



# Deriving an asset-based Wealth Index from refugee and host community survey data

# Acknowledgements

This methodological paper was produced by Sofia Leopold, Arfa Khan and Patrick Brock from the World Bank-UNHCR Joint Data Center with support and inputs from the UNHCR Global Survey Team.

The document benefited from inputs and comments from colleagues, including Ilgi Bozdog, Imane Chaara, Felix Schmieding, Ibrahima Sarr, Steven Ndung'u, and Alessandro Nava. The authors gratefully acknowledge the contributions of UNHCR country operations, the World Bank-UNHCR Joint Data Center on Forced Displacement, and government partners in South Sudan, Cameroon, and Pakistan where the pilot FDS surveys were realized.

Recommended citation: Leopold, S., Khan, A., and Brock, P. (2026). *Deriving an asset-based wealth index from refugee and host community survey data*. FDS Methodological Paper No. 2. Copenhagen, Denmark: United Nations High Commissioner for Refugees.

Cover photo: Sudanese refugee mother determined to make life better for herself and her family in South Sudan.

(©UNHCR/Samuel Otieno)

# Contents

<b>Introduction</b>	<b>4</b>
<b>Why a wealth index for refugee survey data?</b>	<b>4</b>
DHS wealth index . . . . .	4
Forced Displacement Survey Program . . . . .	5
<b>Applying the wealth index to the FDS South Sudan</b>	<b>5</b>
<b>Methodology</b>	<b>7</b>
Categorical variables . . . . .	8
Asset ownership . . . . .	8
Land ownership . . . . .	8
Treatment of binary and continuous variables . . . . .	9
Joint urban/rural principal component analysis . . . . .	9
<b>Results</b>	<b>10</b>
Application in other contexts . . . . .	12
<b>Recommendations</b>	<b>12</b>
<b>References</b>	<b>15</b>
<b>Appendices</b>	<b>16</b>
Descriptive analysis of wealth index variables in FDS South Sudan . . . . .	16
Calculation of the Wealth Index . . . . .	23

# Introduction

This methodology note presents a standardized and replicable approach for calculating asset-based wealth indices using UNHCR Forced Displacement Survey (FDS) data. Building on the Demographic and Health Survey (DHS) Wealth Index as a foundation, it adapts the methodology to displacement contexts and provides practical guidance for its application. The note supports three key goals: generating wealth index results to complement existing FDS reports, establishing a clear and transparent methodology, and promoting a standardized approach across countries and surveys. It also offers survey data collection implementation recommendations to promote comparability in the future.

Using the 2023 South Sudan FDS as a case study, the note outlines how principal component analysis (PCA) can be applied to largely categorical asset and housing data, including strategies for incorporating continuous variables. An in-depth descriptive analysis of the South Sudan data illustrates how the index reflects patterns in household living standards and supports meaningful comparisons across displaced and host populations. The overall aim is to provide a robust tool for enhancing socioeconomic analysis in humanitarian settings and aligning FDS outputs with development-oriented data needs.

## Why a wealth index for refugee survey data?

Efforts to support refugees have traditionally focused on addressing immediate humanitarian needs, such as food, shelter and health. As a result, available data on refugee populations is often geared towards short-term crisis response. Long-term development indicators and data collected for national development planning rarely include refugees. To design development interventions that account for the forcibly displaced, it is crucial to bridge the gap between humanitarian and development data. This requires comparable data on both refugees and host communities that is aligned with international statistical standards and includes indicators beyond immediate humanitarian needs such as assets ownership and housing characteristics.

In South Sudan, a country with over 450,000 refugees as of April 2024 and millions of South Sudanese living below the poverty line, there are currently no estimates of refugee wealth (UNHCR, 2024). The 2024 World Bank Poverty Report indicates that 75.9 per cent of the population lives below the national poverty line and identifies disparities in welfare and food security across the country. Despite better access to services, displaced populations in camps for internally displaced persons (IDPs) face higher poverty rates and significant human capital deficits compared to urban residents (World Bank, 2024). However, the report does not provide sufficiently disaggregated data to assess the situation of refugees in South Sudan.

The 2023 Forced Displacement Survey begins to close this gap by providing data on the socioeconomic and living conditions of refugees and host communities in South Sudan, including the variables used to derive the DHS wealth index (UNHCR, 2024). This presents a rare opportunity to generate development indicators for refugee populations, such as relative well-being, using data collected in a humanitarian context.

### DHS wealth index

The DHS wealth index is a widely used measure of household economic status in low- and middle-income countries. It helps identify disparities in living standards by ranking households based on their ownership of durable goods, housing quality and access to basic services such as water and sanitation. These rankings are typically divided into five wealth quintiles, grouping households from the poorest to the wealthiest within each survey.

Rather than relying on income or consumption data, which can be costly to collect and may fluctuate in

the short term, the DHS wealth index uses assets as proxies for long-term wealth. This approach simplifies the complex dimensions of wealth into a single, standardized measure of household economic status. While this dimensional reduction can obscure some nuances, it offers a practical and intuitive tool for analyzing inequality and targeting interventions, which is why the DHS Wealth Index has become a standard instrument in public health and development research, especially in settings where monetary poverty data is unavailable (Rutstein, 2008).

However, this approach has clear limitations for comparability across countries or over time. Because each DHS wealth index is calculated using assets that are relevant to a specific country and survey year, wealth scores are relative within each survey and cannot be directly compared between different populations. Urban-rural differences in asset ownership also bias results unless separate indices or context-neutral variables are used (Rutstein, 2008). Several studies have examined these comparability issues (McKenzie, 2005; Rutstein, 2015). In response, Smits and Steendijk Smits and Steendijk (2015) developed the International Wealth Index (IWI), a globally comparable index using a fixed set of 12 harmonized variables and standardized weights derived from pooled international data. While the DHS Wealth Index remains useful for within-country analyses, the IWI is better suited for cross-national or longitudinal comparisons.

In contexts like South Sudan, where national household surveys do not typically include refugees and where comprehensive monetary poverty data is difficult to collect, asset-based wealth indices offer a valuable insight into relative living standards across refugee and host communities. This approach helps fill a critical data gap by translating humanitarian survey data into development-relevant insights and supporting the alignment of humanitarian response with long-term development goals. Moreover, asset-based wealth indices are increasingly being used to train machine learning models to predict welfare in data-scarce environments, further expanding their utility in contexts with limited economic data (Yeh et al., 2020; Zheng et al., 2025).

## **Forced Displacement Survey Program**

The Forced Displacement Survey by the UNHCR is designed to streamline and standardize data collection on forcibly displaced populations. It provides high-quality, nationally representative data on the socioeconomic and living conditions of refugees and asylum seekers hosted in low- and lower-middle-income countries and covers host communities living in proximity to refugees and asylum seekers, aiming to inform humanitarian operations, government policy, and development programming (UNHCR, 2024). The FDS aims to bridge the gap between humanitarian response and development planning, offering data that supports durable solutions for those affected by forced displacement. For the first time in a UNHCR-led household survey, the FDS includes detailed multi-topic information on household living conditions and wellbeing aligned with international statistical standards, enabling the construction of asset-based wealth indices. The information collected aligns with the variables used in the DHS methodology, and includes land and house ownership, building materials of dwellings, access to electricity and bank accounts, access to drinking water, sanitation and clean fuel, and the ownership of a long list of assets.

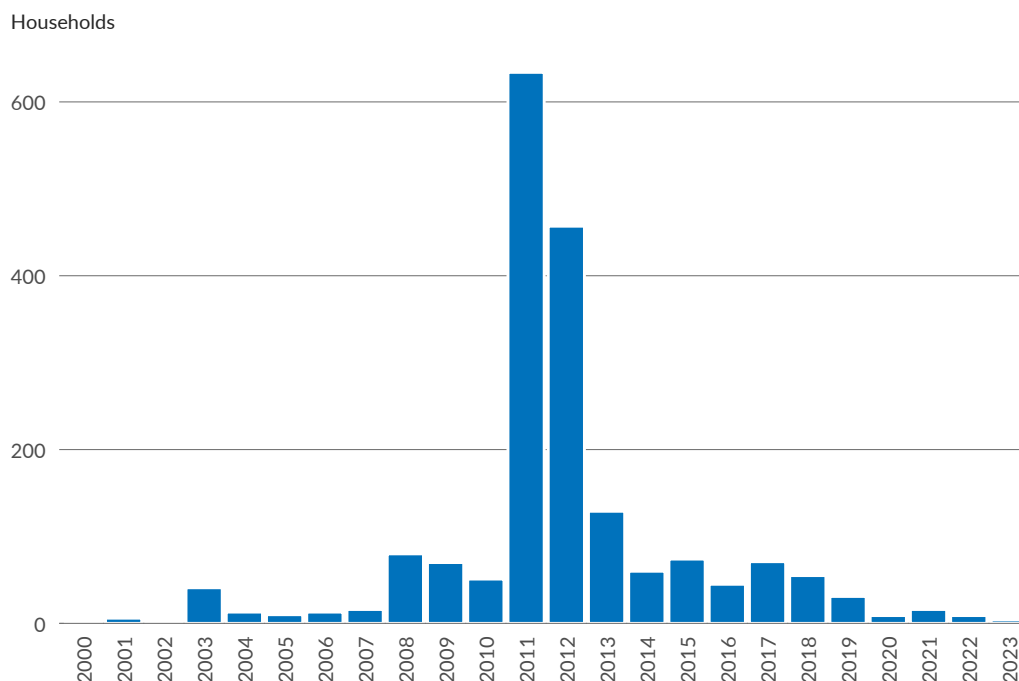
# **Applying the wealth index to the FDS South Sudan**

In displacement contexts, the DHS assumption that assets reflect long-term wealth does not hold in the same way. Displacement often involves the loss of assets, which may lower measured wealth without necessarily reflecting a household's underlying capabilities or potential to rebuild wealth, such as education or professional skills.

Yet, the majority of households in the South Sudan sample report displacement around 2011-2013, corresponding to the period of conflict surrounding South Sudan's independence. Only 3 per cent of the refugees in the sample arrived after 2018 and excluding these observations has minimal impact on the wealth distribution. Among refugees in the south, the proportion of households in the poorest quintile

shifts only slightly from 7.4 per cent to 7.7 per cent, and among refugees in the north, it changes from 15.0 to 14.9 per cent. This suggests that due to the largely protracted displacement situation in South Sudan, the loss of assets during displacement does not substantially skew the relative wealth ranking in the South Sudan FDS.

**Figure 1. Refugee households in South Sudan FDS - year of displacement**



The effects of humanitarian assistance are more difficult to disentangle. Refugees in South Sudan receive varying forms of assistance, both in camps and outside of camps, which can influence indicators used in the index, such as access to services and housing materials. Even host community households may benefit from services provided near refugee settlements. While it is impossible to quantify precisely how humanitarian assistance affects the wealth distribution, households benefiting from these services are in a relatively better position than households without such access, even when access is facilitated through humanitarian support. Therefore, the wealth index developed here is best understood as a measure of current well-being and living conditions, rather than long-term sustainable wealth or self-reliance.

To assess the feasibility of constructing an asset-based wealth index for refugee contexts based on the DHS Wealth Index methodology, we first conducted a descriptive analysis of the variables used in the index calculation for the 2023 Forced Displacement Survey in South Sudan. This analysis examines variation in key asset and housing characteristics across three population groups: refugees in the south, refugees in the north, and members of the host community in the north. It serves as both foundation for producing development-relevant and comparable wealth data for forcibly displaced populations in South Sudan, and as a reference point for interpreting the wealth index results. By examining how variables differ across groups, the descriptive analysis helps to explain the patterns seen in the wealth quintile distributions and sheds light on the main drivers of relative wealth.

