papua, it is important to emphasise that this ToC rests on a series of strong assumptions that OPML unfortunately could not validate with the implementing agency and that bear risks for whether actual impact will materialise.

[26] The team is in conversations with UNICEF to get further information and confirmation on this.
[27] The exact modalities for this integration are not fully clear to the evaluation team at the time of writing.

Figure 35: Aceh intervention ToC



4.2 The focus of the impact assessment in Aceh

In contrast to Papua, the impact assessment in Aceh will mainly rely on secondary data collected by BPS Indonesia via its SUSENAS and RISKESDAS surveys in 2018 and via SUSENAS in 2020. This means that outcome indicators that could be covered in the context of this impact assessment need to be covered by those surveys.

According to OPML's information, in 2018 both the SUSENAS and RISKESDAS surveys in Aceh were implemented in the same households, which means that data from both surveys could, in theory, be linked and a combined analysis performed for the purposes of the baseline. RISKESDAS will, however, not be implemented in 2020.

It is important to emphasise here that the OPML team has not had access to the 2018 RISKESDAS data as at June 2019. It is our understanding from conversations

with UNICEF that, in fact, RISKESDAS data has not yet been shared with BPS, which implies that it will be unlikely that the evaluation team will be able to analyse this data in conjunction with SUSENAS data. In this baseline report, we therefore do not present any results related to indicators that would be constructed using RISKESDAS data.

For the endline stage of this project, this means that our ability to assess the impact of the CG on outcomes will depend on the structure of the SUSENAS surveys in 2018 and 2020. Currently, the SUSENAS questionnaire employed in Aceh in 2018 has been augmented by including components of UNICEF's MICS, in order to cover certain outcomes in the areas of early childhood development and parenting practices. Under the assumption that SUSENAS will be implemented in 2020 in a similar way, the evaluation team assumes that it will be able to assess impacts on the following outcomes:

1. Household consumption
2. Access to services – in particular education
3. Food consumption – at the household level
4. Birth registration
5. Parenting practices or early childhood development.

For this baseline report, we present estimates of indicators related to these areas of interest in sections 4.4 to 4.7.

4.3 Baseline data and analysis methods

4.3.1 Baseline data

In contrast to Papua, this evaluation in Aceh will not collect any primary data for the purposes of the main impact assessment. Instead, the objective is to rely – in as much as possible – on secondary data, collected primarily via SUSENAS and RISKESDAS in 2018 and via SUSENAS in 2020 to implement the impact assessment of the CG.

SUSENAS is the Indonesian National Socioeconomic Survey implemented yearly by BPS in Indonesia, covering a nationally representative sample of households. The survey covers a variety of different subject areas and – importantly – collects information on the outcomes 1 to 4 listed in section 4.2 above. For the March 2018 round of SUSENAS, which is the data that we are using here, and as mentioned above, BPS has included a module in the SUSENAS questionnaire employed in Aceh that augments the questionnaire with components of UNICEF's MICS surveys, covering outcomes related to early childhood development and parenting practices (outcome 5 listed in section 4.2).

SUSENAS data overall – and in Aceh in particular – is generally representative at the district level, but is not limited to households eligible for the CG only. Instead, data is collected from the full population of households. For this impact assessment, however, the subset of households eligible for the CG, i.e. eligible children aged 0–6 and eligible households with children of that age, is of particular interest. For the impact analysis and for this baseline analysis, we therefore limit our sample to that group of observations.

As described above, the original objective of the baseline analysis was to also analyse RISKESDAS data – in conjunction with SUSENAS data – to provide a description of health-related indicators of households eligible for the CG in Aceh. Given that RISKESDAS data has not been made available to the OPML team, this analysis is not included in this report.

The SUSENAS data shared with OPML covers a total of 12 districts in Aceh (see Table 7 below). These districts have also been highlighted in the map in Figure 34. It is important to emphasise here that it is currently not clear to the study team why those 11 districts were selected by BPS in the data shared with us. OPML is currently in conversations with UNICEF and other stakeholders to understand how far other programmes are being implemented in districts across Aceh, concurrently with this evaluation. For the endline analysis, the objective will then be to select a core set of comparison districts to Sabang in which it will be reasonable to assume that no other programmes might affect the robustness of our analysis.

Sabang is the one treatment district where CG payments so far have been confirmed. It is highlighted in grey in the table below. Overall, the data shared with OPML covers a total of 5,502 households. Of those, a total of 2,245 have children aged 0–6 years. In Sabang, a total of 118 households include children aged 0–6 years. These households are considered to be eligible for the CG in Sabang. **It is those households and their children that this baseline analysis focuses on, comparing them to all other children in the eligible age range and their households in the remaining 11 districts for which we have data in Aceh**.

Table 7: SUSENAS sample composition of data available to OPML in Aceh

District Name	Households with children aged 0–6	All households
Simeulue	167	430
Aceh Singkil	192	429
Aceh Selatan	199	514
Aceh Tengah	218	504
Aceh Barat	186	510
Pidie	230	629
Aceh Barat Daya	202	472
Gayo Lues	188	434
Aceh Jaya	166	433
Bener Meriah	191	474
Sabang	118	275
Subulussalam	188	398
Total	2,245	5,502

4.3.2 Baseline analysis

As in Papua, the objective of the baseline analysis in Aceh is to provide a description of households eligible for the CG in the treatment districts (Sabang) and to compare this to households in other districts. In the following paragraphs, we therefore compare the following groups of households:

- Households with children aged 0–6 years in Sabang. **(These are called the 'Treatment' group in our graphs.)**

- Households with children aged 0–6 years in the remaining 11 districts in Aceh. **(These are called the 'Control' group in our graphs.)**

This will allow us to assess, for example, whether eligible households that are due to receive the CG in Sabang are on average poorer than similar households in the other districts in Aceh or whether eligible households receive cash support differentially.

In this report, we present simple descriptive statistics (see section 2). The SUSENAS sample is drawn using clustered sampling, which means that in a first stage PSUs are drawn by BPS and then – in a second stage – households sampled within those PSUs. When producing these descriptive statistics, we take this sampling strategy into account, include sampling weights provided by BPS, and adjust standard errors for clustering.

It should be noted that this simple baseline analysis will be superseded by a more complex econometric analysis at endline, where we will aim at estimating what the effect of the CG was on key outcome indicators. The objective is to use a DID set-up to identify this effect. This baseline analysis provides us with some information on how such a DID analysis could be implemented. We present caveats around this in section 4.8.

4.4 Demographics and background characteristics

Box 7: Key findings on demographics and background characteristics

- When looking at key demographic information and household composition, **households in Sabang eligible to receive the CG are similar to households in the remainder of the sample in Aceh**, except with respect to one key dimension: households in Sabang are more likely to have male household heads.
- The same holds true when looking at a few key demographic indicators for children: **on average, the two evaluation groups are similar to each other.** One exception is that children in Sabang are more likely to live in a household where the household head is the biological or step-parent of the child.

The following sections present key demographic indicators and background characteristics of households eligible to receive CGs in Sabang, households with children aged 0–6 years in the remaining sample in Aceh, and children living in those households.

4.4.1 Households

Overall, households in Sabang are similar to households in the rest of the sample when looking at household size, age dependency ratio, education of household heads, and the gender of the household head. On average, there are about five members living in households in Sabang and the rest of the sample. The age dependency ratio is lower than 100%, implying that households have more than one working age member per dependant member (i.e. children and old people). In fact,

the dependency ratio is at 97% in treatment areas and 92% in the comparison districts. The majority of household heads are men. This proportion is higher in Sabang, however, than in the remaining districts in Aceh: 96% in Sabang compared to 91% in the remaining sample (Figure 36 and Figure 37).

Figure 36: Average household size (left) and age dependency ratio (right)

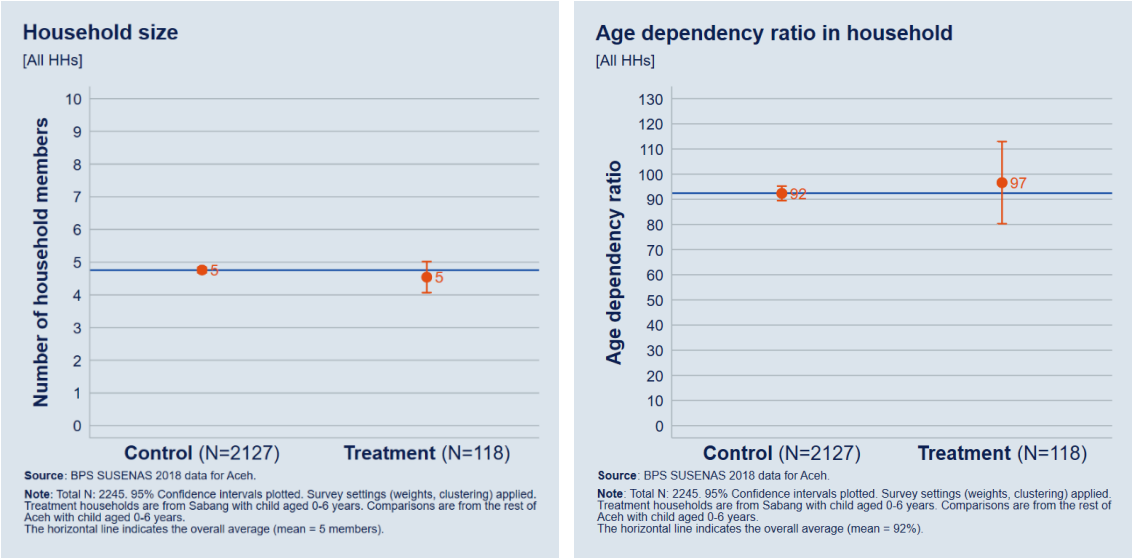
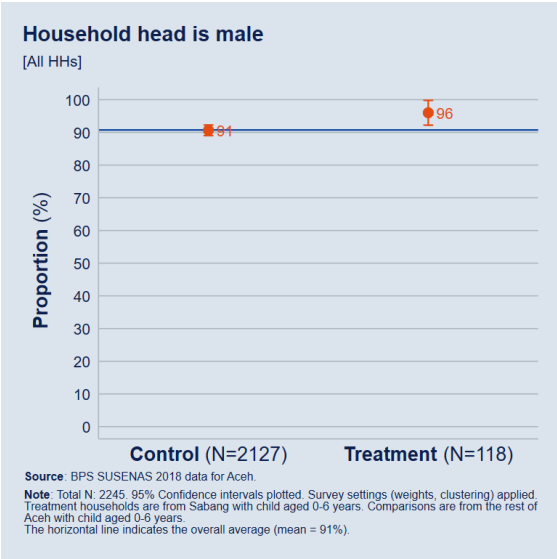


Figure 37: Proportion of male-headed households



4.4.2 Children

Children eligible to receive the CG in Sabang and of similar age in the rest of Aceh are 30 months old on average (Figure 38 left). As could be expected, about 50% are girls. Estimates are similar across evaluation districts (Figure 38 right). The large majority of children are the biological children or stepchildren of the household head, regardless of the child's gender. This percentage is significantly higher in Sabang, where it is around 90% (Figure 39).

Figure 38: Average age of eligible children in months (left) and gender of the eligible child (right)

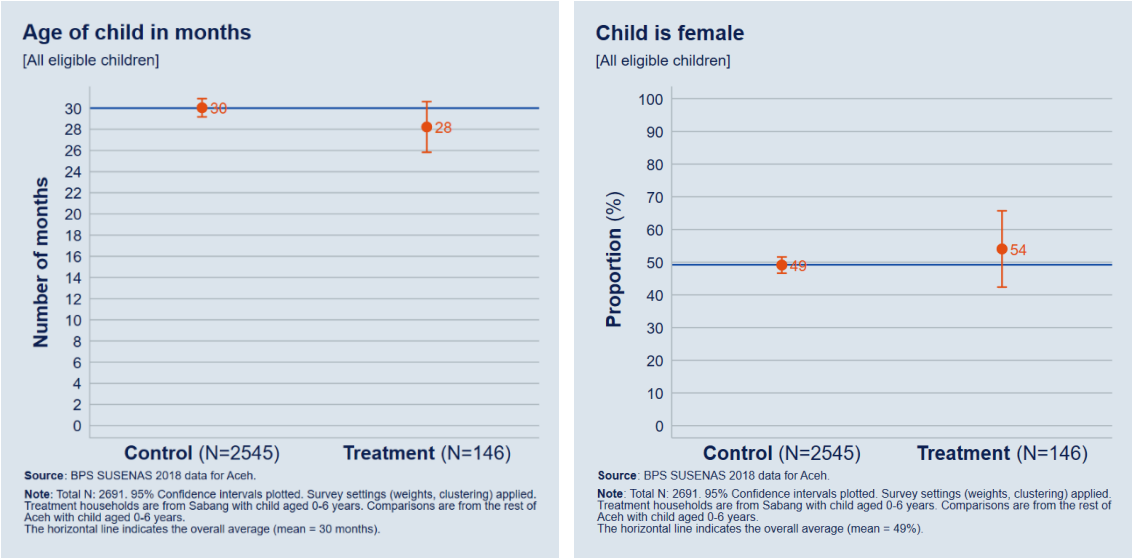
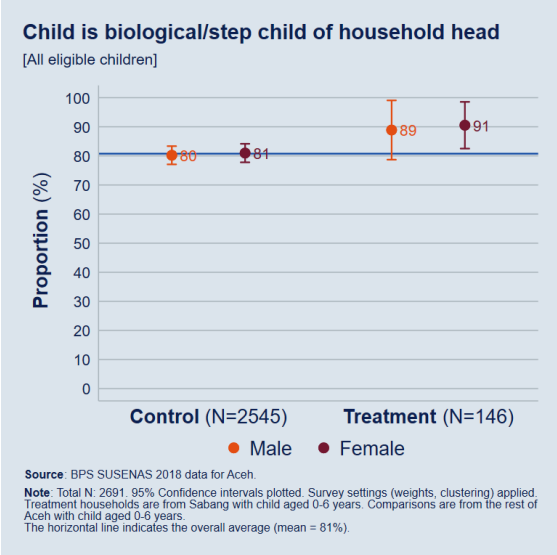


Figure 39: Proportion of biological children or stepchildren



4.5 Exposure to other support programmes

Box 8: Key findings on exposure of households to financial support from other programmes

- **The large majority of households in Sabang are beneficiaries of local government social assistance programmes.** Only 8% of households receive financial support in the remaining districts in Aceh.
- When it comes to nationwide social protection policies, 4% of households in Sabang receive PKH support. The coverage is significantly lower than in the remainder of Aceh. Rastra/Raskin reaches about 56% of all households in the sample, with no significant differences across evaluation groups.

The SUSENAS questionnaire also includes a module where households are asked about financial support from local governments or other support from national support programmes. The below figures presents the results for three of those: support received from local government, PKH support, and whether households were part of Rastra – the rice subsidy programme. Note that recall periods for these indicators vary.

Figure 40 below presents the proportion of households that received support from local government in the year prior to the survey, showing that 72% of households in Sabang have received such financial support compared to just 8% in the remaining districts in Aceh. This difference is highly significant and it corroborates findings produced by Empatika (Jupp *et al*., 2018) that also identified the high prevalence of local support grants in Sabang.

On the other hand, PKH support is not widespread in Sabang, with only 4% of households in Sabang reporting ever having been beneficiaries of this conditional cash transfer programme. Again, this proportion is significantly different in the remainder of Aceh, where on average about 19% of households report having benefitted from PKH (Figure 41 left). Finally, about 56% of all households in the sample received Rastra/Raskin in the four months prior to the survey, with no significant differences across evaluation groups (Figure 41 right).

