Economic Clustering Summary Report: Burundi 2016-17

Andy Kim

14 September, 2023

# Summary of Data

**Country Code-year:** BU1617

**Number of observations:** 15731

**Number of clusters:** 5

**Number of variables used:** 31

**Distance used:** Hamming

**Variables used in the algorithm:** hv206, hv207, hv208, hv210, hv225, hv227, hv230b, hv232, hv232b, hv232y, hv237, hv243a, hv243b, hv244, hv246, hv246h, hv246i, hv247, hv252, sh121g, sh121h, sh121i, sh121j, sh121k, sh121l, water, toilet, floor, roof, cookfuel, wall

# Top Variable Groupings (Sorted by ASW)

| Group | ASW | Var. 1 | Var. 2 | Var. 3 | Var. 4 |
| --- | --- | --- | --- | --- | --- |
| 1 | 0.9592 | has electricity | has television | has mattress | roof |
| 2 | 0.951 | has electricity | has television | has cabinet | roof |
| 3 | 0.9491 | has electricity | has television | anything done to water to make safe to drink | roof |
| 4 | 0.9422 | items present: soap or detergent | items present: ash, mud, sand | items present: none | anything done to water to make safe to drink |
| 5 | 0.9419 | has television | items present: soap or detergent | items present: ash, mud, sand | items present: none |
| 6 | 0.9414 | has electricity | has television | anything done to water to make safe to drink | has cabinet |
| 7 | 0.9412 | has electricity | has television | has bank account | roof |
| 8 | 0.9412 | items present: soap or detergent | items present: ash, mud, sand | items present: none | has hoe |
| 9 | 0.9409 | has electricity | items present: soap or detergent | items present: ash, mud, sand | items present: none |
| 10 | 0.9399 | has television | has cabinet | has mattress | roof |

# Marginal Distributions

| Variable | Description | % Time in Top Clusters | Distribution |
| --- | --- | --- | --- |
| hv206 | has electricity | 60% | Binary, 12.1% 1s (or yes) |
| hv232 | items present: soap or detergent | 40% | Binary, 7.5% 1s (or yes) |
| hv208 | has television | 70% | Binary, 6.7% 1s (or yes) |
| hv232b | items present: ash, mud, sand | 40% | Binary, 6.8% 1s (or yes) |
| sh121i | has cabinet | 30% | Binary, 11.4% 1s (or yes) |
| sh121k | has mattress | 20% | Binary, 32.5% 1s (or yes) |
| hv237 | anything done to water to make safe to drink | 30% | Binary, 5.6% 1s (or yes) |
| hv232y | items present: none | 40% | Binary, 86% 1s (or yes) |
| hv247 | has bank account | 10% | Binary, 15.1% 1s (or yes) |
| roof | roof | 50% | Categorical, 0 (finished) = 83.8%,   1 (natural) = 15.7%,   2 (other) = 0.1%,   3 (rudimentary) = 0.3% |
| sh121l | has hoe | 10% | Binary, 87.1% 1s (or yes) |

# Cluster #1 Configuration

| Cluster Group | has electricity | has television | has mattress | roof | Proportion (%) |
| --- | --- | --- | --- | --- | --- |
| 1 | 1 | 0 | 1 | finished | 5.10 |
| 2 | 0 | 1 | 1 | finished | 6.62 |
| 2 | 1 | 1 | 1 | finished |  |
| 2 | 1 | 1 | 1 | rudimentary |  |
| 3 | 0 | 0 | 0 | natural | 15.01 |
| 3 | 0 | 1 | 0 | natural |  |
| 3 | 1 | 0 | 0 | natural |  |
| 3 | 1 | 1 | 0 | natural |  |
| 4 | 0 | 0 | 0 | finished | 52.52 |
| 4 | 0 | 0 | 0 | other |  |
| 4 | 0 | 0 | 0 | rudimentary |  |
| 4 | 0 | 1 | 0 | finished |  |
| 4 | 1 | 0 | 0 | finished |  |
| 4 | 1 | 1 | 0 | finished |  |
| 5 | 0 | 0 | 1 | finished | 20.74 |
| 5 | 0 | 0 | 1 | natural |  |
| 5 | 0 | 0 | 1 | other |  |
| 5 | 0 | 0 | 1 | rudimentary |  |
| 5 | 1 | 0 | 1 | rudimentary |  |