The disaggregation into three population groups follows the classification of the FDS report:

- **Refugees North:** Refugees living in camps and settlements in Pariang county, Ruweng Administrative Area and Maban county, Upper Nile. This area hosts the large majority of South Sudan’s refugees. Refugees here mostly live in camp settings where humanitarian agencies provide services.
- **Refugees South:** Refugees living in Central Equatoria, Western Equatoria and Jonglei. They represent around 8 per cent of South Sudan’s refugee population. Unlike in the north, refugees in the south live in dispersed settlements with lower levels of management and often integrate into

neighbouring communities. The sample does not include urban refugees in Juba, who represent less than 2 per cent of the refugee population.

- ◆ **Host Community North:** This group includes non-displaced households in Pariang county, Ruweng Administrative Area and Maban county, Upper Nile. Hosts are defined as the national population living in proximity of the refugee population, interacting, and sharing space, resources, and opportunities with the refugee population. Due to logistical and budget constraints, no comparable host sample was collected in the south.

Overall, refugees in the south appear to be better off than both refugees and host community members in the north across a wide range of indicators. In terms of housing materials, refugees in the south tend to live in dwellings constructed with higher-quality materials for walls, roofs, and floors. While natural materials remain common, their use is significantly lower among southern refugees compared to both northern groups. Asset ownership paints a similar picture: refugees in the south report higher rates of ownership for many key items such as hoes, pangas, chairs, mattresses, mobile phones and smartphones than both refugees and host community households in the north. Refugees in the north report better access to improved drinking water (36 per cent) and improved sanitation (61 per cent) than both southern refugees (25 per cent and 50 per cent, respectively) and northern host communities (23 per cent and 15 per cent). This may be driven by facilities provided in northern refugee camps compared to more dispersed, community-based settings in the south.

Together, these findings suggest that while camp-based services may benefit northern refugees in certain areas, southern refugees enjoy broader advantages across housing quality, asset ownership, and living conditions, positioning them as the relatively better-off group among the three. Despite these observed differences, the fact that households across all groups tend to own similar types of assets and rely on comparable fuels, water sources, sanitation facilities and building materials provides confidence that comparisons across groups are both possible and valid. Assets not owned by host community households typically also show very low ownership rates across refugee groups. This supports the validity of calculating a wealth index across the entire sample, from which group-specific distributions can be derived.

As described in the FDS report, the recent conflict in Sudan may have changed conditions since the time of data collection. It will be important to revisit these findings with new survey rounds that includes asset data in South Sudan (UNHCR, 2024). With the standardized methodology from this paper, such updates should be straightforward and allow for comparability over time.

The detailed results of the descriptive analysis can be found in Annex 1.<sup>1</sup>

## Methodology

To construct the wealth index, we apply Principal Component Analysis (PCA), taking into account the high proportion of binary variables in the final dataset. The first principal component is extracted and used as the wealth index, assigning each household a score that reflects its relative position within the asset distribution. Households are then ranked by their wealth score and divided into five equal groups, from poorest (Q1) to wealthiest (Q5) quintile. These quintiles are calculated using the full analytical sample, enabling comparisons across key subgroups such as refugees, host communities, and urban or rural populations.

Several adjustments were made to the standard DHS methodology to better align with the structure of the Forced Displacement Survey data and to simplify future applications across country contexts. A detailed step-by-step description of the methodology can be found in Annex 2 – Calculation of the Wealth Index.

---

<sup>1</sup> Code for Replication of the descriptive analysis and well as standardized methodology for calculating wealth index in refugee/host settings using Forced Displacement Surveys is included with the full wealth index calculation on GitHub. (<https://github.com/jointdatacenter/wealthindex>).

## Categorical variables

Instead of generating multiple binary (dummy) variables for each category of key household characteristics, as done by the DHS, we use aggregate binary indicators for access to improved drinking water, sanitation and clean cooking fuel. Similarly, we classify building materials for walls, roofs and floors into three standardized DHS-defined categories: natural, rudimentary, and finished to reduce the number of variables while retaining the interpretability of housing materials. Moreover, a binary crowding indicator is used in place of the continuous variable of number of persons sleeping in one room used in DHS wealth index calculations. Households are considered overcrowded if three or more members share a sleeping room (Acioly and Horwood, 2011).

These simplifications streamline the analysis and reflect the increasing availability of some pre-calculated indicators in newer FDS datasets. For the South Sudan version these variables were constructed manually. Working with the South Sudan FDS data revealed two data quality considerations. First, the lack of detail for “Other, specify” responses in categorical variables limits interpretability. The dataset did not include the specified responses for these observations, which represent a substantial number of households in some key variables, such as main source of drinking water, type of toilet facility and main roof material. Access to this information could improve the reliability of the analysis.

Second, for respondents living in certain building types – traditional houses and tents – the questionnaire appears to skip the observation of materials used for walls, roofs and floor. Those being the most common housing types across the South Sudan sample, excluding these cases was not an option. Instead, materials were imputed based on assumptions about what is typically used in these dwellings (see Annex 2.1.3 Dwelling Materials Imputation specifying e.g. tents with natural wall, floor and roof materials). Given the large number of observations, these assumptions can have a substantial impact on the results and introduce bias if applied incorrectly. Therefore, including the observations on building materials for all housing types or providing clear guidelines on how to code materials for these two very common housing types in South Sudan could improve accuracy and consistency.

## Asset ownership

Asset and livestock variables were included as provided in each country's FDS dataset, under the assumption that the survey design teams applied informed, context-specific judgment when selecting asset items. All asset variables present and owned by at least one household were retained in the analysis. Assets that were entirely unused (i.e., not owned by any household) were excluded automatically.

While it may be tempting to drop assets with low ownership rates, doing so risks removing highly informative wealth indicators, such as a car, which may only be owned by a few households but still signals clear wealth differentiation. Consistency in our approach means retaining all available and non-empty assets, regardless of prevalence.

## Land ownership

Land ownership is included in the PCA input dataset only in countries where refugees have either a de jure or de facto right to own land. In countries where refugees are legally restricted from land ownership, or where no substantial number of refugee households report owning land, the variable is excluded to avoid introducing bias and ensure comparability between population groups.

Excluding or including this variable was tested for its impact on the wealth index in the South Sudan example. Excluding land ownership had no significant effect on the resulting quintile distribution, supporting the robustness of the index in either specification.

## Treatment of binary and continuous variables

A key consideration in constructing an asset-based wealth index using PCA is the treatment of different variable types, particularly binary indicators, which dominate many forced displacement datasets, alongside occasional continuous variables. It is essential to ensure that no single variable type disproportionately influences the principal components, particularly in datasets where binary variables are the norm.

To address this, the PCA in this analysis is implemented using a correlation matrix that appropriately reflects the mix of variable types. Using R allows the specification of “mixed” correlation, which combines tetrachoric correlations for binary variables and Pearson correlations for continuous ones. This approach ensures that binary variables are treated correctly and prevents continuous variables from disproportionately influencing the wealth index simply because of their scale or distribution. Where the dataset includes only binary variables, it is also possible to specify the use of tetrachoric correlation exclusively.

For example, in the South Sudan dataset, with agricultural land area being the only continuous variable, using a mixed correlation matrix, allowed to preserve the balance across all included indicators. Examination of the variable’s loading on the first principal component confirmed that it had a minimal impact on the index. Further sensitivity checks, such as excluding the variable or recoding it into binary categories, showed that these changes did not meaningfully affect the resulting wealth quintile distribution.

This diagnostic approach, especially reviewing variable loadings, is recommended when considering the inclusion of variables with low prevalence, inconsistent measurement, or potentially disproportionate influence. It supports transparent, data-driven decisions that strengthen the reliability and comparability of the resulting index.

## Joint urban/rural principal component analysis

The DHS wealth index is typically calculated by running separate PCAs for an urban and a rural group, then constructing national scores using regression modelling. This recognizes that asset ownership can differ significantly by settlement type. A similar concern arises when working with data that includes both refugees and host communities, (as well as refugees in and outside of camps) where one might expect systematic differences in access to assets, housing types, and services due to differing legal status, mobility constraints, or humanitarian support.

To assess whether a joint calculation was appropriate, we aimed to compare asset ownership rates between urban and rural and refugee and host households in the South Sudan dataset. However, since the South Sudan FDS data does not include an urban sample, tests included refugees and hosts as well as households living in camps and out of camps (including out-of-camp refugees and hosts). The observed differences in asset ownership across both settlement and population groups were relatively minor. While some variation is expected, the overall similarity in ownership patterns suggested that separate PCA calculations might not be necessary. To test this hypothesis, we implemented the PCA both jointly and separately for refugee and host community households as well as for in-camp and out-of-camp households. The results were highly comparable, reinforcing the conclusion that when asset distributions are similar across groups, a unified PCA approach is methodologically sound.

For additional validation of this logic, we turned to two additional FDS datasets – Cameroon and Pakistan – where both urban and rural households were surveyed. Observed differences in asset ownership between urban and rural households and refugee and host community households were relatively minor and the results of joint and separate PCA calculation for all variations aligned closely.

This reinforces the conclusion that when asset distributions are broadly similar across groups, a unified PCA approach is methodologically sound. Importantly, displaced and host households did not show systematic differences that would warrant separate indices.

This has important practical implications. Despite initial concerns that displaced households might systematically lack access to key assets, our findings indicate that a unified index can be constructed

without compromising methodological robustness, provided that asset distributions are examined in advance. This simplifies the process and facilitates comparability across population groups.

Nevertheless, the applicability of this approach should always be evaluated on a case-by-case basis. While moderate variation across groups is acceptable, if one group consistently lacks access to assets that are widely owned by others, separate indices may still be warranted to avoid distortion. The methodology presented here, including the disaggregation and parallel PCA testing, offers a replicable, evidence-based approach for determining when a unified wealth index is appropriate in forced displacement contexts.

## Results

In South Sudan, notable differences exist between population groups, with host community households appearing worse off than both refugee groups. Geographic variation also emerges as an important factor: refugees in the south appear better off than their northern counterparts. This may be driven by structural and environmental factors. The south offers more arable land and favourable climatic conditions, allowing for greater reliance on subsistence agriculture, which can enhance household self-sufficiency. Additionally, many refugees in the south reside in settlements integrated with host communities, potentially granting better access to livelihoods, markets, and public services. In contrast, refugees in the north are more likely to live in camps, where mobility and economic opportunities are often more restricted.

An interesting observation is that hosts seem to be, on average, worse off than both refugee groups. This reflects in part, the way the wealth index captures current living standards instead of providing a measure of sustainable long-term wealth. Humanitarian assistance contributes to better living conditions for refugees, particularly through access to services such as water, sanitation, and shelter materials, which directly improves their measured well-being. Hosts without comparable support may therefore rank lower on the wealth index, even if refugees remain more aid-dependent and less self-reliant in the longer term. This underscores that humanitarian assistance appears to have a positive impact on household well-being but should not be interpreted as evidence that refugees' economic situation is secure or that they are not in need of support. Importantly, these results reflect only the sample composition of the South Sudan FDS, as described earlier. The host group consists of households from two counties in proximity to refugee camps in the north and is not representative of the entire non-displaced population in South Sudan.