Figure 40: Proportion of households that received financial support from the local government

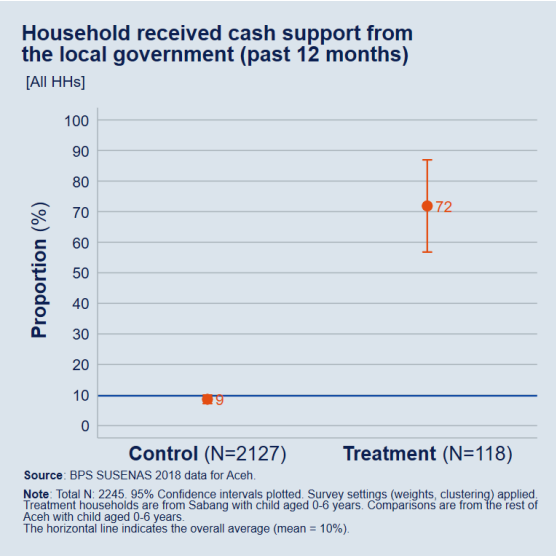
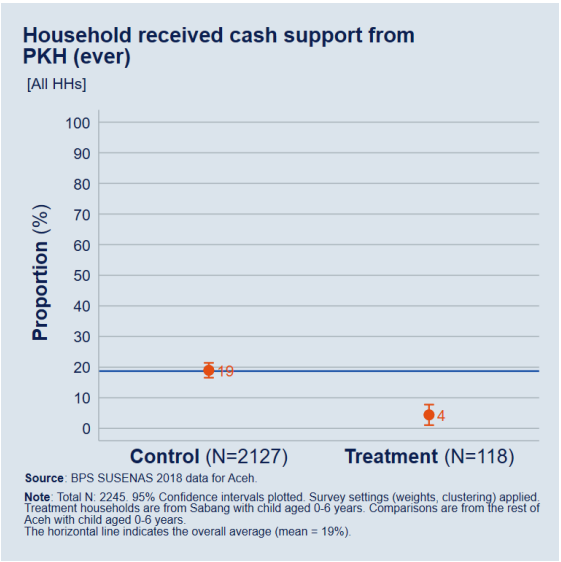
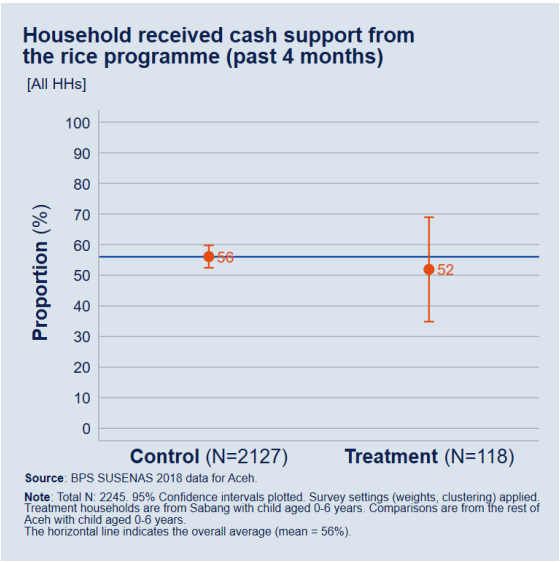


Figure 41: Proportion of household receiving PKH (left) and Rastra/Raskin (right)





These significant differences in the proportion of households receiving local government support or PKH across Sabang and the comparison group pose a difficulty to the impact evaluation strategy adopted in this study. As briefly mentioned in section 4.3.2, this strategy aims at using a DID approach to estimate the effect of the CG on key outcome indicators. Using DID, in turn, implies relying on the parallel trend assumption, which asserts that – without the programme to be evaluated – outcome indicators would develop similarly across the treatment and comparison groups. The existence of differential social support programmes, however, means that it is unlikely that this assumption would hold. This is particularly the case for local government support, where the proportion is much higher in Sabang. At endline, we will try to address this by appropriately controlling for these alternative payments at the household level, as long as they are distinguishable using SUSENAS data.

4.6 First-order outcomes: consumption estimates

Box 9: Key findings on consumption levels

- **Households with children aged 0–6 years in Sabang consume around IDR 1.5 million per adult equivalent per month**. This corresponds to around IDR 5 million per month at the household level. Consumption levels are similar in the rest of Aceh.
- **Households that are due to receive the CG in Sabang consume proportionally less food than similar households in the rest of Aceh.**
- **Household monthly consumption is higher in households with children aged 0–6 years than other households**, with no significant differences between Sabang and the control districts.

As described in section 4.3.1, the consumption analysis in Aceh relies on secondary data collected by BPS Indonesia via its SUSENAS survey in March 2018. In December 2018, BPS shared the SUSENAS data for Aceh with OPML.[28] The dataset includes food consumption aggregates, non-food consumption aggregates, and *per capita* expenditure by household. It also includes information on nutrient intake and indicators for the survey weights, at both household and population level.

The set of indicators we present for Aceh in this section is a subset of the ones covered by the consumption analysis in Papua (section 3.6). This is because we limit our analysis to the secondary data shared by BPS, which do not include consumption expenditure by individual food and non-food items.

In this section, we present results with respect to the following indicators:

- Monthly total consumption expenditure per adult equivalent, as per the OECD equivalence scale;
- Monthly food consumption per adult equivalent, as per the OECD equivalence scale;
- Monthly total consumption and food consumption at the household level; and
- Food share of monthly consumption expenditure.

Consumption indicators have been constructed in real terms, to adjust for geographic differences in prices and to allow comparison across districts. Given that we do not have access to information on the unit prices of food and non-food items, we adjust for price differences using the implicit price index derived by the ratio of each district poverty line and the Aceh poverty line, as estimated by BPS in March 2018.[29]

[28] Block 43 data from SUSENAS.
[29] IDR 464,626 *per capita* per month. Details of the food poverty line can be found at www.bps.go.id/dynamictable/2016/01/18/1123/garis-kemiskinan-makanan-gkm-menurut-provinsi-2015---2018.html and the non-food poverty line at www.bps.go.id/dynamictable/2016/01/18/1124/garis-kemiskinan-non-makanan-gknm-menurut-provinsi-2015---2018.html

Note that the consumption analysis in Aceh does not include estimates of the proportion of households or individuals living under the poverty line. This is because it is not clear to the study team how BPS district poverty lines would need to be adjusted in order to present estimates for Sabang and the remainder of Aceh separately. The evaluation team is currently discussing this with BPS, aiming to produce such estimates at endline.

Figure 42 displays monthly consumption per adult equivalent and by household. Households with children aged 0–6 years in Sabang consume around IDR 1.5 million per month and per adult equivalent. This corresponds to around IDR 5 million overall per month at the household level. This level of consumption does not vary between Sabang (treatment) and the rest of Aceh (control).

Figure 42: Monthly consumption per adult equivalent (left) and total monthly household consumption (right)

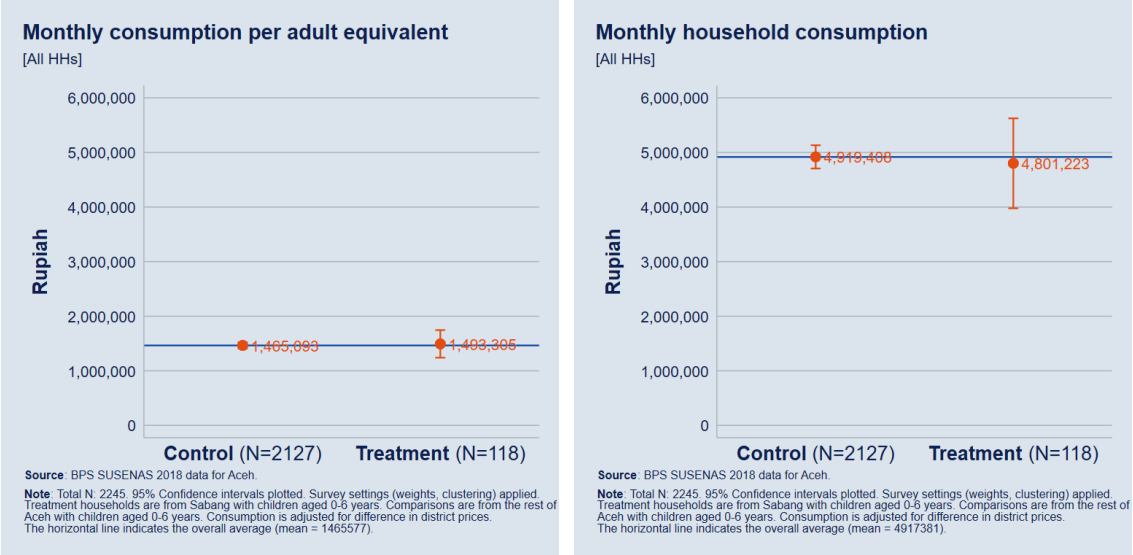


Figure 43 presents consumption levels for the whole sample, i.e. not just for households with children aged 0–6 years but for all households. Monthly consumption per adult equivalent and total monthly household consumption are similar to the restricted sample presented above. Monthly consumption per adult equivalent is somewhat higher in the full sample, while total household consumption is slightly smaller.

Figure 43: Monthly consumption per adult equivalent (left) and monthly household consumption (right), full sample

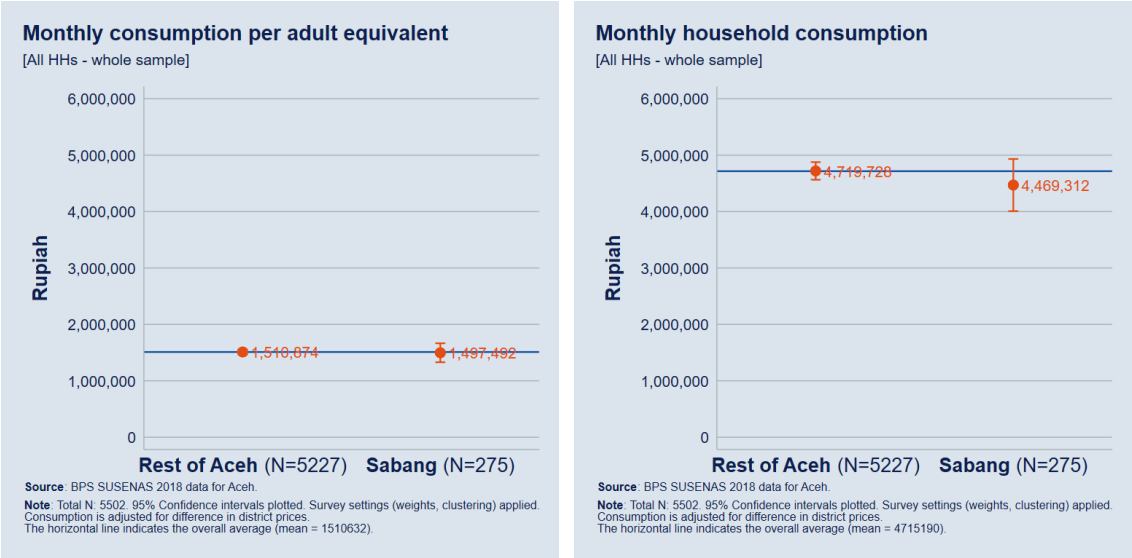
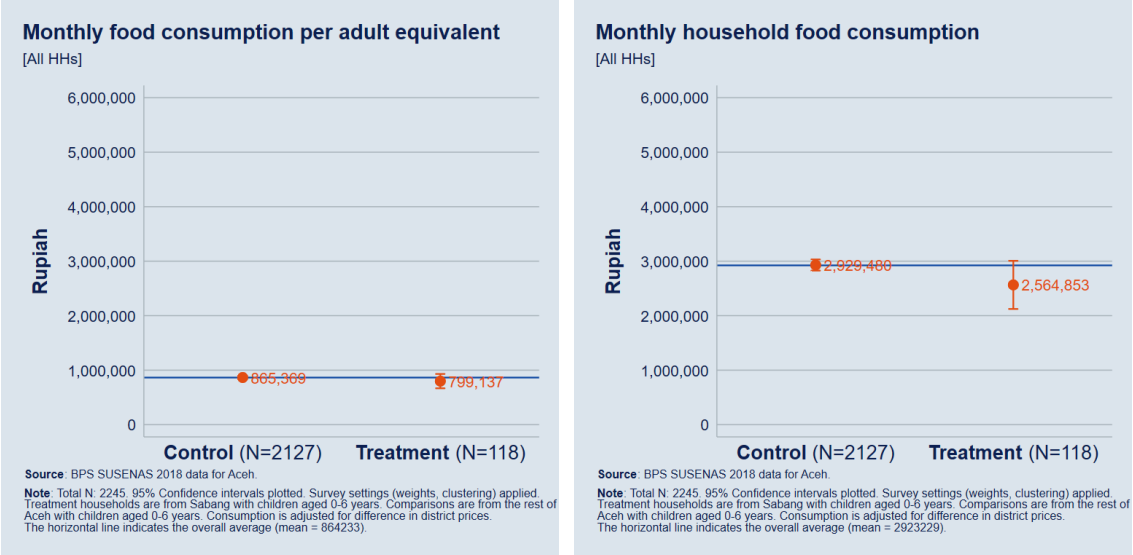


Figure 44 shows that, on average, monthly food consumption per adult equivalent is around IDR 860,000 among households with children aged 0–6 years. Total household food consumption is slightly less than IDR 3 million on average. In Sabang, total household food consumption is lower than in the control group, although the difference is not statistically significant (Figure 44 right).

Figure 44: Monthly food consumption per adult equivalent (left) and monthly household food consumption (right)



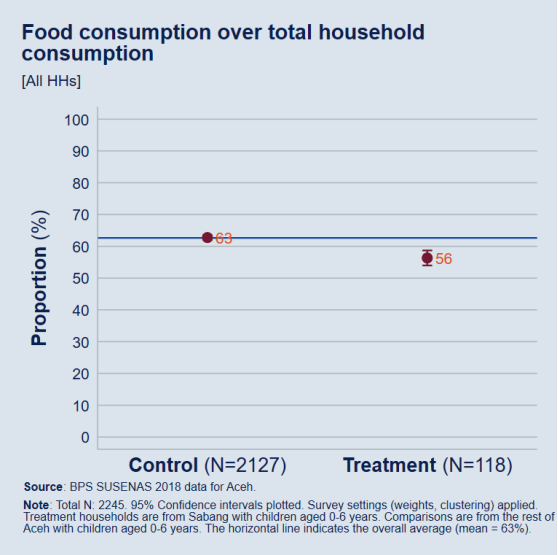
In households with children aged 0–6 years, food consumption makes up about 56% of overall household consumption expenditure in Sabang and 63% in the rest of the districts (Figure 45). This difference is statistically significant, which means that households that are due to receive the CG in Sabang consume proportionally less food than similar households in the rest of Aceh.

This may reflect different consumption patterns between households eligible for the CG in Sabang and the same households in the rest of Aceh. Previous qualitative findings

(Jupp *et al*., 2018) confirmed that households in Sabang are more likely to pay rent than households in the rest of Aceh. Rent expenses would account for a proportion of household overall consumption, potentially reducing the household food share.

In addition, as mentioned above, the similar level of consumption per adult equivalent between treatment and comparison households indicates that households eligible for the CG in Sabang have fewer adult household members (above 18 years old) or fewer children compared to the rest of Aceh. This would explain different food needs between eligible households in Sabang and similar households in the rest of Aceh.

Figure 45: Food consumption as a share of total household consumption



4.7 Second-order outcomes: education, birth registration, early childhood development, and childcare

Box 10: Key findings on child education, birth registration, early childhood development, and childcare

- **Pre-school attendance is low**, with only 10% of children aged 0–4 years being enrolled in pre-school, regardless of the evaluation group. Conversely, on average 80% of children aged 5–17 attend school.
- **Overall, 67% of children have a birth certificate**. There is no significant difference in this proportion across comparison groups nor between genders.
- **Availability of books is significantly higher in Sabang than in the comparison group**, with 34% of children aged 0 to 4 years old having three or more books. The percentage is at 40% among girls, which makes them the group most likely to have three or more books available to them.
- **In contrast, girls in Sabang are least likely to have access to toys.** On average 67% of children have access to two or more toys, while this proportion is at 50% among girls in Sabang.
- **In Sabang, early child supervision seems to be less problematic than in other districts in Aceh**. Only around 6% of children have been left alone for an hour or more in the week prior to the survey.
- **About 52% of children have experienced violence in the month prior to the survey**. We do not observe any significant difference across Sabang and the remainder of Aceh or between male and female respondents.

As in section 3.7 this section presents baseline findings for second-order outcome indicators, i.e. those outcomes that might be indirectly affected by the increased income derived from the CG.

Table 8 presents the indicators this analysis reports on. It should be noted here that this is limited by the availability of data to the study team and that indicators listed in this table are the ones covered by available SUSENAS data.

