STATA

Canadian Stata Conference 2026

6 November 2026
Vancouver, BC, Canada
Share. Learn. Apply.

Building Household Wealth Indices in Stata:

A Reproducible Workflow for Measuring Health Inequalities in Low-Resource Settings

Case Study: Measuring Equity in Access to Cervical Cancer Care in Maputo, Mozambique, Sub-Saharan Africa

Azevedo-Fernandes[1], MD, MPH; Baltazar Chilundo[2], MD, PhD, PHS

[1] Faculty of Social and Economic Sciences, Mozambique Technical University, Mozambique
[2] Faculty of Medicine, Eduardo Mondlane University, Mozambique

Research location and study setting: Oncology Department, Maputo Central Hospital, Maputo, Mozambique (Sub-Saharan Africa)



Health Equity
Reproducible Research
Better Data
Stronger Decisions

1. BACKGROUND

- Socioeconomic position influences access to primary health care, including cervical cancer screening, and may contribute to diagnostic delays and later-stage disease in low-resource settings.
- Household asset-based wealth indices provide a practical and cost-effective measure of socioeconomic position when income data are unavailable or unreliable.

2. OBJECTIVE & HYPOTHESIS

Objective

To demonstrate a reproducible Stata workflow for constructing a household wealth index and to examine its association with prolonged diagnostic delay (≥ 180 days) among women with cervical cancer in Mozambique.

Hypothesis

Women from poorer households have higher odds of prolonged diagnostic delay (≥ 180 days) for cervical cancer.

3. DATA & METHODS

Design: Cross-sectional study of 189 women with cervical cancer.

Setting: Oncology Department, Maputo Central Hospital.

Stage: Known for 165 women (87.3%); 24 (12.7%) missing/indeterminate.

Household asset indicators (14): Housing, utilities, durable assets, transport, communication.

Analysis in Stata: Data cleaning, PCA, wealth quintiles (Q1–Q5), and logistic regression for prolonged diagnostic delay (≥ 180 days).

4. FROM ASSETS TO WEALTH INDICES IN STATA

1. **Prepare data**
   - Recode asset variables (0/1)
   - Check missing data
   - Standardize variables
2. **Principal component analysis (PCA)**
   - *pca* asset1–asset14
   - Assess adequacy (KMO, Bartlett's test)
   - Obtain wealth score (PC1)
3. **Create quintiles**
   - Predict PC1 score
   - Divide into quintiles (Q1–Q5)
4. **Analyse association**
   - Logistic regression (delay ≥ 180 days)
   - Adjust for key covariates (age, education, residence, distance to PHC, pre-diagnosis visits, borrowing money)

PCA model (first principal component)

$$W_i = \sum_{k=1}^{K} \alpha_k \left(\frac{x_{ik} - \bar{x}_k}{s_k} \right)$$

W_i = wealth score for household i
x_{ik} = value of asset k
$\bar{x}_k$ = mean of asset k
s_k = standard deviation of asset k
α_k = PCA loading (weight)

Logistic regression (prolonged diagnostic delay)

$$\text{logit}\left(P\left(Y_i = 1\right)\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k$$

$Y_i = 1$ (delay ≥ 180 days) 0 (delay < 180 days)