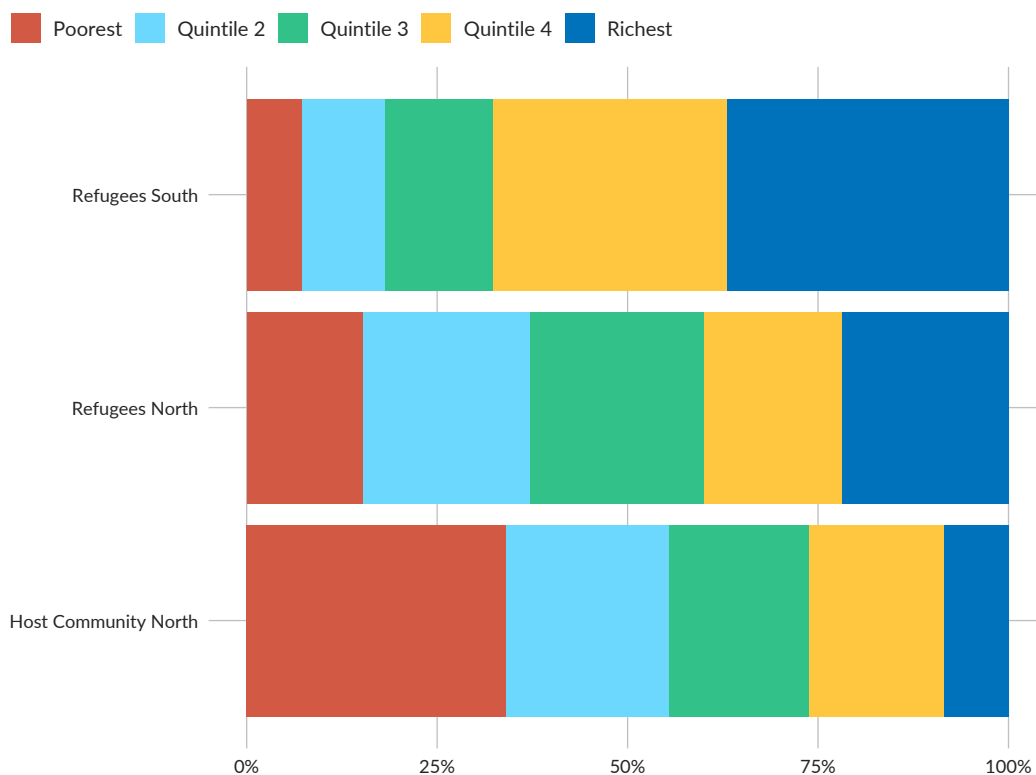
These findings are consistent with the descriptive analysis of asset ownership in Appendix, which shows refugees in the south faring better across multiple indicators (such as asset and livestock ownership rates and quality of building materials). Together, the results suggest that both geographic location and displacement status play central roles in shaping household well-being outcomes.

The first principal component, which is used to construct the index, explains approximately 20 per cent of the variation in the underlying dataset. While this may appear low, it falls within the expected range for PCA-based wealth indices using binary asset data. It reflects the multidimensional nature of household well-being, which a single component cannot fully capture, but still provides a meaningful and interpretable ranking of households along a relative wealth spectrum.

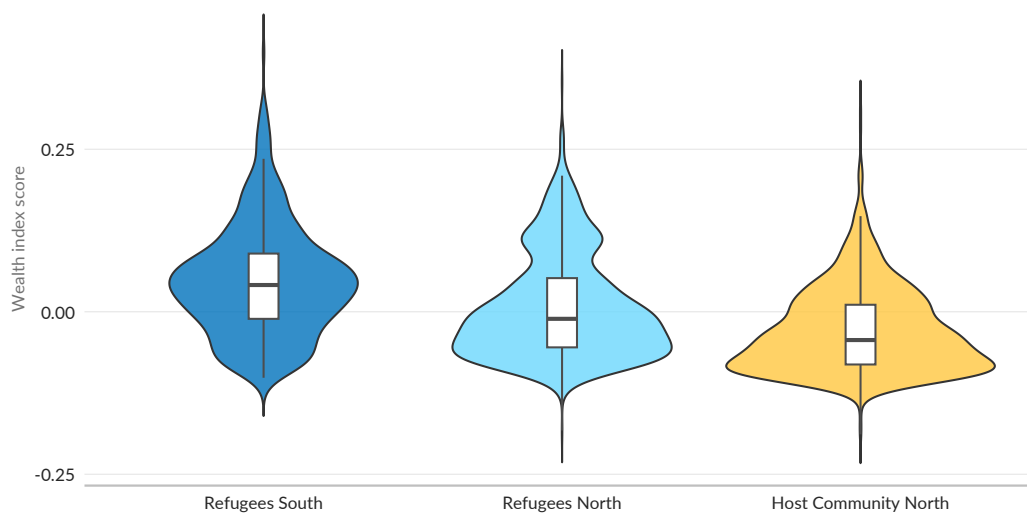
The component loadings of the first principal component, are coherent and interpretable. Assets and housing characteristics associated with better living standards, such as improved materials, access to services, or ownership of consumer goods, load positively on the index, while indicators of deprivation, such as natural building materials or overcrowding, load negatively. Similar types of variables behave consistently, and no single variable dominates the index. These patterns provide reassurance that the index meaningfully captures a broad spectrum of household well-being.

At the same time, some of the variables with the highest loadings, such as ownership of a car, tractor, or Wi-Fi router, are those with the lowest incidence in the population. These rare but high-value assets are highly effective in distinguishing the wealthiest households from the rest of the population. While this improves the index's ability to capture variation at the top end of the wealth distribution, it also means that some of the most influential variables are not broadly informative across the full population. For

**Figure 2. Wealth distribution in South Sudan**



**Figure 3. Wealth Index scores distribution**



example, any household owning a tractor or router is almost certainly found in the top quintile, but this does not help differentiate among lower-wealth households where such assets are entirely absent.

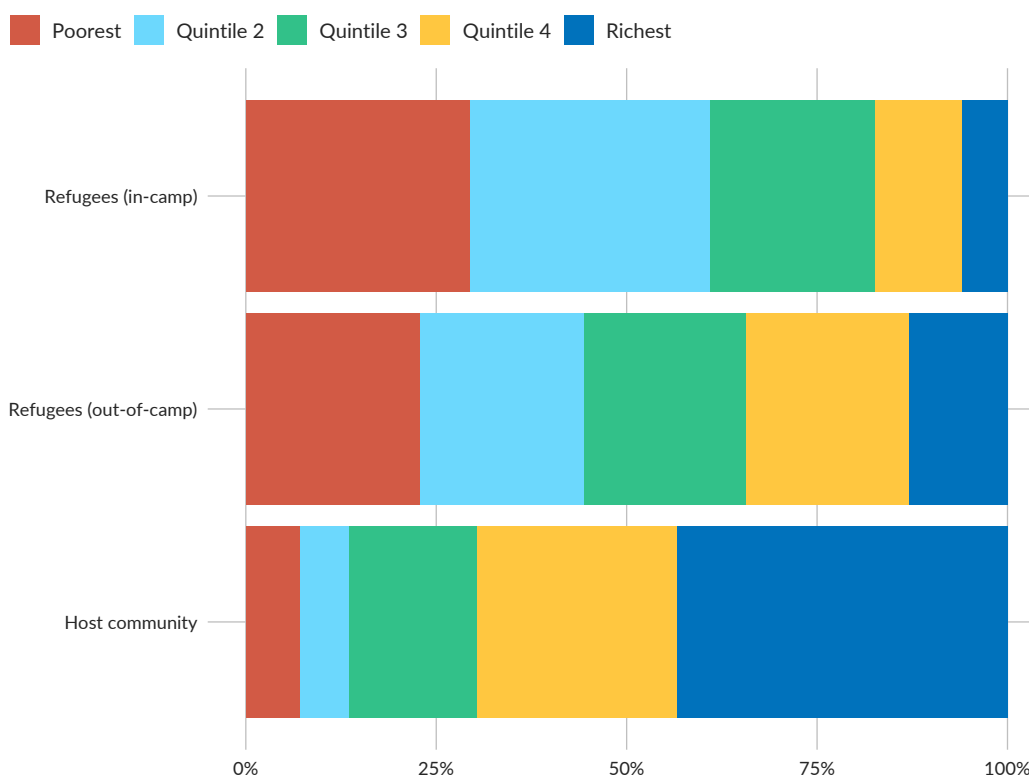
To analyse the magnitude of loadings further, the code provided for the South Sudan wealth index calculation creates a table with loadings for all variables used in the principal component analysis.

## Application in other contexts

The methodology developed using the South Sudan FDS data, was also applied to Pakistan and Cameroon. These applications are not intended as full analyses, but rather as tests of the joint versus separate PCA calculation and to demonstrate that the approach can be transferred across displacement contexts. More comprehensive wealth index analyses for both countries are expected in forthcoming FDS reports.

In Pakistan, the wealth index clearly separates population groups by displacement status. The host community emerges as the wealthiest, followed by out-of-camp refugees, with in-camp refugees being the poorest.

**Figure 4. Wealth distribution in Pakistan**



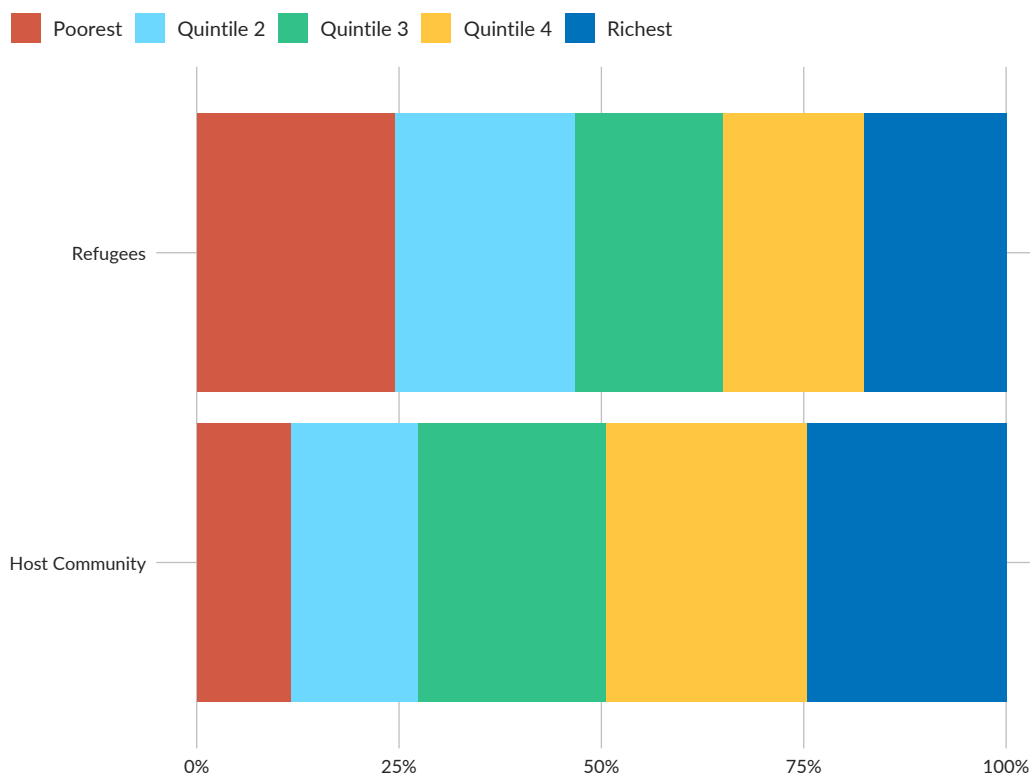
In Cameroon, direct comparison is limited in some regions due to a much larger sample of refugee households. Nonetheless, the available data suggests that host households are generally better off if compared across the full sample. Further analysis within the refugee population may reveal interesting regional variation.