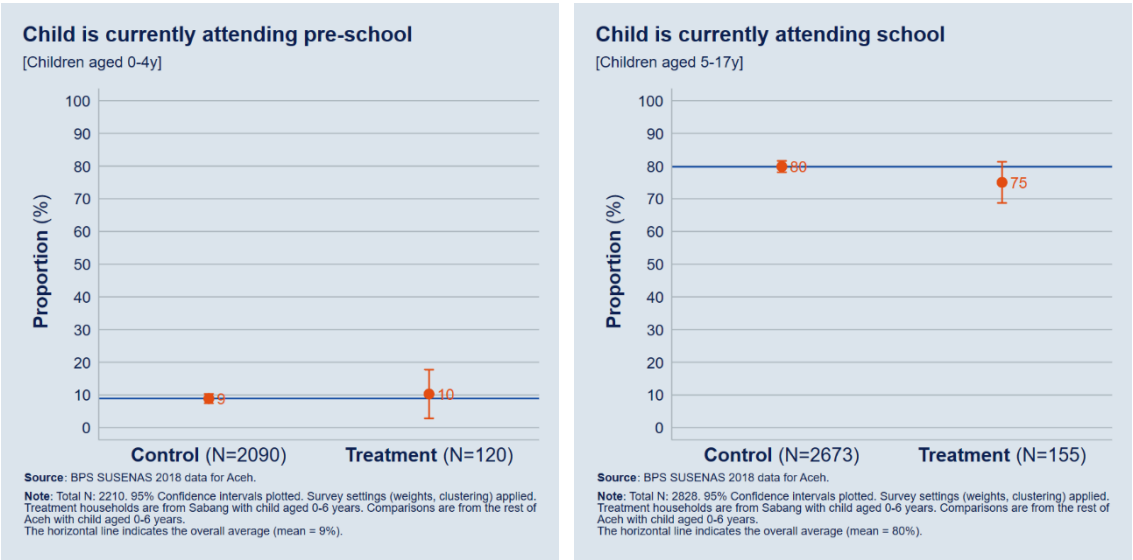
Table 8: Second-order outcome indicators

Dimension	Indicator
Access to education	• Percentage of children currently attending / ever been to pre-school • Percentage of children attending school
Birth registration	• Percentage of children under age 17 who have a birth certificate from the civil registry office
Early childhood development and childcare	• Availability of children's books: Percentage of children under age four who have three or more children's books available in their household • Availability of playthings: Percentage of children under age four who have two or more types of plaything available in their household • Inadequate supervision: Percentage of children under age four left alone or under the supervision of another child younger than 10 years for more than one hour at least once in the last week • Violent discipline: Percentage of children under age four who experienced any violence by caregivers in the past one month

4.7.1 Access to education

In the two graphs below, we present the proportion of children who are estimated to be in pre-school and school. The left-hand panel in Figure 46 shows that about 9% of all children aged 0–4 years are currently attending pre-school. There is no difference between the treatment and control sample. The right-hand panel, in contrast, shows that about 80% of all children aged 5–17 years are currently attending school in our study sample. This proportion is 5% lower among households eligible for the CG in Sabang – although this difference is not statistically significant.

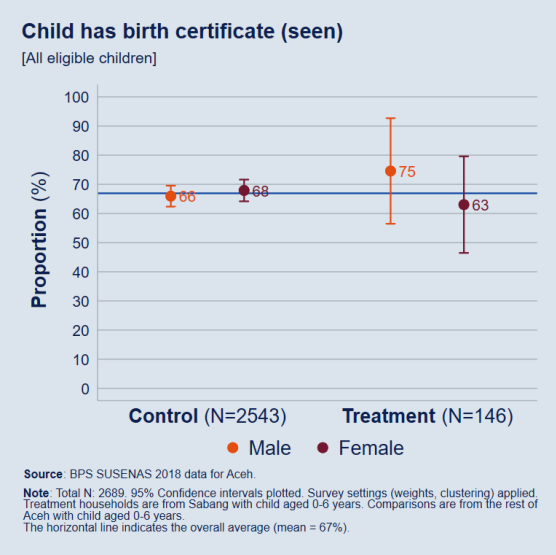
Figure 46: Proportion of children aged 0–4 years currently attending pre-school (left) and proportion of children aged 5–17 years currently attending school (right)



4.7.2 Birth registration

Figure 47 below plots the proportion of children aged 0–6 who have a birth certificate from a civil registry office that was also seen by the enumerator, by comparison group and gender. Overall, 67% of children have a birth certificate. There is no significant difference in this proportion across comparison groups nor between genders. This indicates that in Sabang and the rest of Aceh birth registration is fairly common and it happens to a much higher extent than, for example, in Papua (see section 3.7.2).

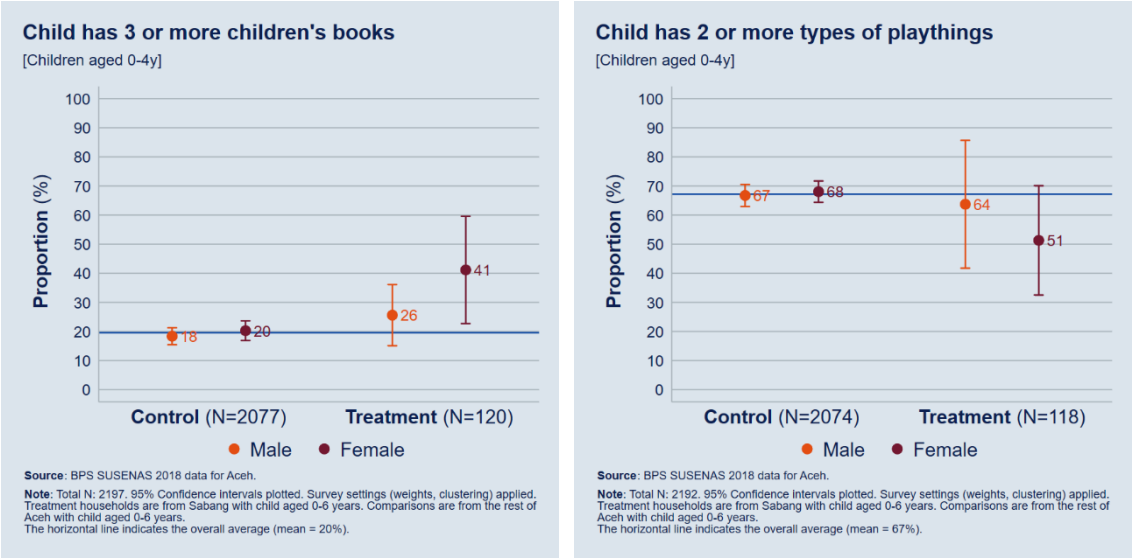
Figure 47: Proportion of children aged 0–6 years who have a birth certificate



4.7.3 Early childhood development and childcare

The left-hand panel in Figure 48 below presents the proportion of children who have three or more children's books available, by gender. Overall, this proportion is at about 20%. In Sabang, the average is significantly higher at about 34% overall (not shown in the graph). This difference seems to be driven mainly by the fact that this proportion is particularly high among female children (40%) in Sabang. The right-hand panel in Figure 48 presents the proportion of children aged 0–4 years who have two or more playthings available, by gender. In contrast to the findings on books, this analysis seems to indicate that girls in Sabang are least likely to have access to playthings. On average, 67% of children have access to two or more playthings, while this proportion is at 51% among girls in Sabang.

Figure 48: Proportion of children aged 0–4 years who have three or more books (left) and who have two or more types of plaything (right)



The proportion of children aged 0–4 years who have been left alone on any day in the week prior to the survey for one hour or more is presented in the left-hand panel in Figure 49 below. The indicator is disaggregated by comparison group and gender. Overall, we find that about 15% of children have been left alone for an hour or more in the week prior to the survey. This proportion is significantly lower in the treatment area – i.e. Sabang – and even more so for girls. This means that in Sabang early child supervision seems to be less problematic than in other districts in Aceh for the group of households with children aged 0–6 years.

The right-hand panel of Figure 49 presents the proportion of children aged 1–4 years who have experienced violence in the month prior to the survey. Overall, this proportion stands at 52% and we do not observe any significant difference across Sabang and the remainder of Aceh or between male and female children.

Figure 50 shows the use of violence toward children aged 1–4 years, by type of violent behaviour. Out of those children experiencing any type of violence, 16% of children aged 1–4 years experienced psychological violence, 36% have been physically punished, and 48% suffered both types of violence in Sabang.[30] The breakdown is somehow different in the remainder of Aceh. Children aged 1–4 years in the rest of Aceh districts are more likely to experience psychological violence and less likely to experience physical violence than the same cohort of children in Sabang.

[30] Physical punishments include: shaking the child; spanking, hitting or slapping the child on the bottom with the bare hand; hitting the child on the bottom or elsewhere on the body with something like a belt, hairbrush, stick or other hard object; hitting the child on the face, head, or ears; hitting or slapping the child on the hand, arm, or leg; beating the child up, i.e. hitting the child over and over as hard as one could. Psychological aggression includes: shouting, yelling, or screaming at the child; calling the child dumb, lazy, or another name like that (UNICEF, 2019).

Figure 49: Proportion of children aged 0–4 years who have been left alone in the past week (left) and proportion of children aged 1–4 years who have experienced violence in the past month (right)

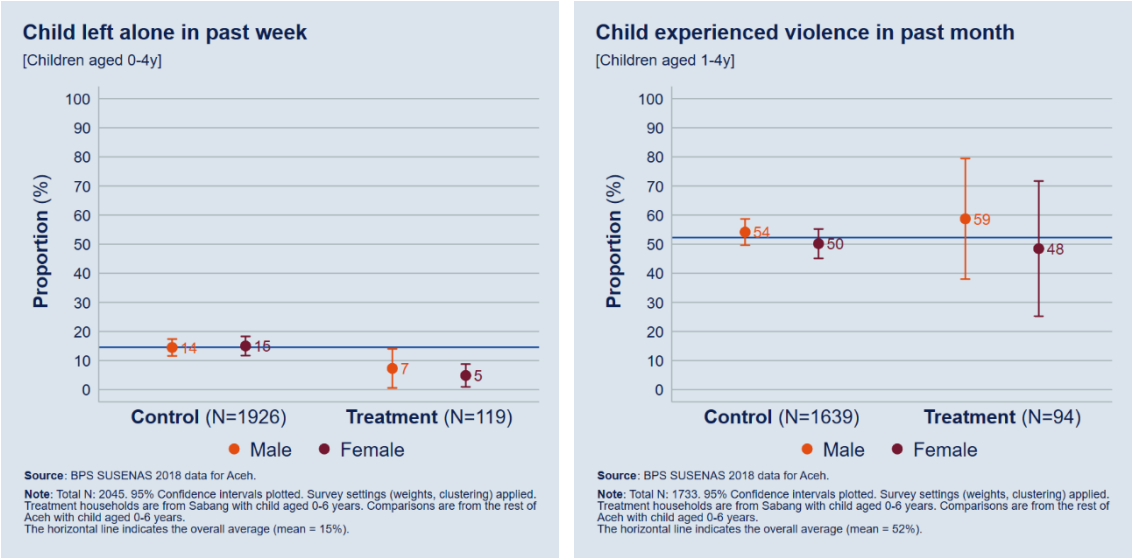
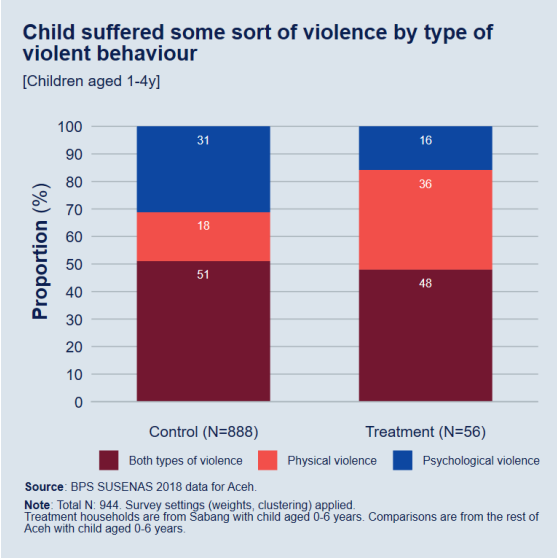


Figure 50: Proportion of children aged 1–4 years who have experienced violence in the past month, by type of violent behaviour



4.8 Methodological caveats and implications

This section of the report has presented the characteristics of the population of households and children in Sabang who will likely be beneficiaries of the CG and compared these to the characteristics of a similar group of households and children in the rest of Aceh, using SUSENAS data from March 2018. The intended evaluation strategy to assess whether the CG will have had an effect on the lives and wellbeing of households and children in Sabang is to use a DID estimation. This means that we would compare measurements in the treatment group (Sabang) with measurements in the comparison group (the other districts in our data) between baseline (SUSENAS March 2018) and endline (SUSENAS from March 2020). The underlying assumption

for this to work is that – in the absence of the CG programme – trends in measurements between the two groups would be the same (i.e. the 'parallel trend assumption'). In essence, DID will assess whether there is a difference in how indicators moved over time between the two treatment groups. The analysis implemented for the purposes of this baseline report has, however, raised a set of important methodological caveats that need to be mentioned here:

- **Limited availability of data:** First, section 4.1 shows that the CG in Aceh is assumed to affect a large set of dimensions of households' and children's lives. To assess whether this is indeed the case, the evaluation team would require access to data that covers indicators across those dimensions. Unfortunately, during this baseline phase, no data on nutrition and malnutrition indicators (from RISKESDAS) was shared with the evaluation team, preventing any related analysis. It is likely that the evaluation team will be in a similar situation at endline, which means that we will only be able to assess impact on a limited set of indicators, as presented in this report. Note that any analysis at endline will depend on SUSENAS data being made available to the team.

- **Small sample size in the treatment group:** This evaluation was originally designed under the assumption that the CG will be implemented in four districts in Aceh. The objective was to identify average treatment effects across those four districts. Currently, however, the only confirmed district in which the CG will be implemented is Sabang. As can be seen in the results presented above, this has implications for the sample size of 'treatment' households and children, which is at around 120 for key household-level indicators in this round of SUSENAS. This low sample size in turn implies that, for some indicators, the level of uncertainty around estimates is high (which can be seen in large confidence intervals) and that it might therefore be difficult to identify treatment effects due to the CG.

- **Difficulty seeing effects at endline:** As in Papua, it is important to emphasise that the ToC underlying the intervention in Aceh includes some strong assumptions around why impact might materialise. It is clear to the evaluation team that some of those assumptions might not, in fact, come into being, which means that there is a real possibility that endline analyses might not find any effects of the CG in Sabang. It should also be noted that for Aceh this issue is exacerbated by the fact that CG payments are planned to start by the end of July 2019. This means that the time in which effects could materialise between the two key survey rounds (March 2019 and March 2020) will be smaller than anticipated (around eight months between July 2019 and March 2020 compared to an original estimate of around 20 months).

- **Parallel implementation of other components and cash transfer programmes:** Finally, as discussed further above and presented in section 4.1, the intervention in Aceh is a bundled intervention, of which the CG is only one component. At endline, the evaluation team will try to assess the extent to which this bundling actually happened and how the targeting of the different components of the intervention overlapped with CG targeting. This will allow us to understand how far these different bundles overlapped and might hence together be driving change in Sabang. For the evaluation, this means that we might not be able to distinguish the effects of the CG separately from the other bundled interventions.

In addition, however, similar to the situation in Papua, we find that a significant proportion of households in Sabang are already recipients of some sort of financial support (see section 4.5). As discussed there, this means that it is unclear whether the parallel trend assumption required for the DID analysis at endline will hold. It might be possible to control for these alternative cash transfers if SUSENAS 2020 data will allow us to distinguish these from each other.

5 Conclusion and next steps

This baseline report describes the characteristics of households and children prior to the disbursement of any grants in treatment and comparison areas in Papua and Aceh. Households and children in the districts where the CG is implemented or is expected to be implemented are compared to households and children in areas not covered by the CG. This is to assess differences and similarities across the two groups before the implementation of the two CG programmes that are the focus of this impact assessment.

Summaries of key findings related to each of the expected programme outcome areas are presented in blue boxes at the beginning of each subsection above. In addition, sections 3.10 and 4.8 present the methodological implications of our findings for the endline phase of this study, separately for Papua and Aceh. To avoid repetition, we do not summarise these findings and implications here again.

