Rural home-grown school feeding program in Nyanda

The analysis of the 2023-2025 education dataset of students in grade 1-6 in Nyanda shows that there are particularly persistent gaps among students in education attainment based on households' incomes, rural residence, attendance, and to certain extent basic parental education. Additionally, variation in students' performance seems to derive from within schools and not between schools or between districts variation.

A home-grown school feeding program targeting rural communities in Nyanda is an effective educational reform aiming at reducing learning inequalities associated with households' income, rural residence, and attendance.

Evidence shows[1] that HGSFP by improving child nutrition and supporting local food systems can directly enhance learning equity by creating stronger incentives for higher attendance. At first, providing regular nutritious meals based on the national diet impact student concentration, cognitive development and overall ability to learn. These benefits are particularly important for children from poor and food insecure households. Second, school meals have been shown to increase school attendance and reduce dropout rates, because school meals provide a tangible incentive for parents to send their kids to school. Third, by sourcing food from the local community it favors the development of the local economy, strengthening rural livelihoods by channeling important resources to local communities. Finally, targeting rural communities often translates in low-income communities, ensuring that intervention reach the most vulnerable groups, where gains in inequities are higher.

The available education data for Nyanda allows us to predict the expected test scores results for both reading and mathematics and for any income group given an increase in attendance. Currently test scores for the last two income groups are between approximately -0.3 and – 0.1 standard deviations the national mean for the lowest two income quantiles. A meta-analysis of the impact of school feeding on children health, nutrition, and educational outcome, shows that school feeding results in a significant increase in the percentage of school days attended (2.6%; 95% CI = 1.2%, 3.9%; $P < 0.001$) (Dongging, et al. 2021). Therefore, given our data, it is realistic to assume a 0.3 standard deviation increase in attendance due to the introduction of school feeding, which translates into a 2.2% increase in attendance, or 4.2 days in a school year. This prediction is in line with the literature and therefore plausible to be expected. If this was to be realized, the student in the second to lowest income group would pass the national average, while the student in the lowest income group would benefit enormously but still be approximately -0.18 standard deviation below the national average.

This conclusion was drawn from a predictive analysis performed with the education data for Nyanda. The prediction was made from a multilevel model that included a random intercept for school, district, and student-level variance, for both reading and math test scores.

[1] WFP **School-based Programmes Impact Evaluation Window**

In the second step, the dataset was modified to simulate an improvement in attendance by 0.3 standard deviation, applied only to students in the lower income quantile (1 and 2) and living in rural settings. Third, the updated dataset was used to predict the new reading and math outcomes.

In the final stage, the results were aggregated by income quantile to showcase how reading and math outcomes would change based on an increase in attendance. The percentage increase in attendance was obtained multiplying the standard deviation of attendance rate by the expected standard deviation improvement in attendance thanks to the introduction of the school feeding program. While the increase in attendance days was derived by multiplying this percentage by an approximate number of school days per year (190).

Bibliography

Barro, D., C. Bogaards, P. Christian, E. Kelley, R. Khincha, F. Kondylis, M.P La, et al. 2025. "Impact Evaluation of the Home-Grown School Meals Programme in The Gambia. World Food Programme Office of Evaluation." OEV/2022/038.

Dongging, Wang, Shinde Sachin, Young Tara, and Fawzi Wafaie W. 2021. "Impacts of school feeding on educational and health outcomes of school-age children and adolescents in low- and middle-income countries: A systematic review and meta-analysis." *Journal of Global Health* doi: 10.7189/jogh.11.04051.