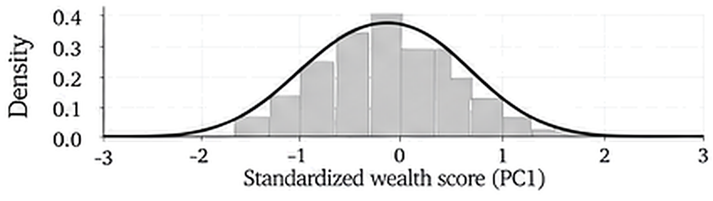
5. RESULTS

5.1 PCA results (N = 189)

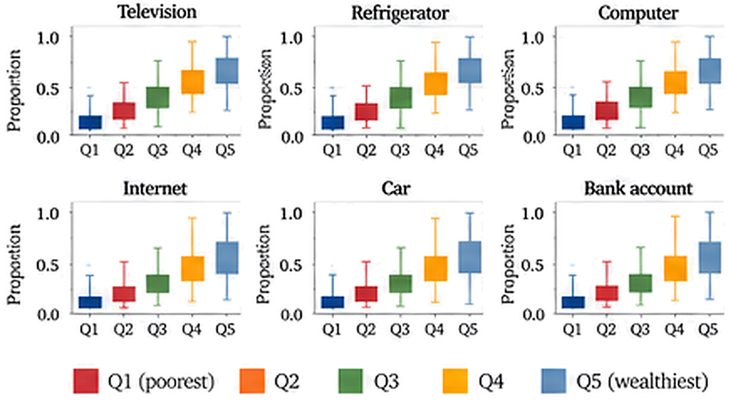
Component	Eigenvalue	Variance explained (%)	Cumulative variance (%)
1	3.46	24.7	24.7

KMO = 0.749 | Bartlett's test of sphericity: $\chi^2(91) = 483.94$, $p < 0.001$

5.2 Distribution of wealth scores (N = 189)



5.3 Selected household assets by wealth quintile (N = 189)



5.4 Multivariable models for prolonged diagnostic delay (diagnosis ≥ 180 days, N = 142)

Outcome	Comparison	Model 1 aOR (95% CI)	Model 2 aOR (95% CI)	Model 3 aOR (95% CI)	p-value (Model 3)
Diagnostic delay (≥ 180 days)	Q2–Q5 vs. Q1 (poorest)	1.34 (0.55–3.29)	1.48 (0.57–3.83)	**1.57 (0.60–4.13)**	0.360

No independent association between household wealth and prolonged diagnostic delay (diagnosis ≥ 180 days).
Model 3: aOR 1.57 (95% CI 0.60–4.13), p = 0.360.

Adjustments: age, education, residence, distance to PHC, pre-diagnosis visits, borrowing money. Household wealth specified as Q2–Q5 versus Q1 (poorest). All models are two-sided.

6. KEY FINDINGS

- PCA produced a reliable household wealth index (KMO = 0.749; 24.7% variance explained).
- Household wealth was not independently associated with prolonged diagnostic delay (diagnosis ≥ 180 days) (Model 3: aOR 1.57, 95% CI 0.60–4.13; p = 0.360).
- Qualitative findings indicated that diagnostic delays reflected shared health system barriers (e.g., navigation failures, repeated visits, administrative issues and financial constraints) across socioeconomic groups.

7. CONCLUSIONS

- A reproducible Stata workflow can construct and apply a household wealth index in a low-resource setting.
- Household wealth was not independently associated with prolonged diagnostic delay among women attending tertiary oncology care, suggesting that common health system bottlenecks affect women across socioeconomic groups.

Interpretation: This is a tertiary-care cohort and may not reflect population-level equity.

8. FULL STUDY (OPEN ACCESS)

Navigating cervical cancer care in Maputo: a mixed-methods analysis of socioeconomic inequities in patient pathways and health system barriers. Azevedo-Fernandes & Chilundo. 2026. BMC Women's Health 26:353. https://doi.org/10.1186/s12905-026-04550-8

9. REFERENCES

1. Filmer D, Pritchett LH. Estimating wealth effects without expenditure data—or tears. Demography. 2001;38:115–132.
2. Vyas S, Kumaranayake L. Constructing socio-economic status indices: how to use principal components analysis. Health Policy Plan. 2006;21:459–468.
3. Azevedo-Fernandes F, Chilundo B. Navigating cervical cancer care in Maputo: a mixed-methods analysis of socioeconomic inequities in patient pathways and health system barriers. BMC Women's Health. 2026;26:353.

Data for a more equitable world. stata.com