As described above, the PCA was also run separately for urban and rural areas and then combined into a national index from there (following the standard DHS approach) to test the robustness of the approach. Across all three datasets, results remained highly consistent. The overall wealth rankings of groups did not change regardless of whether the index was constructed jointly or separately. This consistency builds confidence in the method and supports the use of a unified wealth index when initial descriptive tabulations show broadly similar patterns of asset ownership across subgroups.

## Recommendations

This methodology offers a practical, standardized tool for measuring relative socioeconomic status within countries using Forced Displacement Survey data. Building on the DHS Wealth Index framework,

**Figure 5. Wealth distribution in Cameroon**



the approach has been adapted to reflect the realities of displacement contexts, capturing variation across refugee and host populations while remaining replicable across countries and surveys.

The wealth index is not designed to produce absolute poverty measures or enable direct comparisons across countries. Asset ownership patterns are inherently shaped by local contexts and removing context-specific variables to enforce comparability would compromise the index’s ability to reflect actual living conditions. Instead, this methodology enables meaningful within-country comparisons, allowing for a clear ranking of households or groups along a relative wealth spectrum. When applied consistently, it can also contribute to broader regional insights into how factors like displacement settings and settlement types relate to household well-being, without implying direct comparisons of absolute wealth levels between countries.

The analysis shows that standardized wealth measurement tools can be meaningfully applied in humanitarian contexts. By translating detailed, often dichotomous asset data into development-relevant indicators, the wealth index enables more accurate assessments of living conditions among displaced and host communities. This not only fills a critical data gap in FDS reporting but also supports the integration of humanitarian data into long-term development efforts, enhancing the value and impact of socioeconomic analysis in displacement settings.

Several lessons emerge for future rounds of FDS and for teams applying this methodology elsewhere.

- ◆ Survey design recommendations for FDS teams:
  - Some response categories currently coded as “Other, specify” contain a substantial share of households, but do not include the specified responses. Making these specified responses available in the dataset would increase interpretability.
  - For common dwelling types such as traditional houses or tents, direct observation of wall, roof, and floor materials would avoid the need for imputation and improve consistency across contexts.
- ◆ Methodological recommendations for constructing a wealth index in displacement settings:

- Examine asset ownership patterns across subgroups (urban/rural, refugee/host, in-camp/out-of-camp). If distributions are broadly similar, a joint PCA is appropriate; where systematic differences exist, separate PCAs are more reliable.
- Pay attention to missing values on key housing variables. Imputation should be guided by contextual knowledge of typical materials and practices in each country.
- Land ownership requires particular attention in displacement contexts. Reported ownership may not reflect actual access and sometimes refer to land outside the current country of residence. For example, in South Sudan some refugee households own land across the border in Sudan, the accessibility of which depends heavily on the political situation. Sensitivity checks in this study confirmed that including or excluding land ownership did not change results, but such analyses should always be conducted, as the effect may vary by context. If land ownership meaningfully alters results, researchers may need to consider alternatives such as access-to-land indicators, even if this departs from the DHS framework.
- To ensure that binary variables are handled appropriately in the calculation, the specification of tetrachoric correlation in the PCA command in R is vital.

Together, these recommendations reinforce the strength and adaptability of the approach. With small adjustments to survey design and careful application of PCA, the wealth index can serve as a reliable, replicable tool for measuring relative socioeconomic status in displacement contexts, helping to align humanitarian and development analysis. The standardized methodology facilitates straightforward analysis and should be used to revisit these findings with future survey data taking into account evolving conditions in South Sudan.

# References

- Acioly, C. and Horwood, C. (2011). *A practical guide for conducting: Housing profiles. Supporting evidence-based housing policy and reform*. UN-Habitat. UNON.
- McKenzie, D. J. (2005). Measuring inequality with asset indicators. *Journal of Population Economics*, 18:229–260.
- Rutstein, S. O. (2008). The DHS wealth index: Approaches for rural and urban areas. *DHS Working Paper*.
- Rutstein, S. O. (2015). Steps to constructing the new DHS wealth index. Calverton, Maryland: DHS Working Paper.
- Smits, J. and Steendijk, R. (2015). The international wealth index (IWI). *Social indicators research*, (122).
- UNHCR (2024). *Forced Displacement Survey – South Sudan 2023*. Copenhagen, Denmark: United Nations High Commissioner for Refugees.
- World Bank (2024). *Republic of South Sudan - Poverty and equity assessment (English)*. Washington, D.C. : World Bank Group.
- Yeh, C., Perez, A., Driscoll, A., Azzari, G., Tang, Z., Lobell, D., Ermon, S., and Burke, M. (2020). Using publicly available satellite imagery and deep learning to understand economic well-being in Africa. *Nature Communications*, 11(2583).
- Zheng, Z., Wu, T., Lee, R., Newhouse, D., Kilic, T., Burke, M., Ermon, S., and Lobell, D. (2025). Dynamic, high-resolution wealth measurement in data-scarce environments. *World Bank Policy research working paper*, (11058).

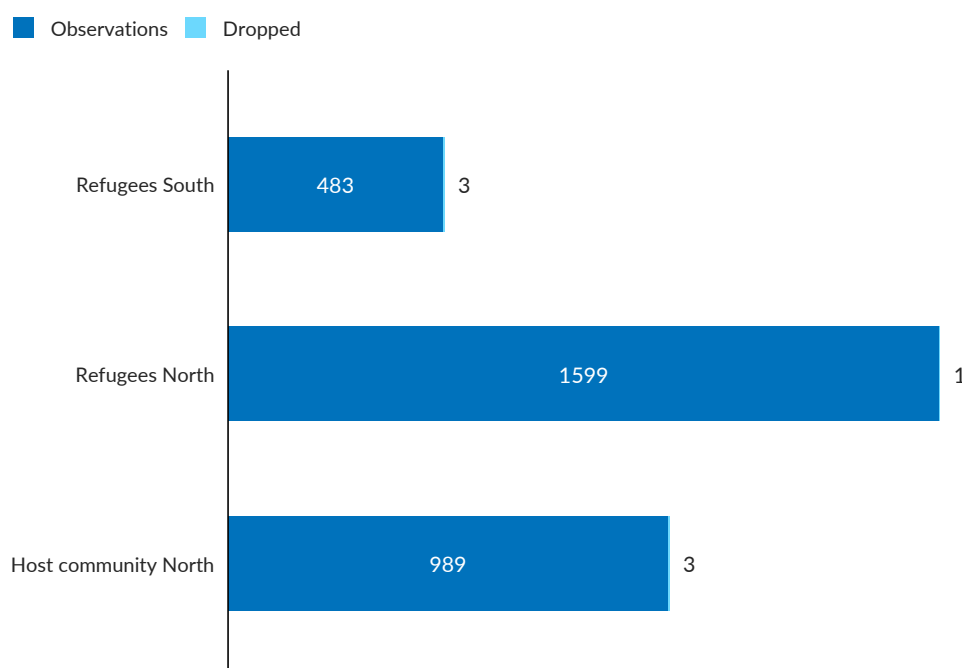
# Appendices

## Descriptive analysis of wealth index variables in FDS South Sudan

After cleaning the data, the total number of observations is 3,071. This includes 483 observations for refugees in the south, 1,599 observations for refugees in the north, and 989 observations for the host community in the north. The values shown in this descriptive analysis incorporate household sampling weights (*wgh\_samp\_resc\_str*). These weights account for unequal probabilities of selection at various stages of sampling, including adjustments for proximity to refugee camps and for households owning multiple buildings. They are rescaled to preserve the sample size within each stratum and are intended for comparative analysis among and within strata<sup>2</sup>.

Importantly, these weights are not applied before the principal component analysis (PCA) but are only used after the PCA, during the calculation of wealth indices for each population group (Rutstein, 2015). The application of these weights in the descriptive analysis does not substantially change the observed differences between groups in the comparisons presented.

**Figure 6. Remaining observations by population group**



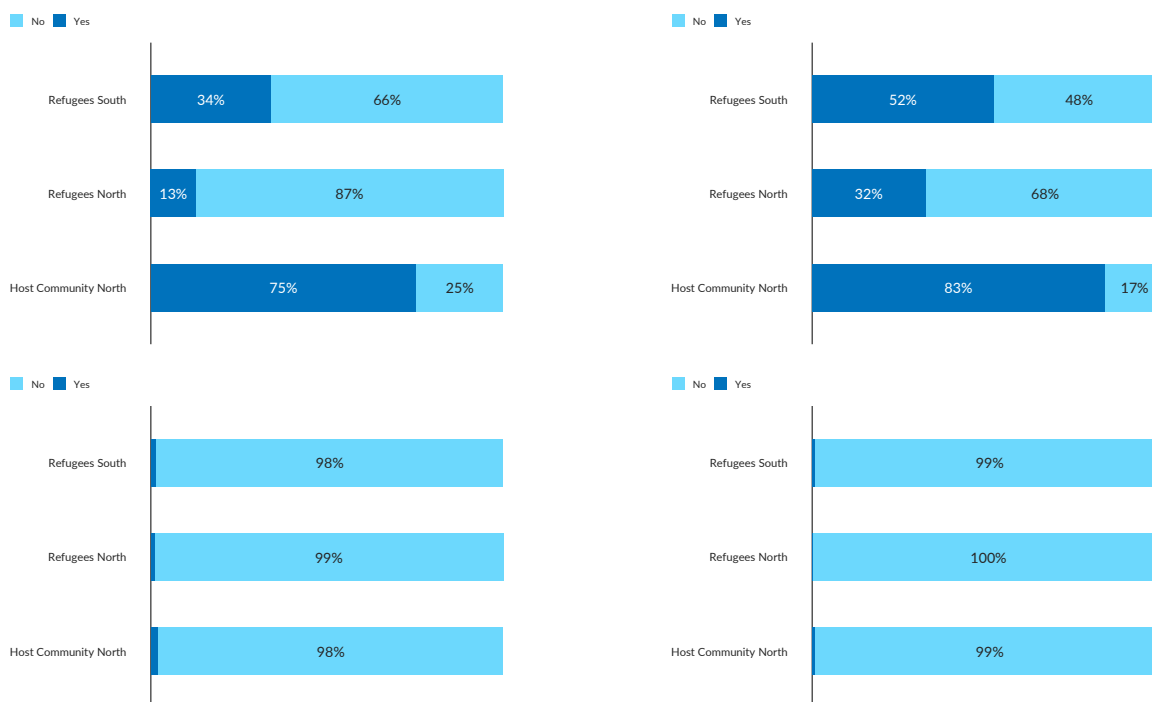
### Land & house ownership, access to bank accounts & electricity

Land and house ownership are high among host community households but lower for the other groups. Refugees in the north have the lowest land ownership, with 13 per cent of households owning land and 32 per cent owing their house. Refugee households in the south are more likely to own land and houses, with more than half owning their own house, but their ownership rates are still below those of the host community in the north.

The number of households with access to a bank account is very low, with less than 2 per cent across all groups and as low as 1.1 per cent for refugees in the north. Access to electricity is also severely limited,

<sup>2</sup> Forced Displacement Survey microdata, South Sudan 2023

**Figure 7. Land & house ownership, access to bank accounts & electricity**



with less than 1 per cent of households in all groups having access to consistent sources of electricity, as also reported in the FDS report.

### Main source of drinking water

Refugees in the north predominantly use public taps or standpipes, with over 80 per cent relying on this water source. In contrast, the main water sources for the other groups are more diverse. Refugees in the south primarily use tube wells, boreholes, or hand pumps, and to a lesser extent, unprotected wells, public taps, standpipes and surface water. Host community households in the north are somewhat evenly split between using public taps or standpipes and wells, boreholes or hand-pumps, with a small portion using surface water.