Appendix

```
#______________________
# b) Multilevel model with random intercept for school, district, and student
-level variance
#______________________
# Running this multilevel model to account for:
# 1. school variance - differences between schools
# 2. district variance - differences between districts
# 3. student-level variance - differences between students withing schools

# load the data
merged_centered = file.path(out_data_fld,"df_merged_centered.xlsx")
df_merged_centered = read_excel(merged_centered)

# transform the data in numeric
df_merged_centered <- df_merged_centered %>%
  mutate(
    reading_z = as.numeric(reading_z),
    math_z = as.numeric(math_z)
  )

# Multilevel model on reading z-scores
mixed_reading <- lmer(
  reading_z ~ factor(household_income_quintile) + female + rural +
    disability_status + parent_education_level + attendance_centered  +
    (1 | school_id) + (1 | district_id) + factor(year),
  data = df_merged_centered,
  weights = weight
)

## boundary (singular) fit: see help('isSingular')

# Summarize the results -READING
summary(mixed_reading)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerM
odLmerTest']
## Formula: reading_z ~ factor(household_income_quintile) + female + rural +
disability_status + parent_education_level + attendance_centered +      (1 |
school_id) + (1 | district_id) + factor(year)
##    Data: df_merged_centered
## Weights: weight
##
## REML criterion at convergence: 522823.4
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.7497 -0.6529  0.0008  0.6487  4.5279
##
## Random effects:
##  Groups      Name        Variance Std.Dev.
##  school_id   (Intercept) 0.0991   0.3148
##  district_id (Intercept) 0.0000   0.0000
##  Residual                0.9124   0.9552
## Number of obs: 192000, groups:  school_id, 9993; district_id, 25
##
## Fixed effects:
##                                        Estimate    Std. Error
df t value            Pr(>|t|)
## (Intercept)                           -0.267441      0.007122  73340.8935
18 -37.551 <0.0000000000000002 ***
## factor(household_income_quintile)2      0.168922      0.005463 191981.3145
51  30.919 <0.0000000000000002 ***
## factor(household_income_quintile)3      0.351070      0.006383 191819.8315
15  54.997 <0.0000000000000002 ***
## factor(household_income_quintile)4      0.532559      0.007972 191081.1308
34  66.800 <0.0000000000000002 ***
## factor(household_income_quintile)5      0.713616      0.012081 189548.3085
06  59.071 <0.0000000000000002 ***
## female                                  0.006998      0.004324 191976.9945
78   1.619               0.106
## rural                                  -0.255341      0.004534 191704.4287
50 -56.318 <0.0000000000000002 ***
## disability_statusYes                    0.017494      0.010723 191982.1533
51   1.631               0.103
## parent_education_levelPrimary           0.006538      0.005186 191985.7972
04   1.261               0.207
## parent_education_levelSecondary        -0.004582      0.005653 191984.5426
00  -0.811               0.418
## parent_education_levelTertiary         -0.009528      0.010074 191985.4359
38  -0.946               0.344
## attendance_centered                     0.817648      0.034595 187438.9993
91  23.635 <0.0000000000000002 ***
## factor(year)2024                        0.119088      0.005010 181880.4352
43  23.772 <0.0000000000000002 ***
```

```
## factor(year)2025                             0.239488      0.005016 182182.2507
49  47.747 <0.0000000000000002 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##
## Correlation matrix not shown by default, as p = 14 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it

## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```

```r
# Multilevel model on math z-scores
mixed_math <- lmer(
  math_z ~ factor(household_income_quintile) + female + rural +
    disability_status + parent_education_level + attendance_centered  +
    (1 | school_id) + (1 | district_id) + factor(year),
  data = df_merged_centered,
  weights = weight
)
```

```
## boundary (singular) fit: see help('isSingular')
```

```r
# summarize the results - MATH
summary(mixed_math)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerM
odLmerTest']
## Formula: math_z ~ factor(household_income_quintile) + female + rural +
disability_status + parent_education_level + attendance_centered +      (1 |
school_id) + (1 | district_id) + factor(year)
##    Data: df_merged_centered
## Weights: weight
##
## REML criterion at convergence: 523501.9
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.4622 -0.6491  0.0014  0.6541  4.6891
##
## Random effects:
##  Groups      Name        Variance       Std.Dev.
##  school_id   (Intercept) 0.103281091581 0.32137376
##  district_id (Intercept) 0.000000000325 0.00001803
##  Residual                0.914315170047 0.95619829
## Number of obs: 192000, groups:  school_id, 9993; district_id, 25
##
## Fixed effects:
##                                           Estimate    Std. Error
df t value             Pr(>|t|)
```