Rather, this conclusion focuses on overarching takeaways that are relevant for the next phase of this evaluation and that relate to both study components in Papua and Aceh:

- **In both Papua and Aceh the implementation of the CG programmes faces significant challenges.** For example, at the time of submitting this baseline report (July 2019), grants have not yet been disbursed in Aceh and only the Government of Sabang has allocated funding to the implementation of the CG in this province from April 2019 onwards. This differs significantly from plans described in the original Terms of Reference to this evaluation, which specified that the CGs in Aceh were supposed to be disbursed in four districts from 2018 onwards. In Papua, the remoteness of beneficiary communities and security concerns complicate grant distribution in treatment districts. For example, as described in section 3.1, as at May 2019 about 40% of all targeted children had been registered in Paniai. Participants in a validation workshop in Jayapura in May 2019 also reported communities refusing to be part of the CG programme in some areas of Papua. Taken together, these challenges indicate that there is a possibility that the effects of the two CG programmes on the population of households and children that form part of this study might be limited.

  This quantitative impact assessment will not be able to provide evidence on all of these challenges. As currently designed, this impact evaluation will aim to track outcomes that are part of each programme's ToC and will attempt to provide an assessment of whether the programmes have had an effect on those or not. It will also be able to provide some evidence on whether some of the challenges listed above might have limited the potential effects that could have resulted from the grants. For example, in Papua we will ask households at endline whether they ever received grants from BANGGA Papua, which will give us an indication of the coverage that the programme achieved there. However, we will not be able to track all assumptions underlying the causal chain depicted in each programme's ToC or assess all challenges that the programmes might have faced.

  We do think, however, that gathering evidence on the process by which the CG is implemented on the ground and independently tracking the implementation of each programme would be beneficial to understand why impact might materialise and how. It would greatly increase the usefulness of our endline analysis. In this regard, **OPML has suggested to UNICEF that it considers conducting an**

Implementation Review of the CG programmes that could closely link our endline analysis to qualitative evidence provided on how in both Aceh and Sabang the CG programmes were implemented.

- The baseline phase of this evaluation confirmed the complexity of the context within which the CG programmes and this evaluation are being implemented in both provinces. For the evaluation, this relates to both data collection efforts and interpretation of results. In this regard, regular communication with BPS was very helpful in designing our evaluation, accessing data, and interpreting results. For example, BPS provided significant inputs in the process of designing our baseline survey in Papua. Similarly, UNICEF and BPS collaborated closely to provide us with SUSENAS data for the analysis in Aceh.

  At endline, it will therefore be important to continue collaborating closely with BPS, especially given that the impact evaluation in Aceh will rely on SUSENAS data again. Similarly, BPS inputs will be instrumental in supporting OPML and UNICEF in the design of the endline data collection in Papua.

- **It is important to mention that the overall evaluation of the CG programmes is composed of both a quantitative and qualitative component.** The qualitative component is being implemented separately from this quantitative component (Jupp *et al*., 2018). Hence, findings from the quantitative baseline analysis presented here should be interpreted in conjunction with qualitative insights in order to contextualise and triangulate results. OPML collaborated with Empatika to disseminate evaluation findings as well as to produce a set of briefs to foster information sharing, learning, and dialogue among CG stakeholders. OPML aims to continue working with the qualitative evaluation team to ensure endline findings provide a comprehensive understanding of the effects of the CG on beneficiaries.

- Lastly, it should be noted, again, that in all treatment districts that this evaluation is currently focusing on (i.e. Sabang in Aceh and Paniai, Lanny Jaya, and Asmat in Papua) cash transfer and social assistance programmes are already being implemented. Most households in Sabang and about 50% of all households in the treatment districts in Papua already receive some sort of cash support. We summarise the methodological implications of this situation in sections 3.10.1 and 4.8. It should be noted here, however, that it is the evaluation team's perspective that this overlap could create the basis for beneficial complementarities between the CG and other social assistance programmes, which are worthy of further investigation.

References

BPS (n.d.) 'Poverty line concept'. www.bps.go.id/subject/23/kemiskinan-dan-ketimpangan.html#subjekViewTab1 [accessed 3 June 2019]

--- (2017) *Survei Sosial Ekonomi Nasional (Susenas)*, Jakarta.

--- (2018) *Survei Sosial Ekonomi Nasional (Susenas)*, Jakarta.

Kementerian Kesehatan (RI) (2017) *Riset Kesehatan Dasar (RISKESDAS)*, Jakarta.

Jupp, S. *et al*. (2018) 'Universal Child Grant baseline qualitative review'. Jakarta: Palladium.

OPML (2018) 'Quantitative Impact Assessment of the Child Grants in Aceh and Papua Provinces Inception report'. Oxford: OPML.

Priebe, J. (2014) 'Official Poverty Measurement in Indonesia since 1984: A Methodological Review', *Bulletin of Indonesian Economic Studies*, 50:2, 185–205.

Rubin, D. B. (2001) 'Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation', *Health Services and Outcomes Research Methodology*, 2, 169–188.

Rosenbaum, P. R. and Rubin, D. B. (1985) 'Constructing a control group using multivariate matched sampling methods that incorporate the propensity score', *The American Statistician*, 39(1), 33–38.

UNICEF (2019) 'Multiple Indicator Cluster Surveys 6'. www.mics.unicef.org/tools#survey-design [accessed 3 June 2019]

World Bank (2010) 'Poverty Measurement Methodology for 2010, Indonesia'. www.siteresources.worldbank.org/EXTPOVERTY/Resources/idn2010.pdf [accessed 3 June 2019]

WHO (2008) 'Indicators for assessing infant and young child feeding practices: conclusions of a consensus meeting held 6–8 November 2007 in Washington D.C., USA.' Washington D.C., USA.

--- (2008b), 'Indicators for assessing infant and young child feeding practices part 2: measurement.' Washington D.C., USA.

--- (2019) 'Child growth indicators and their interpretation'. www.who.int/nutgrowthdb/about/introduction/en/index2.html [accessed 3 June 2019]

--- (2010) 'Nutrition Landscape Information System (NLIS) country profile indicators: interpretation guide'. Geneva.

--- (2007) 'Indicators for assessing infant and young child feeding practices: conclusions of a consensus meeting held 6–8 November 2007 in Washington D.C., USA'.

WHO and UNICEF (2009) 'WHO child growth standards and the identification of

severe acute malnutrition in infants and children. A Joint Statement'. Geneva. www.apps.who.int/iris/bitstream/handle/10665/44129/9789241598163_eng.pdf?ua=1 [accessed 3 June 2019]

Annex A Methodological annex

A.1 Consumption estimates in Papua

This annex provides further technical details on how we constructed consumption aggregates in Papua using the baseline household survey data.

To estimate consumption levels in Papua we sum up the monthly monetary value of a range of food and non-food items as captured by our household survey. We designed our methodology to be as similar as possible to BPS's approach to measuring consumption and monetary poverty, relying primarily on BPS methodological notes, as available online, and on World Bank (2010) and Priebe (2014).

It should be emphasised here that publicly available information does not give an exhaustive overview of BPS methodology. Moreover, while it is good practice to keep consumption estimation coherent over time to ensure comparability of results, minor variations due to methodological improvements as well as changes in consumption habits and technology do take place. We are not able to account for these changes in our approach. Table 9 compares our methodology to the construction of consumption aggregates to the BPS approach. There are two key differences that we would like to highlight here:

- **Reference population**. As per our general survey design, we only focus on Papuan households with children under two years old in a set of accessible areas in treatment and control districts. BPS, however, produces national and provincial level estimates, aiming to represent all parts of Indonesia and the whole of Papua.
- **Consumption items included in the calculation of the consumption aggregate**. Our consumption aggregate includes only items that together comprised 95% of the total household consumption as measured by SUSENAS in Papua in 2017. This is because the whole SUSENAS consumption module would have been too time consuming to administer in our survey. We also do not take rent into account, given the large number of missing values that we observed in our survey. In addition, we exclude government transfers because we have not collected data on this due to time constraints. We also exclude major ceremonies. Due to the short span of our data collection exercise, attempting to include these expenses could result in distorted consumption patterns across areas. Excluding these items from our main consumption aggregate means that our approach differs from BPS's approach, which considers all food and non-food items used for household purposes, including rent, transfers, and major ceremonies as part of the consumption aggregate.

It should be noted that no information is available to the study team to assess the discrepancy between our approach and BPS methodology when it comes to spatial adjustments (i.e. adjustments to differences in local prices) and the treatment of outliers.

Table 9: BPS and OPML methodologies for constructing consumption aggregates

Data			
		BPS	**OPML baseline survey**
1.	Data source	SUSENAS	Primary household survey based on 95% of the list of SUSENAS 2017 consumption items
2.	Population	Papua population	Study area population – indigenous Papuans with children under two years old
3.	Consumption items	Food Non-food including housing, goods and services, education, health, clothing and toiletries, durable goods, taxes and insurance, and festivities and ceremonies	Food Non-food including housing, goods and services, education, health, clothing and toiletries, durable goods, taxes and insurance, and festivities and ceremonies
Important expenditure categories			
4.	Self-produced items	Self-estimate from survey	Self-estimate from survey
5.	Durable goods	Included – no adjustment for depreciation	Included – no adjustment for depreciation
6.	Rent	Self-estimate from survey (no imputation techniques applied to the data)	Excluded due to large number of missing values
7.	Transfers	Included in the consumption aggregate	Excluded because we have not collected data on this due to time and budget constraints. The household survey builds on the SUSENAS food and non-food consumption module, which does not investigate transfers. Data on transfers are collected separately in SUSENAS
8.	Major ceremonies	Included in the consumption aggregate	Excluded. These expenditures are usually excluded from the consumption aggregate. Due to the short span of our data collection exercise, including these expenses would result in distorted consumption patterns across areas
9.	Taxes	Included in the consumption aggregate	Included in the consumption aggregate
10.	Education	Included in the consumption aggregate	Included in the consumption aggregate
11.	Health	Included in the consumption aggregate	Included in the consumption aggregate
Adjustment and treatment of outliers			

12.	Spatial price adjustment to account for geographical differences in prices	Yes – no available information on the methodology for construction of price index	Yes. We use the food price index to adjust for differences in prices for both food and non-food consumption. This is because we do not have unit costs for non-food items due to time and budget constraints. To construct the food price index we rely on the prices of food items as captured by the household survey. Items consumed by less than 3% of the population have been dropped, as well as heterogeneous food categories (i.e. other vegetables, etc.)
13.	Seasonal adjustment to account for temporal differences in prices	No	No
14.	Method of treating outliers/missing values	No available information	Food consumption values are outliers if above or below 3.5 SD of the distribution of the item value by district and urban/rural areas. Criteria are more stringent for non-food items (2 or 1.5 SD). Outliers have been replaced by the minimum or maximum of the related interval of variation, by district and urban/rural area No imputation of missing values has been carried out

As part of the consumption analysis we also compare *per capita* consumption to the Papua poverty line and food poverty line, to assess the level of monetary poverty in our target population. Based on this approach, an individual is poor if he or she does not have enough consumption to put him or her above the Papua poverty line. A person is food poor if his or her consumption is below the Papua food poverty line.

We produce the following measures of poverty:

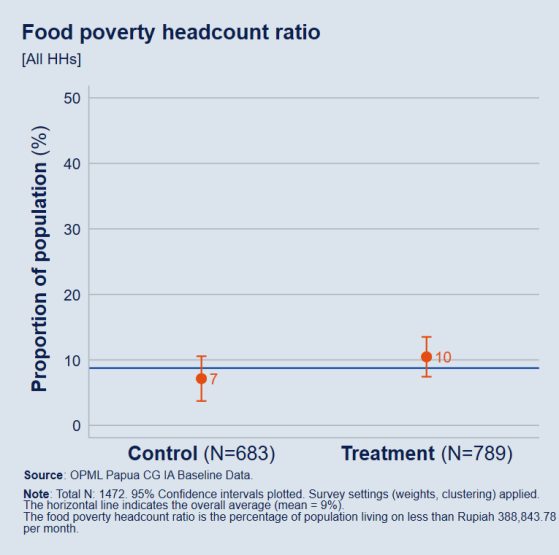
- The poverty headcount ratio, which captures the proportion of the population whose monthly *per capita* consumption is below the Papua poverty line of IDR 518,811.31, as set by BPS in September 2018.
- The food poverty headcount ratio, which captures the proportion of the population whose monthly *per capita* consumption is below the Papua food poverty line of IDR 388,843.78, as set by BPS in September 2018.
- The poverty gap, which measures the depth of poverty, as the average percentage shortfall in income for the population, from the Papua poverty line.

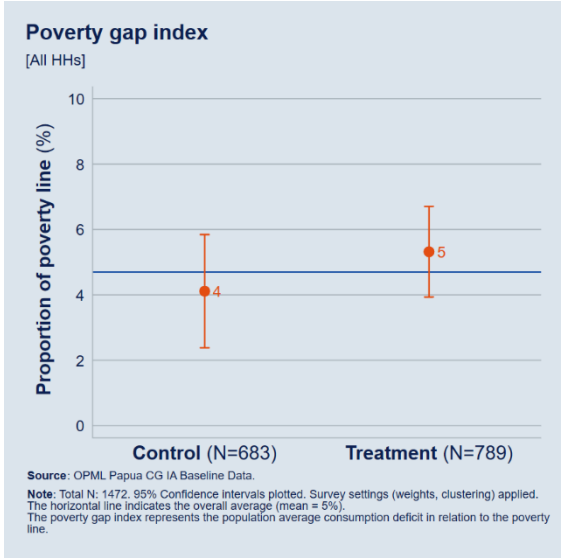
Section 3.6.1 presents the poverty headcount ratios by evaluation group. Figure 51 below presents the food poverty headcount ratio and the poverty gap at IDR 518,811.31 *per capita* per month.

Overall, 9% of the population is food poor, which indicates that about one in 10 people in the study areas do not meet the minimum level of consumption required to acquire a basket of basic food commodities, equating to 2,100 kilocalories per day. Food poverty is slightly higher in the treatment group, although the difference is not statistically significant (Figure 51 left).

The poverty gap index is at 5% on average, with no significant difference between the treatment and control groups (Figure 51 right). This measures the average percentage shortfall in income for the population, from the poverty line. Based on this indicator of the intensity of poverty, an average increase of IDR 26,000 per individual would be enough to raise individual consumption up to the level of the poverty line.

Figure 51: Food poverty headcount and the poverty gap





A.1.1 Robustness checks

To test the robustness of our consumption estimates we study how changes in some of our methodological assumptions would impact these estimates and corroborate our estimates with some additional information. In particular, the following paragraphs present results related to four specific robustness checks:

1) How do changes in consumption aggregates affect our estimates?
2) How do changes in the price indices affect our estimates?
3) What level of caloric intake per person does our estimate for food consumption reflect?
4) Can we observe any time-related trends in consumption estimates?

Overall, our analysis shows that consumption estimates are not particularly sensitive to changes in aggregates or price indices used. However, our analysis of caloric intake reveals that consumption estimates are high and probably higher than expected.

Finally, our time analysis reveals that there is no particularly strong time-trend in our consumption data that would indicate that household consumption was affected by Christmas coming up toward the end of data collection.

In general, the results in this section provide evidence for how difficult it is to collect accurate consumption data via household surveys in Papua. As described in section 3.10.2, one implication for our endline survey will be to try to collect asset and alternative wealth-related data that could allow us to estimate household-level poverty using some form of multidimensional poverty index.

A.1.1.1 Changing the composition of consumption aggregates and the underlying price indices

In a first step, we tested the sensitivity of our results to changes in the construction of aggregated consumption. In particular, we assess how results change when we rely on different consumption aggregates. We compare our preferred model (**M3** in the tables and figures below), which excludes expenses on rent, parties, and ceremonies, to two alternative models:

- **Model M1:** total consumption considering all items captured by the household survey.
- **Model M2:** total consumption *excluding* expenses on durables, rent, and parties and ceremonies.

In addition, we assess how M1, M2, and M3 estimates change relying on alternative price adjustment methods. In models M1, M2, and M3, nominal consumption is adjusted taking into account the prices of all food items, as captured by our household survey. In models M4, M5, and M6, we adjust the nominal consumption by a 'reduced price index'. This has been constructed using the prices of a subset of food items. For this index, we considered only the prices of those items that have less than 3% of outliers.[31] Both indices are presented in Table 10 below.