### Type of toilet facility

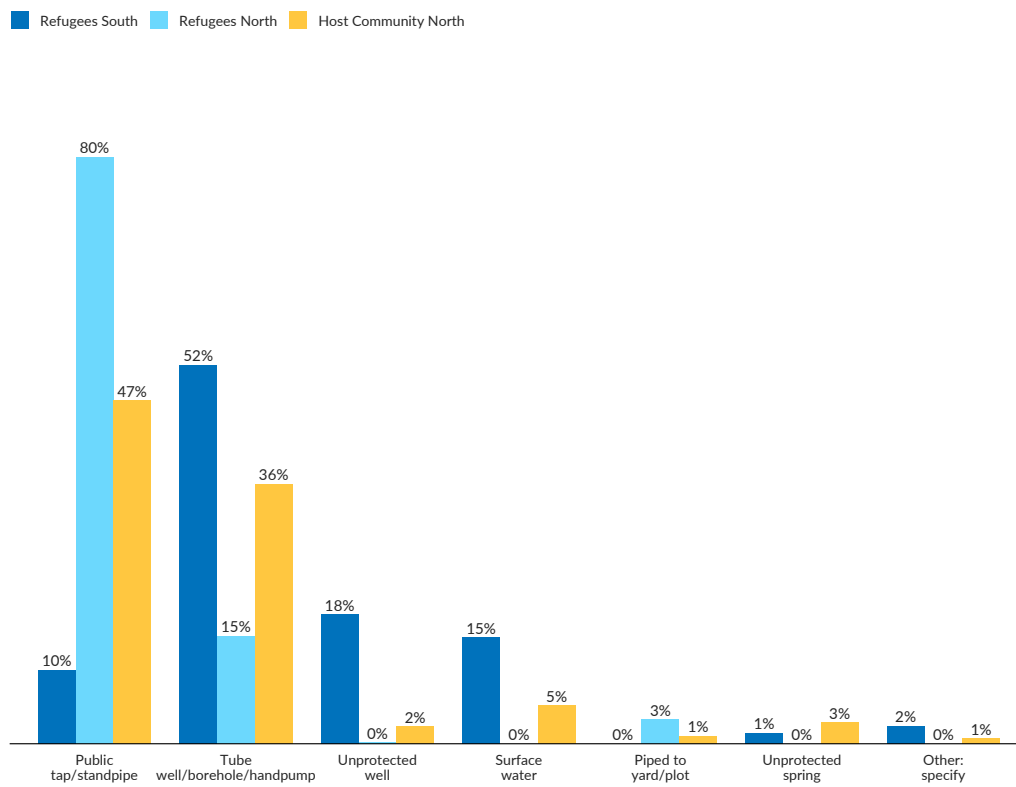
Host community members are the least likely to have access to a toilet facility, with over 70 per cent reporting that they do not typically use any toilet facility. This number is considerably lower, but still above 20 per cent, for both refugee groups. Refugees in the south most commonly use pit latrines with slabs, followed by pit latrines without slabs. In contrast, refugees in the north most commonly use flush-to-pit latrines, followed by pit latrines with slabs. About 5 per cent of refugees in the south report using other types of toilet facilities, though no further specification is provided in the dataset.

While the first figure includes both shared and non-shared facilities, the dataset also provides information on whether households share their toilet facility with others. As shown in the second figure, refugee households in the south are more likely to share toilet facilities than the other groups. For example, 20.5 per cent of refugee households in the south use a shared pit latrine with slab, compared to only 3.6 per cent of refugees and 3.5 per cent of host community households in the north. Shared use is also higher among refugees in the south for pit latrines without slabs (11.3 per cent), while shared flush-to-pit latrines are more common among refugees in the north (7.8 per cent).

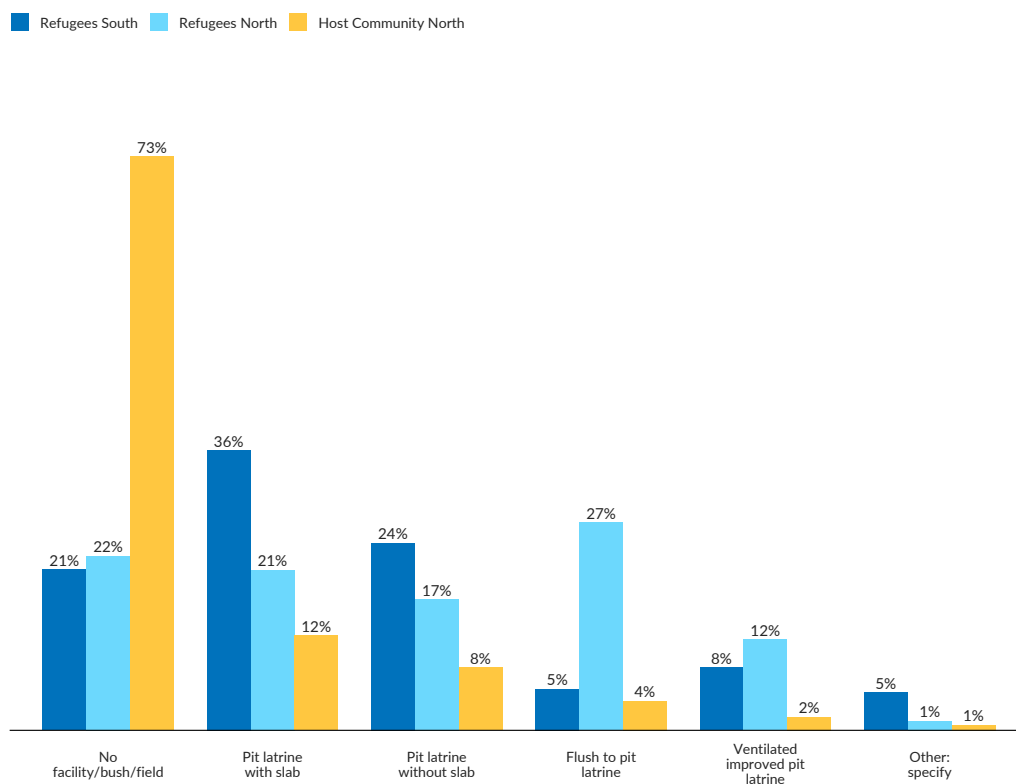
### Type of cooking fuel

The main type of cooking fuel for all three groups is wood, used by more than 70 per cent of all households. Smaller portions of all households in all groups use unprocessed charcoal and unprocessed

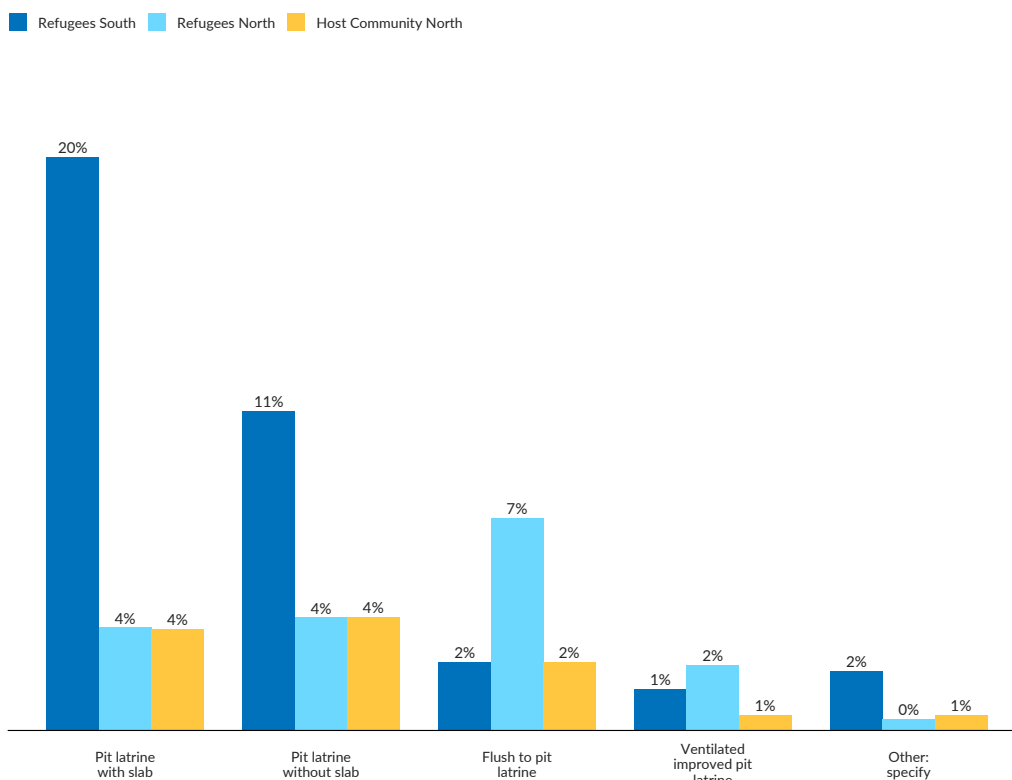
**Figure 8. Main source of drinking water**



**Figure 9. Type of toilet facility**



**Figure 10. Sharing of toilet facility**



coal or lignite. Unlike the other two groups, some refugee households in the north are also using agricultural or crop residue, grass, straw, shrubs or corn cobs, and woodchips as cooking fuel.

For the calculation of the wealth index, some of the detailed dichotomous variables were combined into standardized definitions for improved drinking water, improved sanitation, clean cooking fuel, and overcrowding, following SDG and DHS guidelines. The exact definitions applied can be found in Appendix on Access to Services.

Refugees in the north appear to have better access to both improved drinking water (37 per cent) and improved sanitation (61 per cent) compared to lower numbers for the host community in the same region and refugees in the south.

Access to clean cooking fuel is minimal across all groups, with less than 0.5 per cent of households using clean fuels, reflecting the widespread reliance on wood as the primary cooking source.

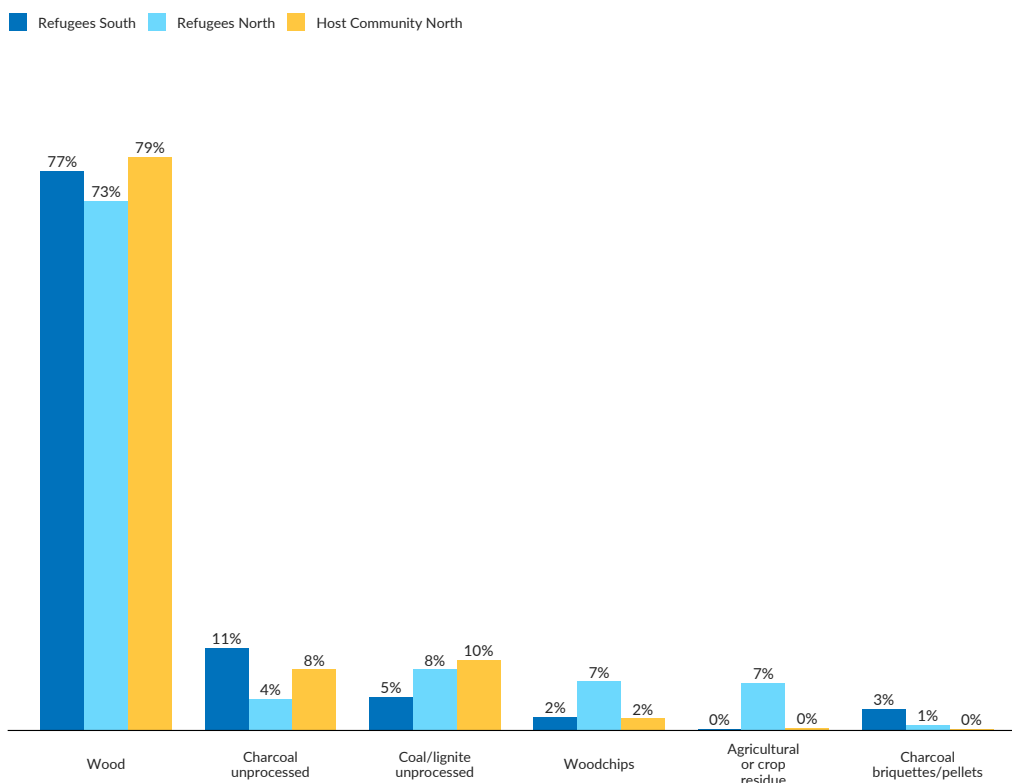
Overcrowding is more severe in the north, where approximately half of all households report more than three people per sleeping room (49 per cent among northern refugees, 51 per cent among northern hosts). In comparison, only 33 per cent of southern refugee households report living in overcrowded conditions.