# Validation Tables

## a.1) Using Children Deceased (Sorted by proportion of 0%)

| Cluster ID/Children Deceased | 0% | 1-33% | 34-66% | 67+% |
| --- | --- | --- | --- | --- |
| 2\* | 1,428 (93.3%) | 76 (5.0%) | 20 (1.3%) | 6 (0.4%) |
| 1 | 820 (93.1%) | 44 (5.0%) | 16 (1.8%) | 1 (0.1%) |
| 5 | 3,230 (84.3%) | 482 (12.6%) | 102 (2.7%) | 17 (0.4%) |
| 4 | 6,968 (78.7%) | 1,413 (15.9%) | 417 (4.7%) | 61 (0.7%) |
| 3 | 1,599 (73.8%) | 383 (17.7%) | 167 (7.7%) | 19 (0.9%) |
| Total | 14,045 (81.3%) | 2,398 (13.9%) | 722 (4.2%) | 104 (0.6%) |
| \*The chi-squared p-value is 0 | | | | |

## a.2) Aggregating proportions greater than 0%

| Cluster ID/Children Deceased | 0% | >0% |
| --- | --- | --- |
| 2\* | 1,428 (93.3%) | 102 (6.7%) |
| 1 | 820 (93.1%) | 61 (6.9%) |
| 5 | 3,230 (84.3%) | 601 (15.7%) |
| 4 | 6,968 (78.7%) | 1,891 (21.3%) |
| 3 | 1,599 (73.8%) | 569 (26.2%) |
| Total | 14,045 (81.3%) | 3,224 (18.7%) |
| \*The chi-squared p-value is 0 | | |

## b) Using Individual Education Level Attained (Sorted by weighted average by row)

| Cluster ID/Education | 0 | 1 | 2 | 3 | 4 | 5 | W. Avg. |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 2\*a | 82 (5.4%) | 204 (13.3%) | 169 (11.0%) | 852 (55.7%) | 56 (3.7%) | 167 (10.9%) | 2.72 |
| 1 | 76 (8.6%) | 134 (15.2%) | 107 (12.1%) | 464 (52.7%) | 38 (4.3%) | 62 (7.0%) | 2.50 |
| 5 | 849 (22.2%) | 883 (23.0%) | 653 (17.0%) | 1,366 (35.7%) | 61 (1.6%) | 19 (0.5%) | 1.73 |
| 4 | 3,661 (41.3%) | 2,362 (26.7%) | 1,250 (14.1%) | 1,550 (17.5%) | 32 (0.4%) | 4 (0.0%) | 1.09 |
| 3 | 1,211 (55.9%) | 634 (29.2%) | 159 (7.3%) | 163 (7.5%) | 1 (0.0%) | 0 (0.0%) | 0.67 |
| Total | 5,879 (34.0%) | 4,217 (24.4%) | 2,338 (13.5%) | 4,395 (25.5%) | 188 (1.1%) | 252 (1.5%) | 1.39 |
| \*The chi-squared p-value is 0 | | | | | | | |
| a0=none,1=incomplete primary, 2=primary, 3=incomplete secondary, 4=secondary, 5=higher | | | | | | | |

## c) Using Primary Healthcare Source (Sorted by % enrolled in public healthcare [ascending order])

| Cluster ID/Primary Healthcare Source | 0 | 1 | 2 | 41 | 42 |
| --- | --- | --- | --- | --- | --- |
| 2\*a | 135 (54.9%) | 25 (10.2%) | 69 (28.0%) | 15 (6.1%) | 2 (0.8%) |
| 1 | 99 (62.3%) | 9 (5.7%) | 43 (27.0%) | 6 (3.8%) | 2 (1.3%) |
| 5 | 433 (76.0%) | 36 (6.3%) | 74 (13.0%) | 22 (3.9%) | 5 (0.9%) |
| 4 | 970 (83.7%) | 85 (7.3%) | 70 (6.0%) | 22 (1.9%) | 12 (1.0%) |
| 3 | 269 (91.5%) | 11 (3.7%) | 8 (2.7%) | 4 (1.4%) | 2 (0.7%) |
| Total | 1,906 (78.5%) | 166 (6.8%) | 264 (10.9%) | 69 (2.8%) | 23 (0.9%) |
| \*The chi-squared p-value is 0 | | | | | |
| a0=public/government, 1=private, 2=other | | | | | |