Table 10: Price indices under different models

District	Original price index	Reduced price index
Asmat	0.862852	0.923432
Boven Digoel	0.706016	0.750141
Keerom	0.851869	0.917189
Lanny Jaya	1.844841	1.500442
Paniai	0.932476	1.073079
Waropen	0.840934	0.94632

Figure 52 presents the results of these robustness checks, disaggregated by treatment and control groups. The left-hand panel presents the average monthly consumption per adult equivalent as estimated by each of the models above. The right-hand panel

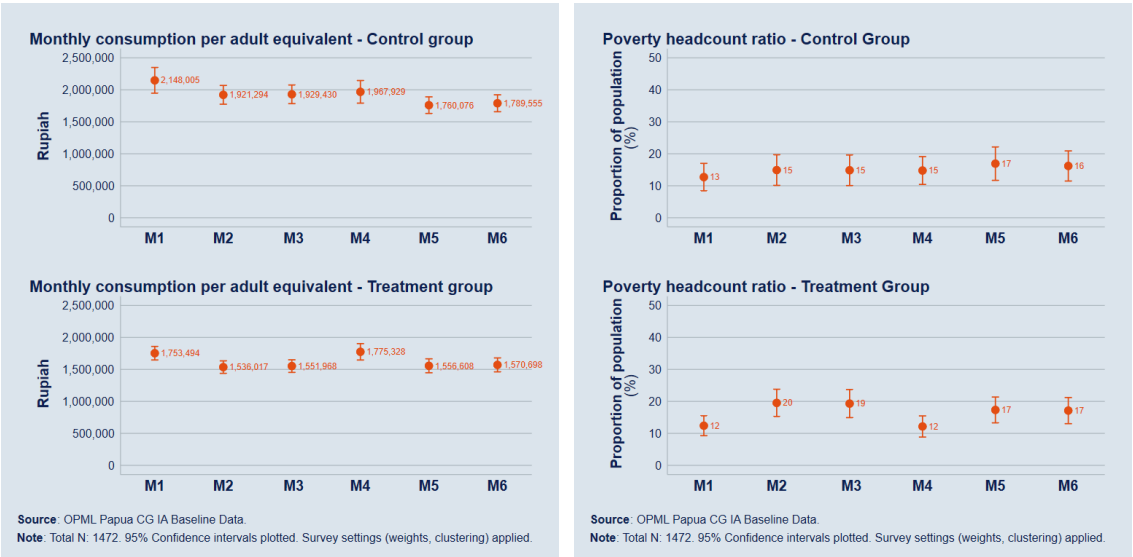
[31] Consumption values are considered as outliers if outside the critical levels defined by +-1.5*SD from the median or three times the median or one-third of the median of the distribution of the value of each consumption item by district and urban/rural area.

shows how monetary poverty changes depending on the consumption aggregate taken into account.

Comparing M1 to M2 and M3, the differences in consumption are not very large. As expected, consumption aggregates under M1 are higher than the consumption aggregates estimated by M2 and M3, which exclude several items. However, this does not affect the poverty headcount ratios very much, except for the treatment group between M1 and the other two models. Expenses on durables, as captured by the difference between M3 and M2, are minor. This suggests rental expenditure and expenditure on ceremonies drive the discrepancy between M1 and the other models in the treatment group. In the treatment areas, less than 1% of households pay rent. This indicates rental expenditure and consumption estimates are mostly based on self-assessments by households on what the rent would be if they were tenants, with high risk of reporting bias. The large number of missing values in the reported rent confirms the complexity of collecting rental information. For this reason, we decided to exclude rent expenses from the consumption aggregate estimation. Similarly, we decided to exclude expenditure on ceremonies to avoid distortions due to location-specific events. We therefore considered M3 as our preferred model.

Comparing M1 to M4, M2 to M5, and M3 to M6, monthly consumption is slightly higher using the original price index compared to the estimates adjusted using the reduced price index. Poverty headcount estimates, however, are consistent across different price index specifications. This indicates that outliers in food prices do not distort the price index, confirming the robustness of the original index. Hence, we remain confident that M3 is our preferred model.

Figure 52: Robustness checks to changes in consumption aggregate construction



A.1.1.2 Estimating daily caloric intake

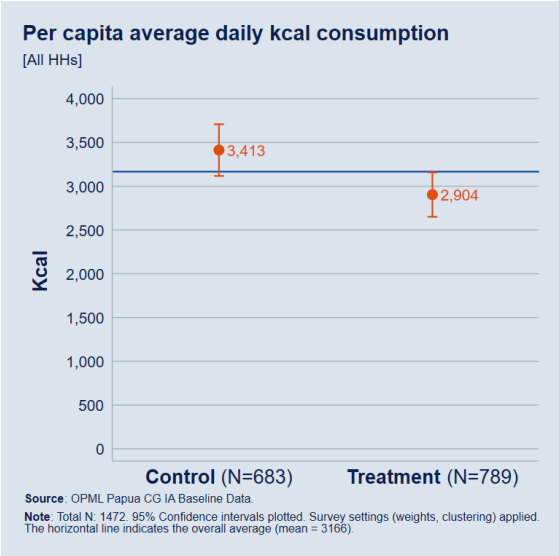
The consumption module of the household survey collected information on both the value and the quantity of food consumed. We use the quantity consumed to triangulate food consumption information to assess whether the quantities reported by households are sensible when transformed into individual energy intake as measured by kilocalories (kcal).

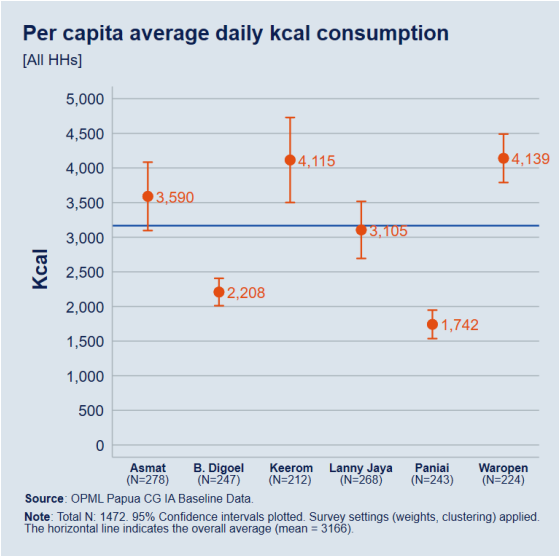
We compute the kcal intake per person per day, as derived by the quantity of food reported to be consumed and the equivalent caloric intake by food item as specified by the SMILING food composition table for Indonesia.[32] On average, our findings are higher than the nutritional requirement for good health of 2,100 kcal per person per day (Figure 53 left). Results vary across districts, with individuals in Waropen consuming more than twice as much as individuals in Paniai (Figure 53 right).

These findings might be related to the specificities of local diets based on high consumption of rice and tubers and roots, seasonality effects, the definition of the household applied by the survey, and difficulties in collecting food quantity information. For example, as was discussed in section 3.4.1, in the context of Papua applying the usual definition of the household was sometimes difficult. This translates into difficulties in identifying household members, which could result in underestimation of household size. This in turn would overestimate the *per capita* caloric intake in any given household.

The triangulation analysis, together with the fact that food accounts for 70% of household consumption on average, indicates that consumption estimates presented in section 3.6.1 are comparatively high.

Figure 53: *Per capita* average daily kcal consumption





A.1.1.3 Checking distribution of consumption levels over time

We also checked the trend in the consumption levels reported during the data collection period, to assess whether our consumption estimates might include exceptional expenses due to district-level or general events coming up toward the end of our data collection period, which was close to Christmas.

Figure 54 below displays consumption levels reported for each day of data collection, per district. The daily median consumption is reported to present changes in consumption over time. The trend is fairly stable in Boven Digoel, Keerom, Paniai, and Waropen. In Asmat and Lanny Jaya, we observe a slight decrease in the reported consumption after the first 10 days of data collection. A regression analysis of the

[32] See www.fao.org/infoods/infoods/tables-and-databases/asia/en/

correlation between the average consumption and the day of interview confirms that consumption estimates in the first 10 days of data collection are significantly higher on average in Asmat and in Lanny Jaya than in the remaining interview days (Table 11).

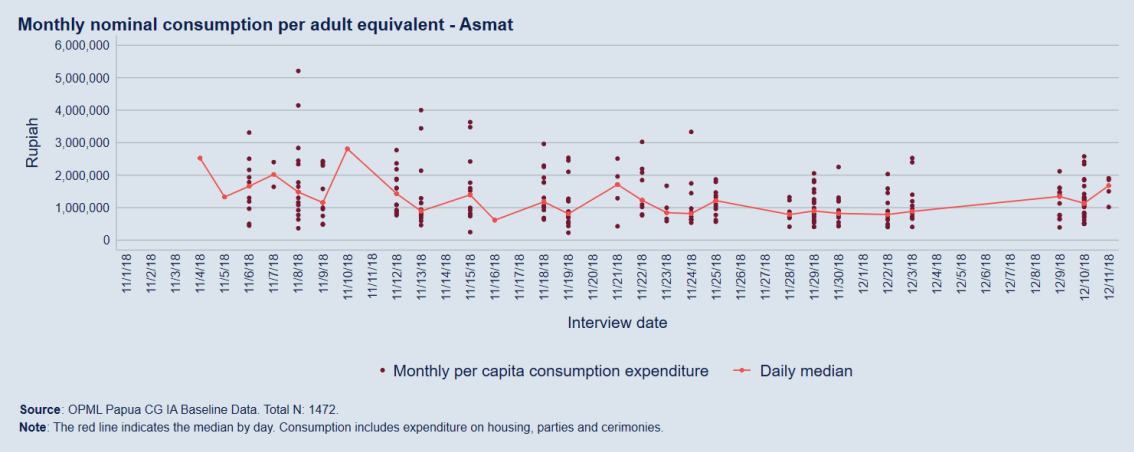
This indicates that our consumption estimates in these two districts could include exceptional expenses due to contingencies or local ceremonies. This fact contributed to our decision to drop expenditure on ceremonies from the calculation of consumption aggregates, to avoid distortions due to systematic variation associated with location or time of the year.

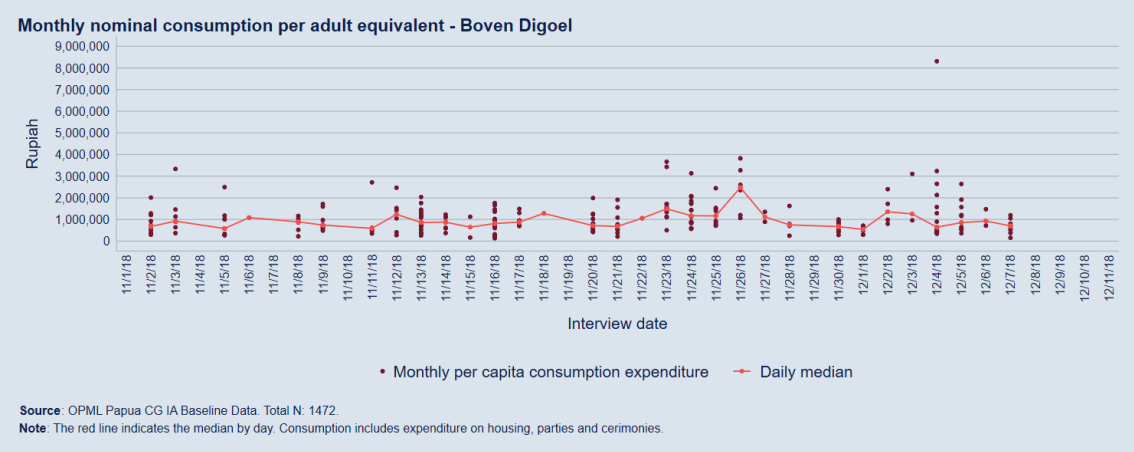
Table 11: Consumption expenditure before and after the first 10 days of data collection (before and after 13 November)

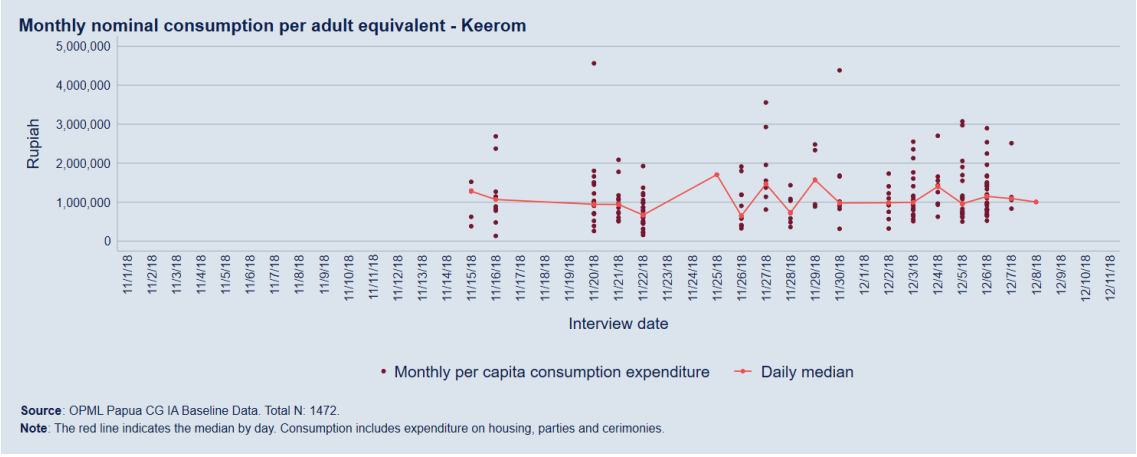
	Overall	Asmat	Boven Digoel	Keerom	Lanny Jaya	Paniai	Waropen
First 10 days of data collection	300,128***	374,285***	-137,892		496,706**	-32,985	154,543
	(103,642)	(86,491)	(115,055)		(199,684)	(123,206)	(136,203)
Constant	1.083e+06***	1.033e+06***	1.011e+06***	984,343***	1.547e+06***	936,697***	1.164e+06***
	(33,435)	(48,057)	(101,000)	(70,288)	(56,163)	(45,980)	(63,162)
N	1,470	278	247	210	268	243	224
R-squared	0.025	0.057	0.006	0.000	0.047	0.000	0.012

Source: OPML Papua CG Impact Analysis Baseline Data
Note: Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

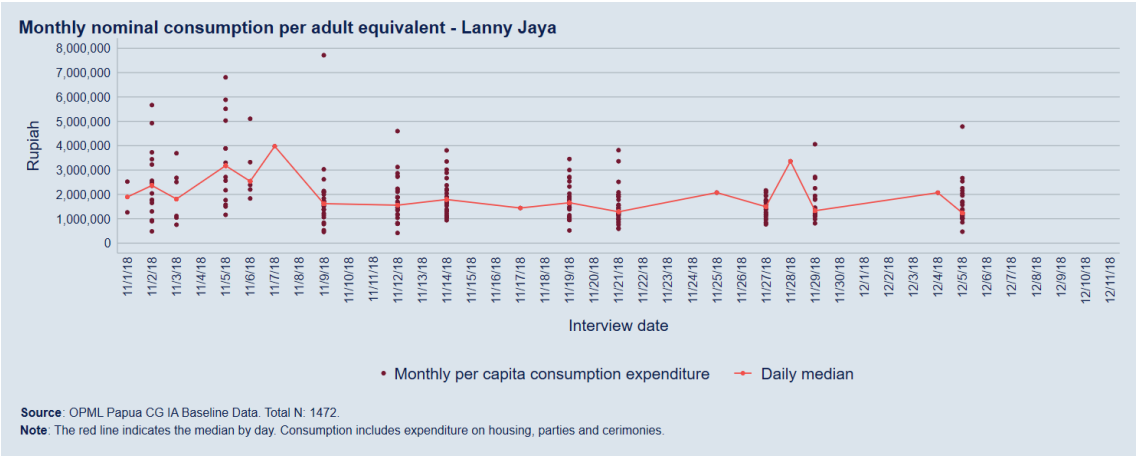
Figure 54: Monthly nominal consumption per adult equivalent by day of interview and district



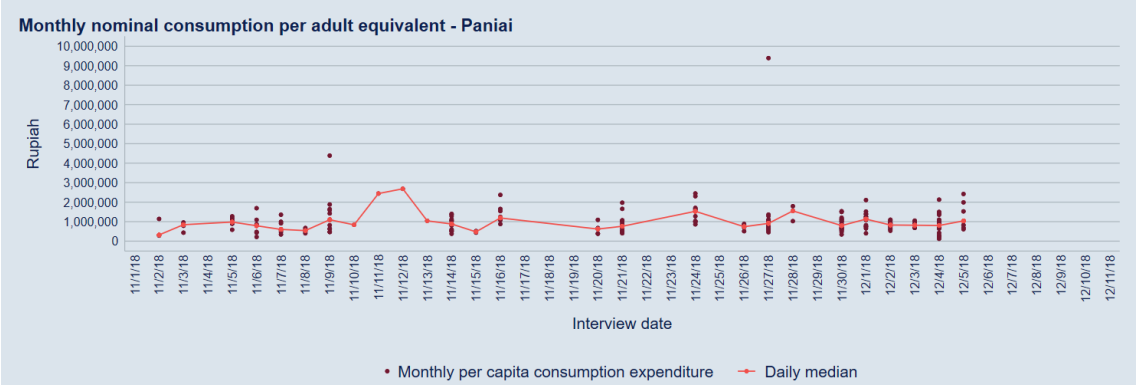




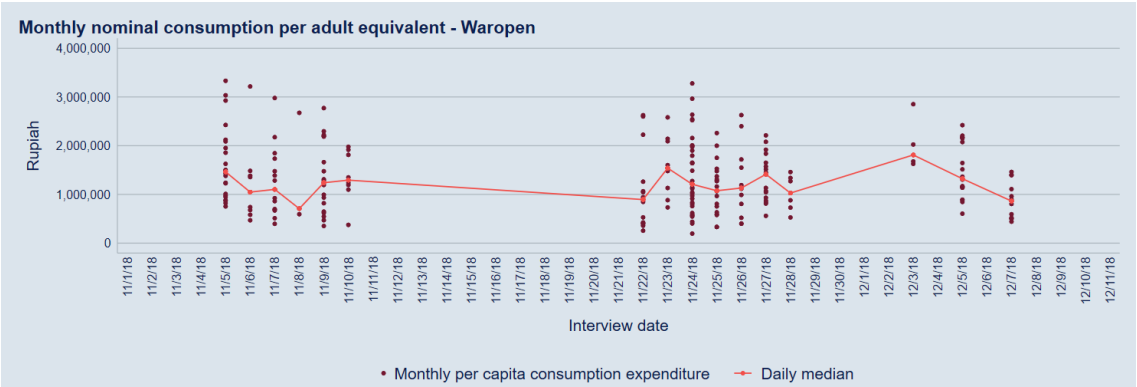
Source: OPML Papua CG IA Baseline Data. Total N: 1472.
Note: The red line indicates the median by day. Consumption includes expenditure on housing, parties and cerimonies.