### Dwelling Characteristics

The most commonly used material for walls across all groups is mud bricks, with 73 per cent of refugees in the south, 57 per cent of refugees in the north, and 41 per cent of the host community households in the north using this material for their dwelling. Interestingly, fewer than 3 per cent of refugees in the south live in dwellings without walls, compared to larger proportions among refugees and the host community in the north (30 per cent and 34 per cent). The share of host community households in the north using mud bricks is slightly lower than for the refugee groups, relying more than other groups on walls made from bamboo with mud with 18 per cent.

For roofs, refugees in the north use significantly less thatch or palm leaf compared to other groups, with

**Figure 11. Type of cooking fuel**



only 48 per cent, compared to 74 per cent among refugees in the south and 67 per cent among the host community in the north. Refugees and host community households in the north are also more likely to live in dwellings without roof (28 per cent and 25 per cent) compared to fewer than 3 per cent of refugee households in the south. Metal roofs are used by 10 per cent of refugee households in the north, compared to only 7 per cent of refugees in the south and 4 per cent of host community households and 8 per cent of refugees in the south and 4 per cent of refugees in the north use other roof materials, which are not further specified in the data.

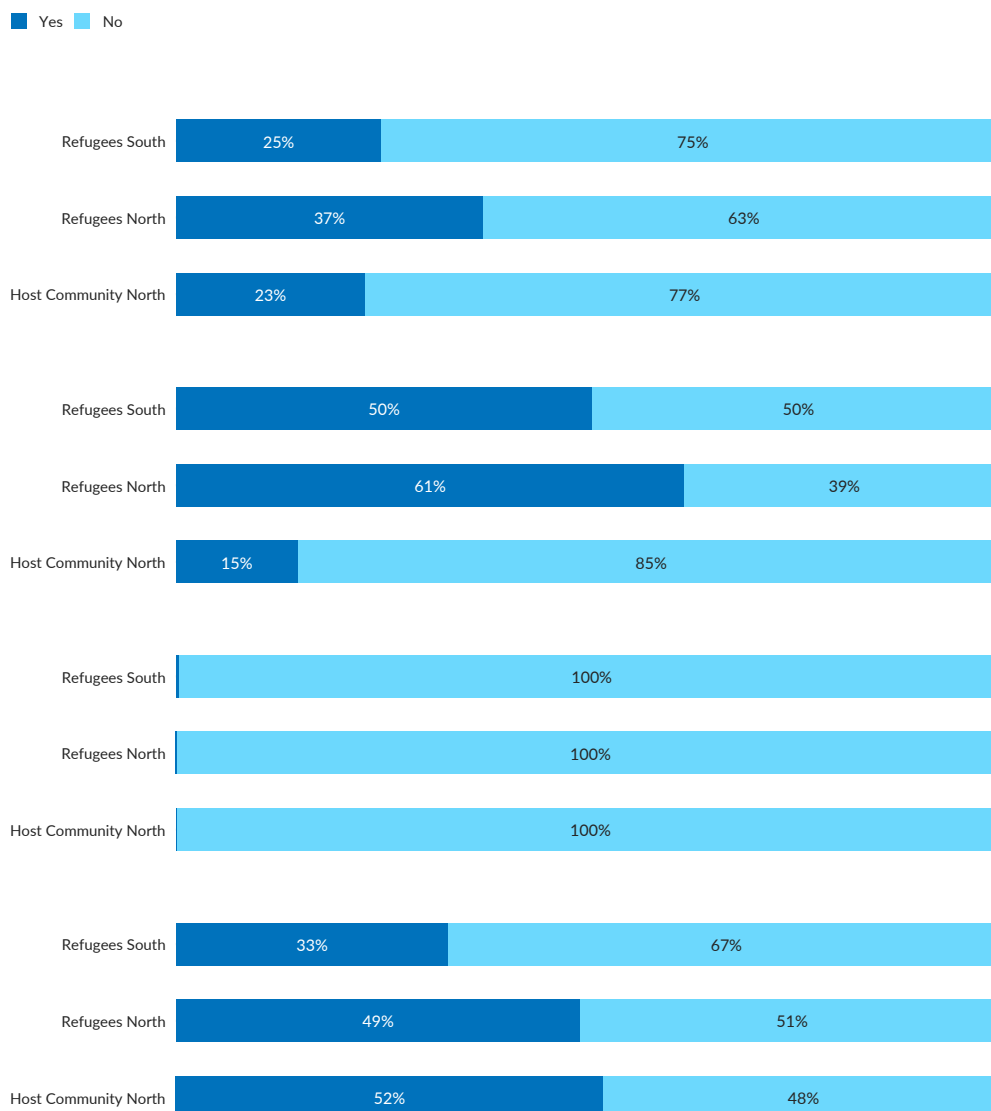
The main floor material for almost all households surveyed is earth or sand. Refugees in the south are the only group in which more than 1 per cent of households use another material such as cement or dung.

For the construction of the wealth index, we use these detailed housing material variables to create three categories for each population group in South Sudan: natural, rudimentary and finished. For each housing feature (walls, roof and floor), households are assigned to one of these categories depending on the material used in their dwelling.

We observe that refugees in the south are more likely to have rudimentary or finished walls compared to both groups in the north, with only 11 per cent of southern refugees using natural floor materials, whereas over 30 per cent of households in both northern groups fall into this category. On the other hand, refugees in the north are more likely to have finished roofs, with 20 per cent compared to 13 per cent among southern refugees and only 5 per cent among northern host households. All three groups rely predominantly on natural floor materials, with over 90 per cent of households falling into this category, reflecting the widespread use of earth and sand floors, as previously discussed.

The assumptions we make to replace missing materials for traditional houses and tents significantly influence these results (see Appendix on Dwelling Materials Imputation). The dominant materials for all three dwelling characteristics categories are those we assign to traditional houses, which are the most common housing type by far. The second most common materials are those we assign to tents. More than 60 per cent of host community and refugee households in the north live in one of these two housing types (approximately 40 per cent in traditional houses and 25 per cent in tents). In the south,

**Figure 12. Access and living conditions by population group**



over 70 per cent of refugees live in traditional houses. This highlights the sensitivity of the variables to our assumptions. It may be important to observe materials in the Forced Displacement Survey even when the housing type is a traditional house or tent.

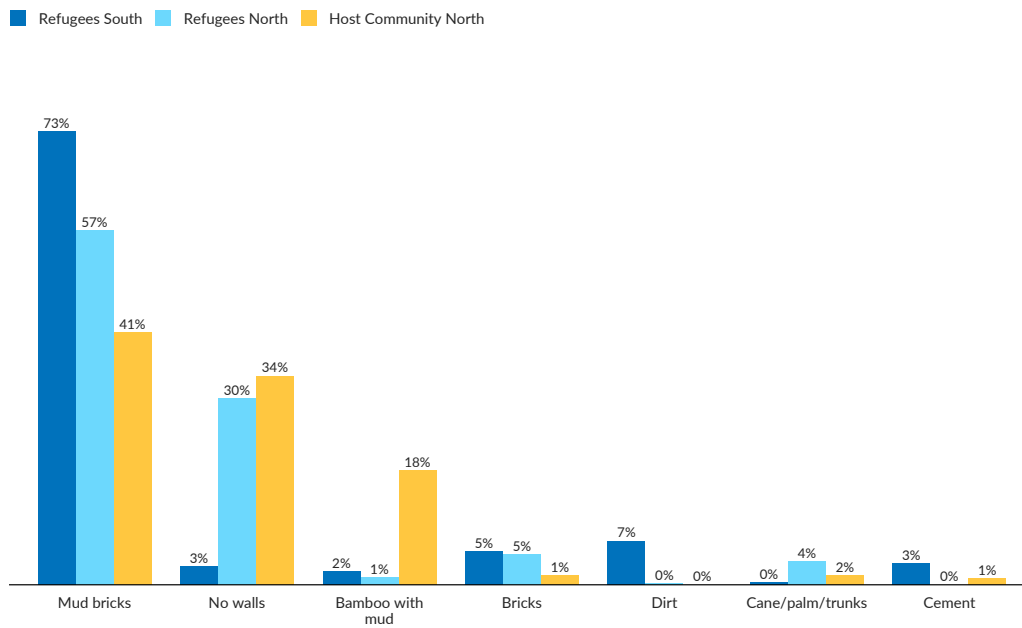
### Asset ownership

The table below shows the percentage of households that own various assets, sorted by the degree of variability between the population groups (measured by the standard deviation of the three percentages). For many asset types, refugees in the south report higher ownership rates than both host community members and refugees in the north. This is the case, for example, for hoes, pangas<sup>3</sup>, chairs, mattresses as well as mobile phones and smartphones, which are more commonly owned by refugees in the south than by the other two groups.

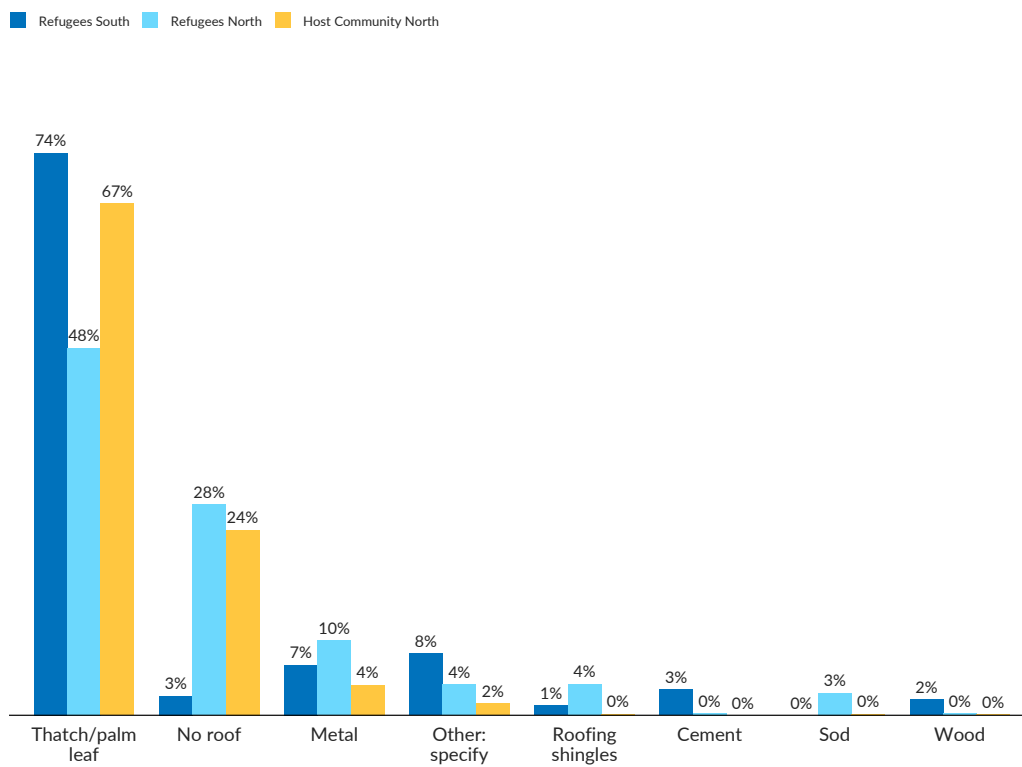
While refugee households in the north appear to have the lowest ownership numbers overall, there are some assets which they are more likely to own than other groups. This includes beds, malodas,