```
## (Intercept)                              -0.276347      0.007161  71899.5131
46 -38.590 < 0.0000000000000002 ***
## factor(household_income_quintile)2        0.185627      0.005473 191964.3620
41  33.918 < 0.0000000000000002 ***
## factor(household_income_quintile)3        0.354298      0.006394 191759.0179
35  55.409 < 0.0000000000000002 ***
## factor(household_income_quintile)4        0.541457      0.007985 190974.7465
46  67.810 < 0.0000000000000002 ***
## factor(household_income_quintile)5        0.718476      0.012098 189429.3374
29  59.388 < 0.0000000000000002 ***
## female                                   -0.002273      0.004331 191985.8175
93  -0.525              0.59979
## rural                                    -0.256621      0.004542 191793.4903
51 -56.495 < 0.0000000000000002 ***
## disability_statusYes                      0.018136      0.010742 191985.6562
62   1.688              0.09135 .
## parent_education_levelPrimary             0.016739      0.005195 191981.7223
63   3.222              0.00127 **
## parent_education_levelSecondary           0.002512      0.005663 191984.2457
78   0.444              0.65731
## parent_education_levelTertiary           -0.009176      0.010091 191983.0143
25  -0.909              0.36321
## attendance_centered                       0.816741      0.034640 187320.2491
97  23.578 < 0.0000000000000002 ***
## factor(year)2024                          0.121560      0.005015 181873.3197
00  24.239 < 0.0000000000000002 ***
## factor(year)2025                          0.244321      0.005021 182168.6967
07  48.658 < 0.0000000000000002 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##
## Correlation matrix not shown by default, as p = 14 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it

## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')

#______________________
# c) Prediction math and reading score given an increase in attendance due to
the proposed intervention
#______________________
# Define values for prediction
sd_increase = 0.3
nbr_school_days = 190

# Simulate an increase of 0.3 for low-income students living in rural setting
s
df_sim <- df_merged_centered %>%
```

```
  mutate(attendance_centered = if_else(low_income == 1 & rural == 1, # low in
come represents households in the first and second quantale
                                       attendance_centered + sd_increase,
                                       attendance_centered))

# Predict outcomes using the mixed model
df_sim$reading_pred <- predict(mixed_reading, newdata = df_sim, re.form = NA)
df_sim$math_pred <- predict(mixed_math, newdata = df_sim, re.form = NA)

# Aggregate predicted outcomes by income quantile
sim_summary <- df_sim %>%
  group_by(household_income_quintile) %>%
  summarise(mean_reading_pred = mean(reading_pred),
            mean_math_pred = mean(math_pred))

print(sim_summary)

## # A tibble: 5 × 3
##   household_income_quintile mean_reading_pred mean_math_pred
##                       <dbl>             <dbl>          <dbl>
## 1                         1          -0.182         -0.188
## 2                         2           0.00862        0.0189
## 3                         3           0.0984         0.0952
## 4                         4           0.303          0.305
## 5                         5           0.512          0.510

# mean and standard deviation of attendance
mean(df_merged_centered$attendance_rate, na.rm = TRUE)

## [1] 0.686

sd(df_merged_centered$attendance_rate, na.rm = TRUE)

## [1] 0.07350024

# presents results on school attendance

# % increase in attendance:
sd(df_merged_centered$attendance_rate, na.rm = TRUE)*sd_increase

## [1] 0.02205007

# Increase in attendance days thanks to school feeding. Assuming 190 days of
school:
sd(df_merged_centered$attendance_rate, na.rm = TRUE)*sd_increase*nbr_school_d
ays

## [1] 4.189514
```