Source: OPML Papua CG IA Baseline Data. Total N: 1472.
Note: The red line indicates the median by day. Consumption includes expenditure on housing, parties and cerimonies.



Source: OPML Papua CG IA Baseline Data. Total N: 1472.
Note: The red line indicates the median by day. Consumption includes expenditure on housing, parties and cerimonies.



Source: OPML Papua CG IA Baseline Data. Total N: 1472.
Note: The red line indicates the median by day. Consumption includes expenditure on housing, parties and cerimonies.

A.2 Papua baseline survey: survey objectives, sampling, weights, and survey settings

A.2.1 Survey objectives

The following paragraphs describe how we implemented primary data collection for the purposes of the Papua component of this evaluation. It should be noted that this data collection exercise had three main objectives:

- First, to collect data from individual children and their households who were eligible to receive the CG in the treatment districts and were estimated to be eligible until the endline data collection, which implied collecting data from Papuan children aged 0–23 months.
- Second, to collect data in a way in which estimates were representative for this population of children and their households in the areas in which the survey was implemented, which implied some form of random sampling.
- Third, to collect data in comparison areas in a way that allowed appropriate comparison to data in the treatment areas for the purposes of the impact assessment and attribution of impact, which would include the use of PSM at endline.

The following paragraphs explain how we sampled respondents for this survey in order to achieve the above objectives.

A.2.2 Sampling

PSM helps overcome selection bias by comparing eligible households and children to non-eligible households and children who are observationally similar. Because the CG can affect characteristics of children and households at many different levels, this similarity needs to be established before the CG starts, i.e. at baseline. We show that it is possible to establish this similarity using PSM and baseline data in section 3.9.

It is important to emphasise here that eligibility, in this context, is defined both geographically and by individual children's characteristics: first, no children in the comparison areas will be eligible for the CG, because the grant is only paid out in the treatment districts for now: Asmat, Paniai, and Lanny Jaya. Second, for the purposes of the CG programme, children (and their households) are considered eligible when they are aged 0–3 years old (under the age of four) and have a Papuan parent or Papuan main caregiver. Because the present study focuses on children who will be eligible for the CG throughout the study period, the eligible children for our survey were thus 0–1 year (0–23 months) old. Note that being Papuan in this context means being indigenous Papuan, which parents or caregivers self-identify as for the purposes of our study.

Because the assumption is that the CG will be rolled out universally to all eligible households within districts, and we are interested in comparing outcomes across children in similar age groups, we compare outcomes in the three districts where the programme will be first implemented (Lanny Jaya, Asmat, and Paniai) with outcomes in districts where the programme is not scheduled to roll out during the period of the study, i.e. only from 2020 onwards. In order to be able to do this, this study's sample had to be drawn using a similar approach in both treatment and comparison areas. We have therefore implemented a multi-stage sampling process in the following manner:

Step 1: Selection of comparison districts

To find appropriate comparison districts to visit, OPML implemented a district matching exercise using available data in order to identify those districts that were as similar as possible to the treatment districts on a range of variables that were also initially used to identify the three districts of Lanny Jaya, Asmat, and Paniai. Note that this analysis was limited to districts for which UNICEF documentation indicated that the CG would only be implemented from 2020 onwards.

The variables used were as follows – all taken from data provided by UNICEF to the OPML team, originally sourced from BPS SUSENAS or simulations conducted by UNICEF:

- Proportion of rural population (2015)
- Population density (2015)
- Estimated income *per capita* (GRDP) at 2013 prices (IDR) in 2015
- Proportion of 0–7 year olds in the district (2015)
- Proportion of the population living below the poverty line (2015)

- Human Development Index (2015)
- Life expectancy (2015)

We used this data to construct a matrix that contained the Euclidean distance of each non-treatment district in Papua from each of our three treatment districts taking all seven dimensions above into account, using the *'dist'* function in the statistical package R.[33] From this matrix, we identified control districts that were closest to the treatment districts and shared information on this with UNICEF. The final selection of control districts was then made taking logistical issues of accessibility, i.e. the possibility of doing fieldwork in the districts, into account.

Table 12 below lists each of the treatment districts and the selected comparison districts.

Table 12: Selected treatment and comparison districts

Treatment districts	Comparison districts
Lanny Jaya	Keerom
Asmat	Boven Digoel
Paniai	Waropen

Step 2: Selection of sub-districts within selected districts

Within treatment districts, OPML then selected a set of sub-districts within which the survey would be implemented. These sub-districts were selected based on accessibility, security situation, population size, and their geographical dispersion. In each district, the objective was to select accessible and safe sub-districts that were spread across the district's area (i.e. to avoid clustering in one area if possible) and to include the district capital in the survey.

It should be noted here that selecting sub-districts for the purposes of this study was a complex process due to changes in the administrative structure of districts in Papua in 2014/15 and the difficulty of finding consistent information about the size and number of sub-districts in each of the treatment and comparison areas. For example, publicly available information on Asmat produced by BPS ('Kabupaten Asmat 2018 in figures') and by the Ministry of Internal Affairs (http://gis.dukcapil.kemendagri.go.id/peta/) are not consistent with respect to the number of sub-districts that exist in Asmat.

In addition, the change in administrative structures in 2014/15 has also led to a significant increase in the number of sub-districts in Lanny Jaya and Paniai (which means that 'new' sub-districts are now smaller than 'old' ones): in Lanny Jaya and Paniai, BPS information from 2015/16 listed 10 sub-districts in each of the districts, while data from the Ministry of Internal Affairs from 2017/18 lists 39 and 23 sub-districts respectively.[34] The change in Asmat was from a total of 23 to 19. In comparison districts, such significant change did not happen. It is unclear how exactly this

[33] See http://stat.ethz.ch/R-manual/R-devel/library/stats/html/dist.html for how this is implemented in R. The Euclidean distance between two vectors $\mathbf{a}$ and $\mathbf{b}$, here defined as vectors that contain the variables above for two districts $\mathbf{a}$ and $\mathbf{b}$, is defined as: $d(\mathbf{a}, \mathbf{b}) = \sqrt{\sum_{i=1}^{n}(a_i - b_i)^2}$, where in the present case the sum is over the number of variables included.

[34] BPS information retrieved from 'Kabupaten in figures' publications by BPS. Information from the Ministry of Internal Affairs retrieved from http://gis.dukcapil.kemendagri.go.id/peta/.

extensive re-drawing of borders that happened in Paniai and Lanny Jaya affected the population size in each of the sub-districts there.

The original objective was to select five 'old' sub-districts in each treatment district. However, given the re-drawing of boundaries, and using the most recent data from the Ministry of Internal Affairs, the evaluation team decided to select, for Paniai and Lanny Jaya, a more comprehensive list of sub-districts, rather than just five as originally intended. For Asmat, the team selected a total of four sub-districts.

Within comparison districts, the objective was to select five sub-districts that in as much as possible matched the characteristics of the sub-districts covered in the treatment areas. Together with Myriad, OPML selected these sub-districts using publicly available data from the National Team for the Acceleration of Poverty Reduction on the welfare status of households living in these sub-districts. Given the difference in the number of sub-districts selected in treatment and comparison areas, this was not a one-to-one matching exercise. Rather, the welfare statistics in sub-districts that were selected in comparison areas corresponded in as much as possible to the range of values found in the treatment areas.

A final adjustment to the list of sub-districts surveyed was done at the beginning of fieldwork, when field teams were in the districts. Security assessments were then implemented by the team in order to understand whether visiting a sub-district was possible or not, given tribal conflict and other security concerns.

The purposeful nature with which sub-districts were selected in treatment and control districts implies that indicator estimates are representative of the target population in those sub-districts only. OPML estimates, derived from Ministry of Internal Affairs data, suggest that in treatment areas this means that we covered about 40% to 55% of the total population there, if these data are accurate.[35] The geographical spread of these areas ensures that the survey covers a variety of different areas in the treatment and control districts and is not clustered around one central area only. A final list of selected sub-districts can be found in Table 13 below.

Table 13: List of selected sub-districts

Treatment area	
District	**Sub-district**
Asmat	Agats
	Suator
	Akat
	Sawaerma
Paniai	Paniai Timur
	Paniai Barat
	Bibida

Comparison area	
District	**Sub-district**
Boven Digoel	Mindiptana
	Iniyandit
	Fofi
	Mandobo
	Sesnuk
Keerom	Skanto
	Arso

[35] Available upon request.

Treatment area	
Lanny Jaya	Tiom
	Yiginua
	Yugungwi
	Gelokbeam
	Kolawa
	Makki
	Kully Lanny
	Muara
	Karu
	Poga

Comparison area	
	Web
	Waris
	Arso Timur
	Arso Barat
Waropen	Wapoga
	Inggerus
	Masirei
	Demba
	Risei Sayati
	Waropen Bawah
	Oadate

Step 3: Sampling of villages

OPML then implemented a random sampling approach within the subset of sub-districts selected in each of the treatment and comparison districts in order to sample villages within which the survey would be implemented. These villages are our so-called PSUs. The list of villages from which to draw the sample was defined as all the villages that were located in the sub-districts selected in step 2 above and that were considered to be accessible based on information received in Jayapura from local fieldwork teams. Note that accessibility refers here to both geographical accessibility (e.g. we excluded villages that could only be visited using helicopters) and accessibility with respect to security concerns.

To do the sampling, we implemented a stratified (per district) random sampling approach (sampling probability proportional to size) to sample villages using publicly available PODES 2011 data and information from 2017 from the Ministry of Internal Affairs (http://gis.dukcapil.kemendagri.go.id/peta/). Given that between 2011 and 2017 the administrative boundaries in Papua were updated (as discussed above), this meant matching the list of 2011 villages in PODES (for which PODES provides estimates of the number of households living in them) with a 'new' list of villages provided by the Ministry of Internal Affairs. From this list we excluded any villages that we knew would not be accessible given security and logistical concerns. This final matched list constituted our sample frame from which to draw the final sample of villages. Per district, a total of 20 villages was originally sampled. The sample also included 10 replacement units for cases where villages in our main list turned out to be inaccessible, e.g. due to tribal conflicts or other security concerns.

It is important to mention here that, depending on the size of villages that the field teams encountered on the ground, a segmentation or combination protocol for villages was implemented. This has meant that the final list of villages sometimes consisted of combinations or segmented villages from our original sampling frame. It was necessary to implement these protocols because information on the size of villages from PODES

2011 or other publicly available sources was not always accurate. The combination or segmentation protocol was implemented as follows:

- First, the field team obtained an estimate of the size of the village (in terms of households) from the village head or sub-administrative unit official (*Rukun Warga* (RW) or *Rukun Tetangga* (RT)).
- Second, the decision of segmentation or combination was made as follows:
    - Villages that were estimated to have more than 200 households were segmented in two by randomly picking a sub-administrative unit. This sub-unit became our PSU for listing purposes (next step).
    - Villages that were estimated to be smaller than 120 households were combined with a neighbouring village where possible. If the neighbouring village was too far away (more than two days required to combine and list) or inaccessible, then no combination was implemented. The closest accessible village was then used as a unit to combine with.

Step 4: Sampling of children and households

Within each of the selected PSUs, a listing exercise was conducted to identify eligible children for the survey and their households. As described above, eligibility for the purposes of our survey was defined as being 0–23 months old and having a Papuan caregiver or parent. The listing exercise listed all households and all eligible children within each of the PSUs.[36] In a second step, a random 'main' sample of 15 children and of 10 'replacement' children was drawn for the purposes of data collection.

Caveats

It is important to emphasise, for the present analysis, that selection and sampling of villages in the sub-districts listed above faced significant challenges. In some instances, villages were selected and visited, but data collection could then not be implemented because villagers did not allow enumerators to access the village. Similarly, information and concerns coming from the field led to an adaptation of the sampling and analysis approach for three districts:

- In Keerom, a very high proportion of trans-migrant population in an early selected sub-district (Skanto) led the field teams to decide to not continue implementing data collection in this sub-district. This sub-district was dropped from the analysis, which means that we have included a sample of 14 villages from Keerom in our analysis.
- In Waropen, early low proportions of eligible households per village and non-accessibility of other villages led the team to sample and visit two additional villages in this district. In total, the sample from Waropen is therefore distributed across 22 villages.

[36] For the purposes of this study, we are defining households as a group of individuals who share a common cooking place (i.e. kitchen and 'budget' for food) and who identify one common person as the household head.

- In Asmat, the total number of sampled villages visited that were actually within the selected sub-districts were 19, reducing in effect the final number of sampled villages there.

Table 2 presents target and achieved sample sizes for this survey in detail.

A.2.3 Sampling weights and survey settings in analysis

The sampling approach explained above implies that – in effect – sampling was purposeful in the first two steps of selecting districts and sub-districts within which the survey was implemented.

Steps 3 and 4, however, involved a random sampling approach that allows us to calculate the probability of each household or children being sampled and – therefore – also allows us to calculate sampling weights for children and households. Using these weights allows us to produce estimates that are representative of the population of eligible households and children in accessible villages in the sub-districts listed in Table 13.

Household-level weights

Household-level weights are based on a combination of two probabilities: the probability of a village being sampled and the probability of a household within the village being sampled. In our survey, the probability of villages being sampled can be defined as follows:

$$(1)\ p_i^{PSU} = \frac{n_i^{PSU}}{N_i^{PSU}},$$

where n_i^{PSU} is the number of PSUs (villages) sampled per district i , $N_i^{PSU}$is the total number of PSUs on the list of PSUs from which the sample is drawn, and p_i^{PSU} is the resulting probability of each village to be sampled. As explained above, our sampling approach was to sample 20 PSUs from the list of accessible PSUs in our sub-districts. However, for the purpose of creating weights, we adapted n_i^{PSU} for three districts (Asmat, Keerom, and Waropen) so that it reflected the actual number of villages that were sampled and used in this baseline analysis per district, i.e. to 22 for Waropen, 19 for Asmat, and 14 for Keerom (see above).