<sup>3</sup> tool used by farmers, like a shovel

**Figure 13. Main wall material**

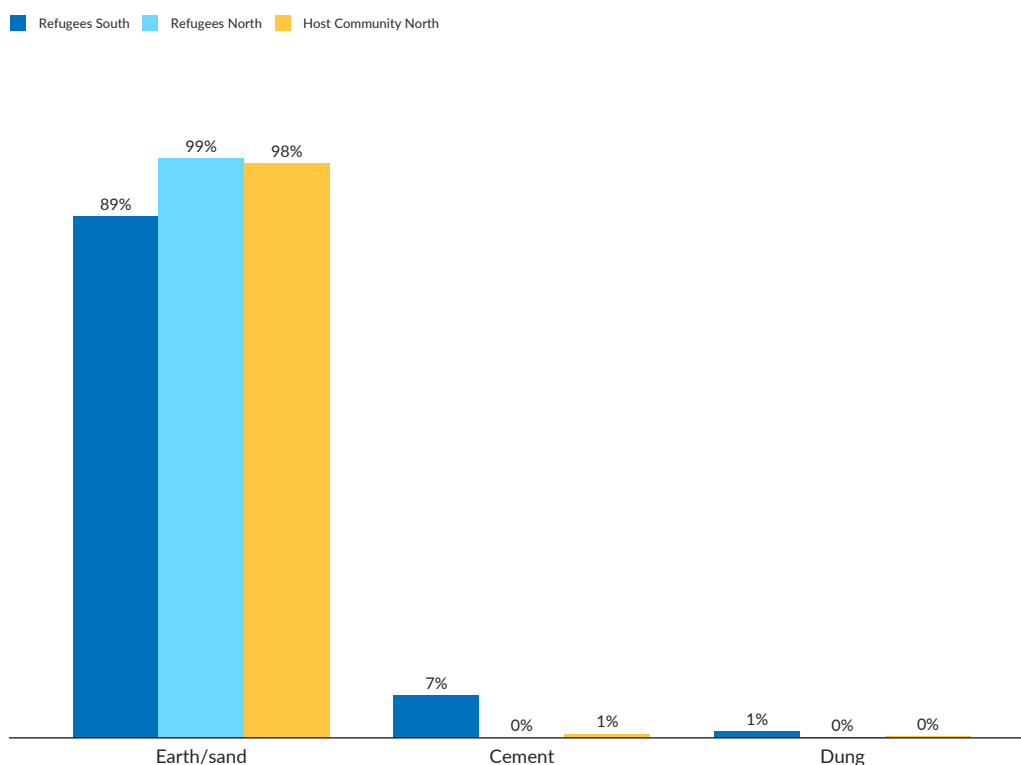


**Figure 14. Main roof material**



traditional stoves, axes, spades and harrows. They also own more hoes, pangas and seeds for planting than host community households in the north, suggesting that refugee households are more likely to own tools for agriculture than the host community.

**Figure 15. Main floor material**



### Livestock ownership

Refugees in the south seem to own the least livestock, particularly lacking cows and bulls, horses, donkeys and mules, and sheep. Both refugees and host communities in the north own more animals overall. For some categories like goats and barnyard animals, refugee households in the north are more likely to own 1-4 animals, while host community households tend to own larger number of animals. Refugees in the north own more sheep overall, whereas host community households own more cows and bulls.

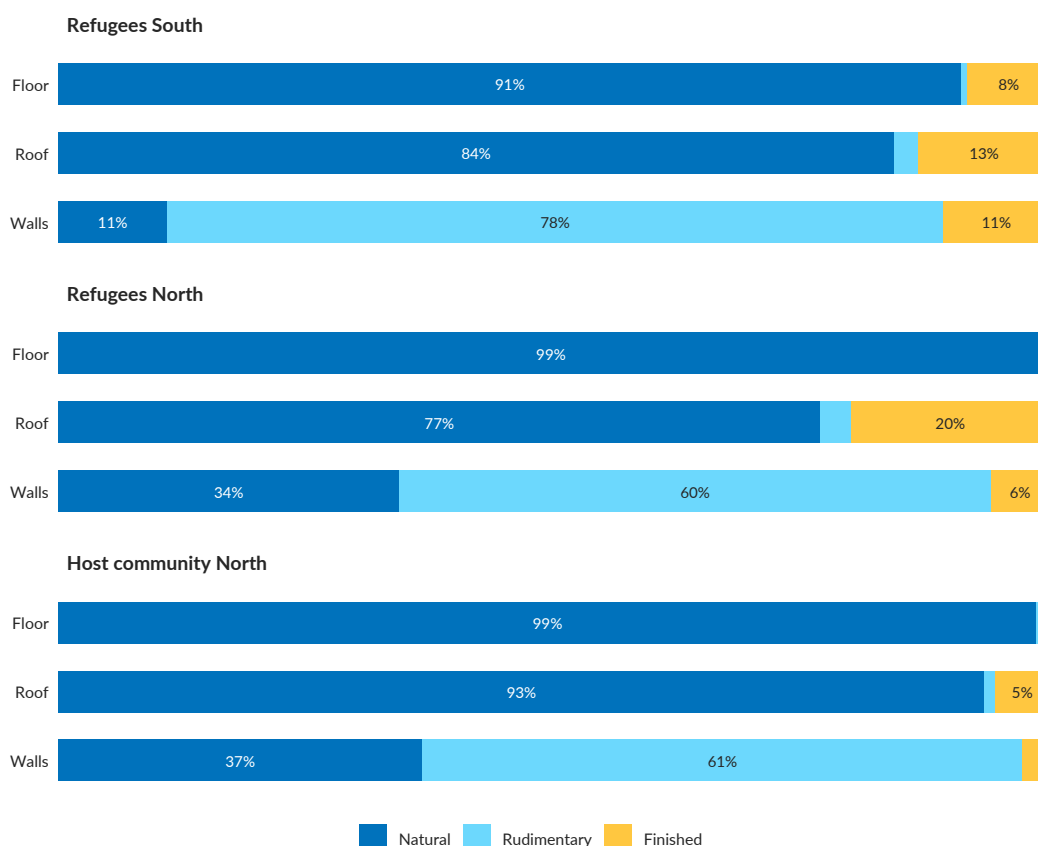
The variation in selected variables reveals meaningful differences in asset ownership and housing characteristics across the three population groups studied. Despite these differences, the general tendency for households across all groups to own similar types of assets and rely on comparable fuel, water, sanitation, and building materials provides confidence that cross-group comparisons are both possible and valid.

Assets rarely owned by host community households also tend to be rare among refugee groups, mostly below 1 per cent. This analysis supports the validity of calculating a unified wealth index across the full sample, from which group-specific indices can be derived. This method offers a replicable way to generate comparable relative wealth measures for refugees and host communities in South Sudan and similar contexts.

### Calculation of the Wealth Index

This section outlines the steps followed to construct a wealth index from FDS data, following the core principles of the DHS methodology, while adapting to the FDS context. The methodology is designed to be replicable and adaptable across countries where FDS or similar survey data are available.

**Figure 16. Housing material**



### Variable Selection and Preprocessing

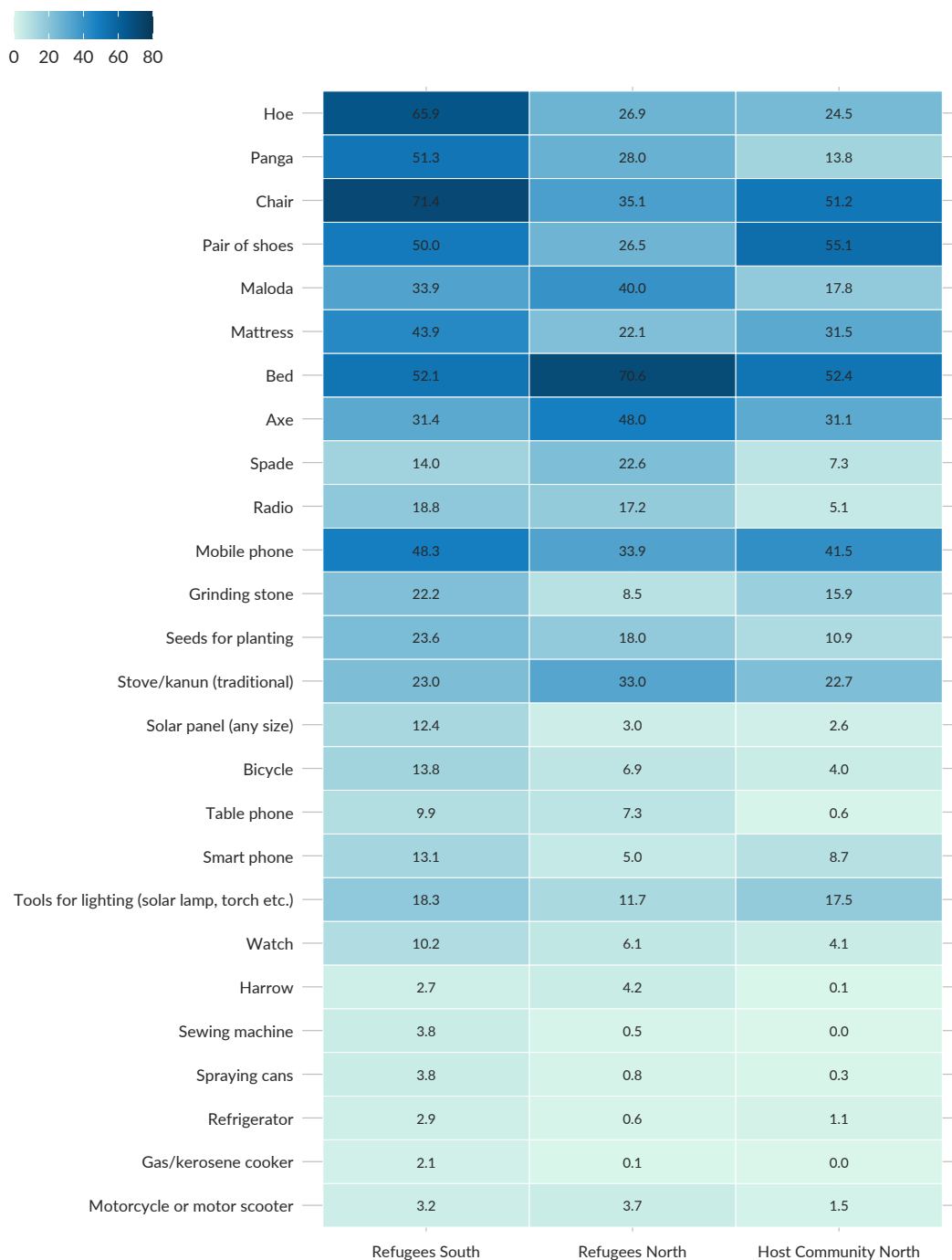
The initial pool of variables is selected from the Forced Displacement Survey dataset, aiming to reflect core dimensions of household wealth as used in DHS surveys. This includes access to utilities, housing characteristics, ownership of durable goods, livestock, and financial access.

Where relevant, FDS variables are transformed or adapted to improve comparability and analytical validity.

**Access to Services** If not already included in the dataset (check for FDS datasets after 2024), the following binary indicators are constructed:

- ◆ **improved\_dw:** Access to an improved drinking water source, which includes piped water (into dwelling, yard/plot, piped to neighbour, or to a public tap/standpipe), tube wells/boreholes, hand-pumps, protected wells and springs, rainwater, tanker-trucks with water from protected source, carts with small tank, water kiosks, bottled water, and sachet water.
- ◆ **Improved\_dw\_final:** Access to an improved drinking water source located on premises or reachable in less than 30 min.
- ◆ **Improved\_san:** Access to improved sanitation facilities, which includes flush toilets, ventilated improved pit latrines, pit latrines with slab and composting toilet
- ◆ **Clean\_cookingfuel:** Use of clean cooking fuel, which includes cooking with solar cookers, electric stoves, piped natural gas stoves, biogas stoves, LPG/cooking gas stoves or using ethanol. (note: commonly found as variable RBM20802 in FDS).
- ◆ **crowding\_cat:** Binary indicator for overcrowding, defined as more than three household members sleeping in the same room.

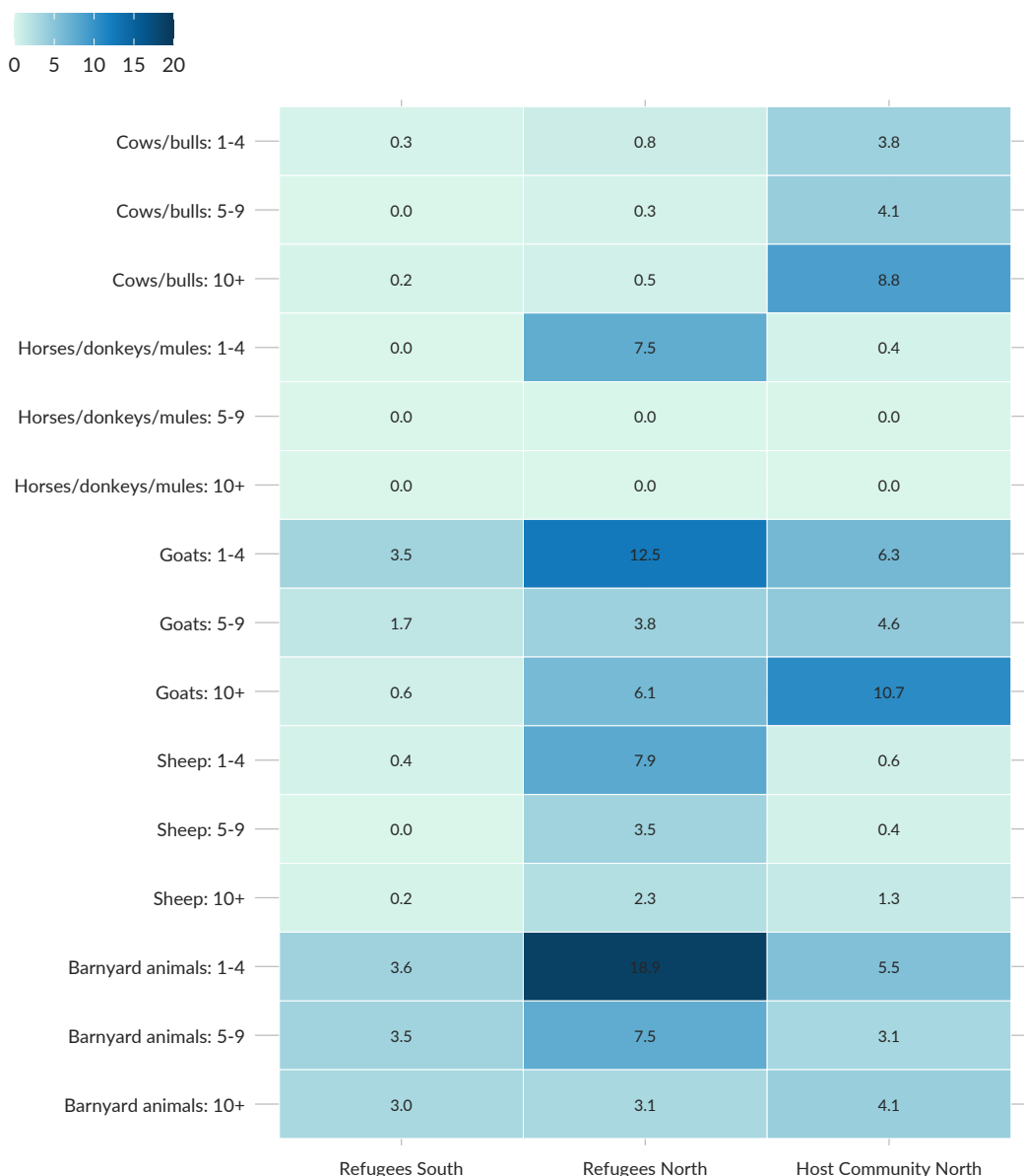
**Figure 17. Asset ownership**



These variables are constructed manually where not already present. In newer FDS datasets, they may already be included and require no additional calculation.

- Toilet Sharing (*shared\_san*):  
 The aggregate *Improved\_san* indicator does not consider whether a toilet is shared. Therefore, a separate variable *shared\_san* is created. In cases where respondents indicate they do not use a toilet, they are not asked whether the facility is shared. For these cases, missing values in the sharing question (*SAN05*) are replaced with 0 (not shared), under the assumption that these households are not using shared facilities.
- Electricity:  
 The electricity variable follows the SDG 7.1.1 definition of access to reliable electricity. A

**Figure 18. Livestock ownership**



household is considered to have electricity access (Tier 1 or above) if it either (a) receives power from the grid or mini-grid for at least one hour in the evening, or (b) uses an off-grid source (e.g., solar system, generator, battery) capable of powering at least one basic appliance. This indicator will be pre-calculated in future FDS datasets.

**Housing and Land Ownership** Land ownership can be a key indicator of wealth, but its inclusion in the wealth index must be carefully considered in contexts involving displaced populations. In countries where refugees are not legally permitted to own land, including this variable can introduce bias and should be avoided. Therefore, land ownership is only included in the calculation of the wealth index in countries where refugees have a de jure or de facto right to own land. To operationalize this, we apply a threshold requiring that land ownership must be reported by at least 5 per cent of refugee households in the survey data. This helps ensure that land ownership reflects meaningful variation in wealth, rather than simply mirroring structural access barriers. This approach supports consistency across contexts while avoiding skewed results due to legal or institutional constraints.

When land ownership is retained, further adjustments may be required due to the structure of the questionnaire:

- In earlier FDS versions, households that reported owning the land they live on (`Land02 = 1`) were often not asked about house ownership (`Land01`), resulting in missing values for the latter. In such cases, we impute house ownership by assuming that households that own the land also own the house and recode missing `Land01` values as 1 (Yes) if `Land02 = 1`.
- If land area is included as a continuous variable, it is constructed by summing the size of all land plots reported by the household. These are then converted into a consistent unit, such as hectares, resulting in the variable `agricultural_land_area`. Unit conversions must be country-specific, based on how land size is reported in the survey.

**Dwelling Materials Imputation** Information about wall, roof, and floor materials is used to assess housing quality. However, in South Sudan and potentially other countries, respondents living in certain housing types were not asked to specify these materials. In South Sudan, missing values in wall (`HH02`), roof (`HH03`), and floor (`HH06`) materials are imputed based on housing type (`HH01`), using typical material assumptions:

- If `HH01 = 3` (traditional house such as tukul), we assume:
  - Walls = mud bricks
  - Roof = thatch
  - Floor = earth/sand
- If `HH01 = 4` (tent), we assume:
  - Walls = no walls
  - Roof = no roof
  - Floor = earth/sand

These assumptions reflect likely material usage and the importance of these variables for the wealth index.

**Livestock Ownership** Given the non-linear relationship between livestock ownership and wealth, continuous livestock counts (variables starting with `Assets04`) are transformed into categorical bins of 1-4, 5-9 and more than 10 animals. Following DHS practice, this reflects increasing marginal value with larger herd sizes. Categories with ownership rates below 1 per cent are dropped to reduce noise.

**Other Asset Variables** All available asset variables (that are owned by at least one household) will be used in the PCA. These variables have to be recoded from 1/2 (Yes/No) to 1/0 and missing values are replaced with 0, assuming absence of the item. This also applies to bank account access (`HH_30a`).

**Final Data Assembly** Once all transformations are complete, the final dataset is compiled to include a unique household identifier, regional and population group identifiers (e.g., refugee/host, north/south, urban/rural), and all selected asset and housing variables, both binary and continuous.

Before running the PCA, all variables with zero variation, (i.e. variables not owned by any household) must be excluded, as they do not contribute any information to the analysis and cause errors in the computation. This cleaned and fully assembled dataset is then used as the input for the principal component analysis in the next stage of the wealth index calculation.

## PCA and Quintile Calculation

To construct the wealth index, Principal Component Analysis (PCA) is applied using the `principal()` function from the `psych` package in R. Given the mix of binary and continuous variables, the analysis is run with `cor = "mixed"` to ensure appropriate handling of variable types. If the final dataset includes only binary variables, `cor = "tet"` (tetrachoric correlation) can be used instead. Using the default

Pearson correlation (`cor = "pearson"`) can distort results in datasets dominated by binary indicators and should be avoided.

The first principal component is extracted and used as the wealth index. Each household receives a score representing its relative position in the asset distribution.

Households are then ranked by their wealth score and divided into five equal groups, from poorest (Q1) to wealthiest (Q5) quintile. These quintiles are calculated using the full analytical sample to allow for comparisons across subgroups.

Once assigned, the distribution of population groups (e.g. refugees, hosts, urban/rural) across quintiles can be analysed and visualized to highlight differences in living standards.



Published April 2026. Some rights reserved.

This work is made available under the Creative Commons Attribution 4.0 IGO license (CC BY 4.0 IGO) <https://creativecommons.org/licenses/by/4.0/>.

In any use of this work, there shall be no suggestion that UNHCR endorses any specific organization, products, or services.

If you create a translation of this work, please add the following disclaimer: This translation was not created by the United Nations High Commissioner for Refugees (UNHCR). UNHCR is not responsible for the content or accuracy of this translation.

If the work is adapted, please add the following disclaimer along with the citation: This is an adaptation of an original work by the United Nations High Commissioner for Refugees (UNHCR). Views and opinions expressed in the adaptation are the sole responsibility of the author or authors of the adaptation and are not endorsed by the UNHCR.

Please cite the work as follows: Leopold, S., Khan, A., and Brock, P. (2026). *Deriving an asset-based wealth index from refugee and host community survey data*. FDS Methodological Paper No. 2. Copenhagen, Denmark: United Nations High Commissioner for Refugees.

The use of the UNHCR logo is not permitted.

Nothing herein shall constitute or be construed or considered to be a limitation upon or waiver, express or implied, of any of the privileges and immunities of the United Nations including its subsidiary organs, or of the United Nations High Commissioner for Refugees (as a subsidiary organ of the United Nations).

This document along with further information on Forced Displacement Survey is available on UNHCR's statistics website: <https://www.unhcr.org/fds>.



**Forced  
Displacement  
Survey**



©2026 United Nations High Commissioner for Refugees

Some rights reserved (see previous page).

Statistics and Information Management Section  
UNHCR Global Data Service  
UN City, Marmorvej 51  
2100 Copenhagen, Denmark  
stats@unhcr.org

Further information on Forced Displacement Survey  
is available on UNHCR's statistics website:  
<https://www.unhcr.org/fds>