The probability of households being sampled – within villages – can be defined as follows:

$$(2)\ p_j^{HH} = \frac{n_j^{HH}}{N_j^{HH}},$$

where n_j^{HH} is the number of households sampled per PSU j, N_j^{HH} is the total number of eligible households listed per PSU, and p_j^{HH} is the resulting probability of each household being sampled within PSU j. In general, we sampled 15 households per PSU, except in cases where there were less than 15 eligible households present in the PSU. In those cases, all eligible households were sampled.

The above two probabilities defined in (1) and (2) can then be combined to create household-level weights as follows:

$$(3)\ w_{ij}^{HH} = \frac{1}{p_i^{PSU} * p_j^{PSU}}.$$

Here, w_{ij}^{HH} is the resulting weight for households in PSU j within district i. Note that within each PSU, all households have the same weights.

Child-level weights

In some households, there was more than one eligible children present. Hence, for child-level analyses, we also included child-level weights that expand weights as defined in (3) by the inverse of the probability of each child being sampled in their households. For households with just one eligible child, this probability is one. The resulting weights can be defined as follows:

$$(4)\ w_{ijk}^{CH} = \frac{1}{p_i^{PSU} * p_j^{HH} * p_k^{CH}}.$$

This is the weight for a child in household k, sampled in PSU j within district i, and p_k^{CH} is the probability of a child being sampled in household k.

Rescaling and truncating weights

We rescaled both child- and household-level weights so that they sum up to the actual number of observations in our sample and truncated them in order to prevent very large weights skewing our analysis. We present the estimates for key indicators in Papua both including weights and excluding weights in Annex B.2.

A.3 Papua baseline survey: training, piloting, data collection, and data quality assurance

A.3.1 Overview

This section of the annex describes the baseline data collection for the purpose of this evaluation in Papua. It provides details on the fieldwork team, the fieldwork model deployed across the study districts, the data collection timeline, and the data quality assurance protocol.

OPML contracted the Indonesian survey firm Myriad as its fieldwork partner to conduct the baseline survey in three treatment districts – Asmat, Lanny Jaya, and Paniai – and three control districts – Boven Digoel, Keerom, and Waropen – of Papua. The fieldwork involved the collection of data from a sample of households who had a self-identified indigenous Papuan parent or caregiver of a child aged 0–23 months. The household interviews aimed to collect data around household background socioeconomic and demographic characteristics, food and non-food consumption, as well as child nutrition and early childhood development. In addition, anthropometric measures (i.e. height and weight) of children aged 0–23 months were collected to assess malnutrition in the population of the targeted children.

A.3.2 Fieldwork team

Considering the difficulty of doing data collection in remote districts in Papua, OPML and Myriad prioritised recruitment of enumerators and anthropometric specialists from local communities where the survey was implemented. Myriad made use of their

existing network of enumerators in these areas and developed screening criteria to hire competent enumerators, anthropometric specialists, and supervisors in each of the districts.

Specifically, Myriad used the following criteria for the recruitment:

- Applicant must either be an indigenous Papuan or have been born in Papua.
- Applicant must have the ability to communicate with local people in local languages.
- Applicant must be at least a college student in the 7th semester.
- Applicant should ideally have experience in conducting surveys.
- Applicant must be familiar with smartphones or tablets.

During the recruitment, the team focused on recruiting female applicants and was successful in recruiting 30% female enumerators in the data collection team. In addition to the above criteria, there were specific requirements for anthropometric specialist applicants:

- Applicant should ideally be female.
- Applicant should ideally have a background in nutrition or health.
- Applicant should ideally have experience in doing anthropometric measurements.

The fieldwork team was finalised after an OPML and Myriad training course for the data collection exercise.

A two-week enumerator training course for the listing exercise and the household and child/caregiver questionnaires was held in Jayapura from 15 to 27 October 2018.

The training consisted of three differentiated components: listing training, household and child/caregiver questionnaire training, and anthropometric training. All three components included both in-class sessions and piloting exercises to help the team to familiarise themselves with the data collection tools and to test the planned logistics. **For more information on the training exercise, refer to Annex G of the Inception Report** (OPML, 2018).

A total of 72 participants took part in the training. Out of them, 63 were selected into the final field team.

The final team hired to undertake baseline data collection activities consisted of 12 supervisors, 39 enumerators, and 12 anthropometric specialists.

A.3.3 Fieldwork model

Considering the logistical constraints of doing data collection in remote communities, such as lack of transport, inaccessible terrain, security issues, and general distrust of survey activities in local communities, the fieldwork model was carefully planned to address these challenges.

Figure 55 illustrates the fieldwork model deployed across the six districts in Papua.

Within each district, the enumerators and anthropometric experts were organised into two teams. Team composition varied across districts. Each team consisted of one supervisor, one anthropometric specialist, and three to four enumerators. The supervisor was responsible for overall team coordination, including travel, accommodation, sensitisation, and introduction to local communities. The supervisor also ensured data quality and kept track of the interview completion rate to achieve the daily and weekly targets set by field coordinators.

Each district was assigned a field coordinator who oversaw both teams and received daily feedback from team supervisors. Field coordinators liaised with the Myriad data manager and OPML, and provided feedback to the on-the-ground teams as part of the quality assurance protocol. They also reviewed reasons for potential PSU replacement and communicated these upwards to the Myriad data manager and team leader in Jakarta. In addition, field coordinators ensured the regular uploading of the household data for continuous monitoring.

Each of the six field coordinators reported to the Myriad survey manager, who was stationed in Jayapura for the entire duration of the data collection. The survey manager was responsible for the overall fieldwork implementation, coordination, and logistics planning across all six districts.

The Myriad data manager and team leader were OPML's main points of contact within the Myriad team. They liaised with OPML on a regular basis and checked on data quality, enumerator performance, and PSU completion rate. They also coordinated the replacement of PSUs, when original PSUs were deemed inaccessible.[37]

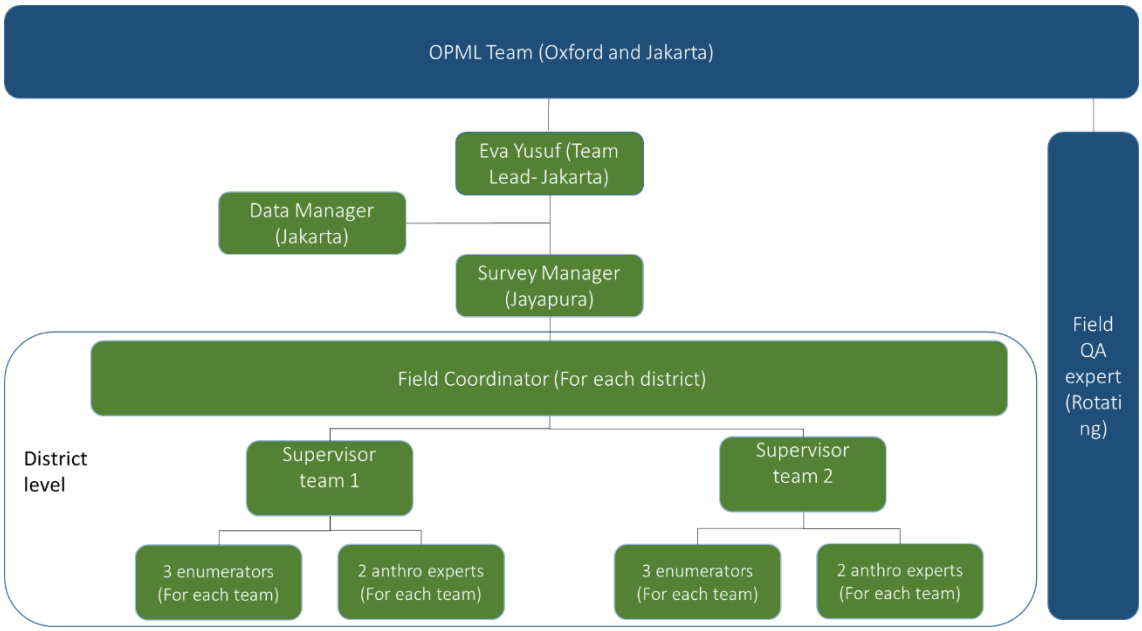
In addition, OPML had also hired an independent field quality assurance expert who floated across teams and districts to monitor the quality of data collection and reported regularly to the OPML team using a monitoring checklist.

[37] Replacement of village/sub-district and related PSUs happened only in the following scenarios:

- Sampled village was not inhabited or the majority of the community was temporarily absent;
- Sampled village was inaccessible due to natural obstacles, collapse of transport routes, or due to security concerns (such as tribal warfare, militancy, or political instability); and
- Majority of the population in the sampled village was not Papuan.

The OPML survey manager supervised the replacement process and made the final decision. The list of replacements was provided by the OPML team, using the sampling and replacement protocol agreed prior to the start of the data collection phase and as described in the Inception Report (OPML, 2018).

Figure 55: Field team model



A.3.4 Data collection timeline

Data collection was rolled out on 1 November and continued until 11 December. Data collection finished before the disbursement of the CG to households.

Each data collection team was assigned 10 PSUs to be completed during the full duration of the survey. The supervisors and field coordinators developed a detailed workplan that outlined the schedule of data collection in each of the assigned 10 PSUs. The survey plan took into consideration travel time, time to get local permissions, and time for community sensitisation. The data collection route was planned such that each team started out from the villages close to the capital and then proceeded to more remote parts of the district.

A.3.5 Fieldwork implementation

Fieldwork implementation involved four steps: Introduction, listing, sampling, and the survey.

Step 1: Introduction

Once the team entered a sampled village, the supervisor met with the village leaders, provided an introduction to the study, presented the permission letters, and provided an overview of the data collection process.

Step 2: Selection of PSU and identification and listing of eligible households

PSUs were randomly selected from the list of RWs and RTs in the village, as identified by the listing team and the village leaders. Annex A.2 provides more information on the

sampling strategy. In each PSU, each and every residential structure was listed and Papuan households with children under the age of two were identified.[38]

Step 3: Sampling of households

Supervisors drew a random sample of households using systematic random sampling, based on the list of households produced during the listing exercise. A total of 15 main sample households and 10 replacement households was sampled for each PSU.

At the end of the data collection, 6% of the sampled households were replaced due to refusals, targeting errors (i.e. no eligible children in the household), or household members not being available up to the third visit. Table 14 presents the breakdown by district:

Table 14: Percentage of replacement by district

District	Percentage of replacement
Asmat	3%
Boven Digoel	12%
Keerom	4%
Lanny Jaya	1%
Paniai	14%
Waropen	2%

Overall, the target sample size for the baseline survey was 1,440 children and their households. See Table 2 for more details on the final sample achieved.

Step 4: Data collection

Data collection started straight after the listing exercise. A fieldwork report describing the data collection, including challenges and risk mitigation strategies, is available on request.

A.3.6 Quality assurance (QA) protocols

In addition to an extensive pre-test, training and pilot, our field data QA model was implemented at five different and complementary levels:

- **In-built consistency checks in the CAPI questionnaire**: the evaluation team built and tested a series of consistency checks in the CAPI tool to minimise data entry errors from enumerators. These include, for example, consistency checks of household roster information, uniqueness of household head, skips within and across different questionnaire modules, and flags for particular out-of-range values.
- **Spot checks** were conducted by field supervisors to assess how survey protocols are implemented and identify and retrain enumerators who require additional support.

[38] 'Residential structure' excludes hostels, hospitals, hotels, and army or police barracks

- **Central dashboard**: OPML developed a data QA system written in Stata and displayed using an online available visualisation tool called Power Bi in order to: i) monitor data collection progress and ensure teams are being deployed effectively; ii) check interviewers' performance on key survey modules and variables; and iii) identify inconsistencies or uncommon patterns in individual interviews to provide timely training and feedback. The main objective of this system was to provide timely feedback and to implement course-correction processes that significantly improve data quality. This dashboard was updated as soon as data were uploaded to the server and insights from this system were fed back to the survey team on the ground. This system was also used to make corrections and do preliminary data cleaning before the preparation of the final dataset.

- **WhatsApp chat groups**: OPML created one WhatsApp group per district to relay feedback from the dashboard to the teams. The feedback focused on specific enumerators and described the issue in detail. For instance: 'Hi Iranti, on 04/11 you visited household ID 20280111. The child you measured is 3 months old but she is only 6.2 cm tall. Can you double check please?' The responses on the WhatsApp group were used to do the initial data cleaning.

- **Weekly reports from OPML's field QA expert**: OPML hired and trained an experienced fieldwork supervisor to do spot checks across the six districts and report back to the Jakarta and Oxford teams on data quality issues. The fieldwork QA expert was provided with a checklist of items focused on introductions, listing protocol, and survey protocols against which he reviewed enumerators' performance. The feedback from this weekly report was communicated with the Myriad team and monitored via dashboard.

A.4 Papua baseline survey: data collection instrument

This section describes the data collection instrument in more details, drawing upon OPML Inception Report (OPML, 2018).

Table 15 below presents the survey modules and related indicators that each module addresses. The questionnaire sections and outcome indicators are mapped against their impact domain, as specified by the Terms of Reference. The table also shows from which sources questions (or modules) will be drawn, in order to ensure consistency between this survey and other data sources (such as, for example, SUSENAS and MICS).

The questionnaire covers the following impact domains: consumption and monetary poverty, access to education, child health and nutrition (including food security, dietary diversity, breastfeeding, stunting, and wasting), birth registration, and early childhood development. It should be noted that modules investigating child-level outcomes will only be administered to one eligible child (randomly selected through CAPI) per household.

Selected key outcomes include those impact domains that are expected to be affected by the programme in the short and medium term.

The full questionnaire can be provided upon request.

Table 15: Survey modules and related indicators

Impact domain	Outcome indicators	Survey module	Module based on
(Not applicable)	(Not applicable)	**COVER AND INTRODUCTION** - Enumerator's introduction to household - Consent form - Filling out interview details - Questionnaire transition (from household questionnaire to mother/caregiver/ and child anthropometric measurement)	**OPML HH surveys**
(Not applicable)	(Not applicable)	**HOUSEHOLD ROSTER:** identification of household members and collection of key demographic information on each one of them	**OPML HH surveys**
(Not applicable)	(Not applicable)	**SOCIAL PROTECTION INFORMATION:** identification of social protection programmes in which the household participates (PKH, Raskin, etc.)	**SUSENAS**
Consumption and Monetary Poverty	• Monthly total expenditure per adult equivalent, adjusted for regional price differences • Monthly food expenditure per adult equivalent, adjusted for regional price differences • *Per capita* monthly health expenditure • Per child monthly education expenditure • Food share of monthly consumption expenditure (% of food expenditure out of total monthly expenditure) • Percentage of people below poverty line • Poverty gap (based on monthly total expenditure per adult equivalent) • Food poverty gap	**HOUSEHOLD CONSUMPTION AND EXPENDITURE** - Food consumption module - Non-food consumption module	**SUSENAS consumption questionnaire**
Access to Education	• % of children (0–10 years) currently attending / have ever been to pre-school • % of children (above 5 years) attending school	**HOUSEHOLD ROSTER**	**OPML HH surveys**
Child Health and Nutrition (meal	• Minimum acceptable diet: percentage of children aged 6–23 months who had at least the minimum dietary diversity[39] and the	**CHILD HEALTH AND NUTRITION**	**MICS and RISKESDAS**

[39] Minimum dietary diversity: Proportion of children aged 6–23 months who receive foods from four or more food groups. The seven foods groups used for calculation of this indicator are:
— grains, roots and tubers
— legumes and nuts
— dairy products (milk, yogurt, cheese)

frequency and dietary diversity)	minimum meal frequency[40] during the previous day ○ breastfed children (breastfed children aged 6–23 months who had at least the minimum dietary diversity and the minimum meal frequency during the previous day) ○ non-breastfed children (non-breastfed children aged 6–23 months who received at least two milk feedings and had at least the minimum dietary diversity not including milk feeds and the minimum meal frequency during the previous day) • Exclusive breastfeeding: percentage of infants under six months who are exclusively breastfed[41]		
Child Malnutrition (prevalence of stunting and wasting)	• Stunting prevalence: percentage of children under age two who fall below (a) -2 SD (moderate and severe) (b) below -3 SD (severe) of the median height for age of the WHO standard • Wasting prevalence: percentage of children under age two who fall below (a) -2 SD (moderate and severe) (b) -3 SD (severe) of the median weight for height of the WHO standard	**CHILD HEALTH AND NUTRITION**	**MICS**
Birth registration	• Percentage of children under age 17 who have a birth certificate from the civil registry office	**HOUSEHOLD ROSTER**	**OPML HH surveys**
Early Childhood Development	• Availability of children's books: Percentage of children under age two who have three or more children's books available in their household • Availability of playthings: Percentage of children under age two who have two or more types of plaything available in their household • Inadequate supervision: Percentage of children under age two left alone or under the supervision of another child younger than 10 years for more than one hour at least once in the last week • Violent discipline: Percentage of children aged 1–2 years who experienced any physical punishment and/or psychological aggression by caregivers in the past one month	**EARLY CHILDHOOD DEVELOPMENT**	**MICS and RISKESDAS**

A.5 Child malnutrition estimates in Papua: rapid data quality assessment

One concern with anthropometric measurements and their use to estimate the prevalence of malnutrition in children is the quality of these measurements.

— flesh foods (meat, fish, poultry and liver/organ meats)
— eggs
— vitamin-A rich fruits and vegetables
— other fruits and vegetables.

[40] Minimum meal frequency: Proportion of breastfed and non-breastfed children aged 6–23 months who receive solid, semi-solid, or soft foods (but also including milk feeds for non-breastfed children) the minimum number of times or more. For breastfed children, minimum is defined as two times for infants aged 6–8 months and three times for children aged 9–23 months. For non-breastfed children, minimum is defined as four times for children aged 6–23 months.

[41] Infants receiving breast milk, and not receiving any other fluids or foods, with the exception of oral rehydration solution, vitamins, mineral supplements, and medicines.

Measurements of poor quality can bias estimates. Following the standard procedures suggested by the WHO, we plot two key data quality indicators for anthropometric data in Table 16 below:

- The proportion of outliers in Z-scores. Outliers are values of Z-scores that are physically implausible. A high proportion of outliers indicates that many measurements are of poor quality.
- The SD of Z-scores. SDs generally fall within a narrow range when comparing different studies that try to assess child malnutrition (see WHO, p. 218 and here). High SDs can indicate poor quality of anthropometric measurements.

The results in Table 16 below indicate that the proportion of outliers both for our sample overall and for the sub-sample of Asmat (where the proportion of wasted children is highest) is low. The SDs are within the ranges found in other studies, although on the higher bounds (see here, p. 445). These results indicate that there is no evidence for significant quality concerns with respect to the anthropometric data collected in this survey.

Table 16: Child malnutrition data quality table

Z-score	Outliers as % of total	SD
Overall		
WLZ	1.3%	1.5
WAZ	0.5%	1.4
HAZ	2.2%	1.9
Asmat sub-sample only		
WLZ	1.0%	1.5
WAZ	1.7%	1.6
HAZ	2.4%	2.0
Source: OPML Baseline Data – Child Grants Impact Assessment in Papua. Unweighted estimates		

Annex B Full statistical tables

B.1 Papua – estimates for selected indicators

	TREATMENT					CONTROL					P-value T vs C *=p<0.1,**=p<0.05, ***=p<0.01
Indicator	N	Estimate	SE	Low CI	Upper CI	N	Estimate	SE	Low CI	Upper CI	
Household-level indicators											
Distance of household to the district capital	779	26.1	2.9	20.4	31.8	672	28	2.6	23	33.1	0.621
Household size	790	4.8	0.1	4.6	4.9	683	5.7	0.1	5.5	6.0	0***
Age dependency ratio	753	110.4	2.8	104.9	115.8	641	124.1	2.7	118.7	129.5	0.001***
Household head has ever attended school	790	87.1	1.4	84.3	89.9	683	95.8	0.8	94.3	97.4	0***
Caregiver is the biological mother	790	91.7	1	89.7	93.6	683	93.6	0.8	92.1	95.1	0.121
Caregiver is female	790	93.9	0.8	92.3	95.6	683	97.2	0.8	95.7	98.7	0.004***
HH received cash support in past 12 months	774	43.5	2.8	37.9	49	680	10.5	1.4	7.7	13.3	0***
Monthly real consumption expenditure per adult equivalent (IDR)	789	1551968	49188.1	1454479	1649457	683	1929430	72604.0	1785531	2073329	0***
Monthly real food consumption expenditure per adult equivalent (IDR)	789	1161969	41064.4	1080581	1243358	683	1317399	46123.6	1225984	1408815	0.013**
Monthly household real consumption expenditure (IDR)	789	4654181	143751.7	4369270	4939092	683	6863696	206705.2	6454013	7273379	0***
Monthly household real food consumption expenditure (IDR)	789	3476942	118562.9	3241954	3711930	683	4674063	127720	4420926	4927200	0***
Monthly *per capita* nominal health expenditure (IDR)	789	2266	350.2	1572	2960	683	5976	664.9	4658	7294	0***
Monthly education expenditure per child [0–17] attending school in the last 12 months (IDR)	359	47078	15146.0	17059	77097	418	54676	11450.5	31987	77371	0.690
Proportion of food consumption expenditure over total consumption expenditure	789	73.2	0.6	71.9	74.4	683	69.3	0.6	68.1	70.6	0***
Poverty headcount	789	19.3	2.2	15	23.7	683	14.8	2.4	10.1	19.6	0.170
Poverty gap	789	5.3	0.7	3.9	6.7	683	4.1	0.9	2.4	5.8	0.284
Food poverty headcount	789	10.5	1.5	7.4	13.5	683	7.1	1.7	3.7	10.6	0.152
Food poverty gap	789	2.3	0.4	1.6	3	683	1.9	0.6	0.8	3.0	0.524
Eligible children-level indicators											
Age of child in months	790	11.9	0.3	11.4	12.5	683	11.2	0.2	10.8	11.7	0.043**
Child is female	790	47.1	1.9	43.3	50.8	683	52	1.6	48.8	55.3	0.050**
Child with minimum dietary diversity [eligible children aged 6–23m]	631	20.3	2	16.4	24.2	511	27	2.1	22.8	31.2	0.023**
Child with minimum meal frequency [eligible children aged 6–23m]	520	53.1	3	47.1	59.1	461	53.8	2.5	48.8	58.9	0.853
Child with minimum acceptable diet [eligible children aged 6–23m]	612	12.5	1.7	9.1	15.9	502	13	1.4	10.3	15.8	0.797
Child has birth certificate (seen)	773	15.9	2.4	11.1	20.6	682	13.4	1.4	10.6	16.1	0.369
Child has three or more children's books	790	2.3	0.5	1.3	3.3	683	1.6	0.5	0.6	2.7	0.328
Child has two or more types of plaything	790	25.8	1.9	22.2	29.5	683	32.9	1.8	29.3	36.4	0.007***
Child has manufactured toys	790	31.3	2.1	27.2	35.4	683	45	2.7	39.6	50.3	0***
Child left alone in past week	481	88.1	2	84.1	92.1	514	64.1	3	58.1	70.1	0***
Child experienced violence in past month [eligible children aged 12–23m]	407	42.1	2.4	37.3	46.8	303	69	2.9	63.3	74.8	0***
Child experienced psychological violence [eligible children aged 12–23m]	407	28.6	2.6	23.5	33.7	303	60.4	2.9	54.7	66.1	0***

Child experienced physical violence [eligible children aged 12–23m]	407	33.7	2.3	29.2	38.3	303	54.2	2.8	48.8	59.7	0***
Child is underweight	781	28.3	2.1	24.1	32.6	682	19.6	1.6	16.3	22.8	0.001***
Child is wasted	772	20.8	1.8	17.2	24.4	677	17.5	1.5	14.5	20.6	0.169
All children-level indicators											
Child is currently attending pre-school [children aged 0–10y]	1724	25.8	1.6	22.7	28.9	1704	18.5	1.6	15.3	21.7	0.002***
Child is currently attending school [children aged 5–17y]	886	31.9	2	27.9	35.9	1057	60.3	2.1	56.1	64.4	0***
Child attended school but stopped now [children aged 5–17y]	886	7.1	1	5.1	9.1	1057	5.7	0.9	4	7.5	0.307
Child has birth certificate (seen) [children aged 0–17y]	2000	15	2	11	19	2131	25.9	1.9	22.1	29.7	0***

B.2 Papua – unweighted estimates for selected indicators

The below table reproduces indicator estimates from Annex B.1 above, without using household and child-level sampling weights.

	TREATMENT					CONTROL					P-value T vs C *=p<0.1,**=p<0.05, ***=p<0.01
Indicator	N	Estimate	SE	Low CI	Upper CI	N	Estimate	SE	Low CI	Upper CI	
Household-level indicators											
Distance of household to the district capital	779	27.4	3	21.5	33.2	672	29.2	2.4	24.5	33.9	0.636
Household size	790	4.8	0.1	4.6	4.9	683	5.7	0.1	5.5	5.9	0***
Age dependency ratio	753	111.4	2.4	106.6	116.2	641	122.4	2.6	117.3	127.5	0.002***
Household head has ever attended school	790	86.2	1.5	83.2	89.2	683	95.9	0.7	94.5	97.3	0***
Caregiver is the biological mother	790	90.8	1	88.9	92.7	683	93.7	0.7	92.3	95.1	0.15
Caregiver is female	790	93	0.8	91.4	94.7	683	97.5	0.7	96.1	98.9	0.0***
HH received cash support in past 12 months	774	45.9	2.6	40.7	51	680	10.7	1.4	8	13.5	0***
Monthly real consumption expenditure per adult equivalent (IDR)	789	1587845	49567	1489604	1686086	683	1962107	59648	1843886.1	2080328	0***
Monthly real food consumption expenditure per adult equivalent (IDR)	789	1185497	41621	1103005	1267988	683	1342659	37815	1267712	1417606.1	0.006***
Monthly household real consumption expenditure (IDR)	789	4758005	143272	4474044	5041965	683	6928376	190644	6550525	7306226	0***
Monthly household real food consumption expenditure (IDR)	789	3545679	119506	3308822	3782536	683	4725788	117909	4492098	4959479	0***
Monthly *per capita* nominal health expenditure (IDR)	789	2603.5	430.8	1749.7	3457.3	683	6014.2	598.4	4828.3	7200.2	0***
Monthly education expenditure per child [0–17] attending school in the last 12 months (IDR)	359	51222	17735.9	16070	86374.1	418	56648.7	13167.5	30551.1	82746.3	0.806
Proportion of food consumption expenditure over total consumption expenditure	789	73	0.6	71.9	74.2	683	69.4	0.6	68.3	70.6	0***
Poverty headcount	789	17.1	1.7	13.7	20.4	683	14.2	2	10.2	18.2	0.276
Poverty gap	789	4.4	0.5	3.3	5.5	683	3.7	0.7	2.4	5	0.415
Food poverty headcount	789	8.3	1.1	6.1	10.5	683	6.4	1.3	3.7	9.1	0.293
Food poverty gap	789	2	0.3	1.3	2.6	683	1.6	0.4	0.8	2.4	0.459
Eligible children-level indicators											
Age of child in months	790	12	0.2	11.5	12.4	683	11.3	0.2	10.8	11.8	0.043**
Child is female	790	48.1	1.6	45	51.2	683	53.1	1.5	50.1	56.2	0.024***
Child with minimum dietary diversity [eligible children aged 6–23m]	631	21.4	2	17.5	25.3	511	27	2	23	31	0.049**
Child with minimum meal frequency [eligible children aged 6–23m]	520	52.1	2.9	46.3	57.9	461	53.4	2.5	48.4	58.4	0.747
Child with minimum acceptable diet [eligible children aged 6–23m]	612	12.9	1.8	9.4	16.5	502	12.9	1.4	10.2	15.7	0.986
Child has birth certificate (seen)	773	14.1	1.7	10.7	17.5	682	13.8	1.5	10.7	16.8	0.891

Child has three or more children's books	790	2.4	0.5	1.4	3.4	683	1.3	0.4	0.5	2.1	0.095*
Child has two or more types of plaything	790	27.5	1.6	24.3	30.7	683	32.9	1.7	29.7	36.2	0.019
Child has manufactured toys	790	32.9	1.8	29.4	36.4	683	45.5	2.5	40.6	50.4	0***
Child left alone in past week	481	87.5	1.9	83.7	91.4	514	64.6	2.9	58.9	70.3	0***
Child experienced violence in past month [eligible children aged 12–23m]	407	41.5	2.3	37.1	46	303	69	2.7	63.7	74.3	0***
Child experienced psychological violence [eligible children aged 12–23m]	407	28.5	2.2	24	33	303	59.7	2.6	54.5	65	0***
Child experienced physical violence [eligible children aged 12–23m]	407	33.4	2.2	29.1	37.7	303	55.4	2.7	50	60.9	0***
Child is underweight	781	27	2	23	31.1	682	20.1	1.4	17.3	22.9	0.006***
Child is wasted	772	20.1	1.7	16.6	23.5	677	17.9	1.3	15.2	20.5	0.317
All children-level indicators											
Child is currently attending pre-school [children aged 0–10y]	1724	25.3	1.5	22.3	28.3	1704	19.4	1.5	16.3	22.4	0.007***
Child is currently attending school [children aged 5–17y]	886	32.3	1.8	28.7	35.9	1057	60.8	2	56.9	64.8	0***
Child attended school but stopped now [children aged 5–17y]	886	7.1	1	5.2	9	1057	5.8	0.9	3.9	7.6	0.323
Child has birth certificate (seen) [children aged 0–17y]	2000	13.9	1.7	10.5	17.2	2131	25.8	2.1	21.6	30	0***

B.3 Aceh

	TREATMENT					CONTROL					P-value T vs C *=p<0.1,**=p<0.05, ***=p<0.01
Indicator	N	Estimate	SE	Low CI	Upper CI	N	Estimate	SE	Low CI	Upper CI	
Household-level indicators											
Household size	118	4.5	0.2	4.1	5	2127	4.8	0	4.7	4.8	0.372
Age dependency ratio in HH	118	96.6	8.3	80.3	112.9	2127	92.4	1.5	89.5	95.2	0.614
Household head is a male	118	96	1.9	92.2	99.8	2127	90.6	0.8	89	92.3	0.011**
HH received Rastra/Raskin (past four months)	118	51.9	8.7	34.9	69	2127	56.1	1.9	52.5	59.8	0.636
HH received PKH (ever)	118	4.4	1.7	1.1	7.8	2127	19	1.2	16.6	21.4	0***
HH received financial support from local government (past 12 months)	118	71.9	7.7	56.8	86.9	2127	8.6	0.7	7.3	10	0***
Monthly household real consumption expenditure (IDR)	118	4801222	419235.4	3977711	5624734	2127	4919407	108377.5	4706520	5132295	0.785
Monthly household real food consumption expenditure (IDR)	118	2564853	225020.3	2122842	3006865	2127	2929480	52065.2	2827208	3031753	0.115
Monthly real consumption expenditure per adult equivalent (IDR)	118	1493305	128669.5	1240557	1746053	2127	1465093	30660.5	1404866	1525320	0.831
Monthly real food consumption expenditure per adult equivalent (IDR)	118	799137	66961.5	667604	930671	2127	865369	14368.0	837146	893592	0.334
Proportion of food consumption expenditure over total consumption expenditure	118	56.3	1.2	54	58.7	2127	62.7	0.4	61.9	63.5	0***
Eligible children-level indicators											
Age of child in months	146	28.2	1.2	25.8	30.6	2545	30	0.4	29.2	30.9	0.162
Child is female	146	54	5.9	42.4	65.7	2545	49.1	1.3	46.6	51.6	0.417
Child is biological/stepchild of household head	146	89.8	3.6	82.6	96.9	2545	80.6	1.3	78	83.1	0.017**
Child has birth certificate (seen)	146	68.3	8.4	51.8	84.9	2543	66.9	1.5	63.9	69.9	0.867
Child has three or more children's books [children aged 0–4y]	120	33.9	5.7	22.6	45.2	2077	19.3	1.3	16.8	21.8	0.013**

Child has two or more types of plaything [children aged 0–4y]	118	57.1	7.7	42	72.1	2074	67.4	1.4	64.6	70.2	0.187
Child left alone in past week [children aged 0–4y]	119	6	2.1	1.9	10.2	1926	14.7	1.3	12.3	17.2	0***
Child experienced violence in past month [children aged 1–4y]	94	53	8.5	36.3	69.6	1639	52.2	1.9	48.5	56	0.935
All children-level indicators											
Child is currently attending pre-school [children aged 0–4y]	120	10.3	3.8	2.8	17.7	2090	8.9	0.7	7.5	10.3	0.719
Child is currently attending school [children aged 5–17y]	155	75	3.2	68.7	81.3	2673	79.8	0.9	78.1	81.6	0.121