

Closing the Universal Health Coverage Gap through Efficiency: A Bias-Corrected Bootstrap DEA Analysis of Health-System Performance in Sub-Saharan Africa

Ernestina Armah^{1*}; Reindolph Osei-Anim²; Lydia Agbeko-Dawuso³; Alice Aluwah Blay⁴; Janet Sintim Aboagye⁴

¹Department of Nursing and Midwifery, Christian Service University, Kumasi, Ghana.

²Department of Accounting and Finance, Christian Service University, Kumasi, Ghana.

³Department of Nursing and Midwifery, Fountainhead University College, Tema, Ghana.

⁴Department of Nursing and Midwifery, Christian Service University, Kumasi, Ghana.

*Corresponding Author's Email: earmah@csuc.edu.gh, oseianim@rocketmail.com

Abstract

Although there is a significant gap in universal health coverage (UHC) in Sub-Saharan Africa, it is unclear how much of the region's current health resources could be used to increase service coverage through increased efficiency. An output-oriented variable-returns-to-scale data envelopment analysis (DEA) was estimated at a 2022 benchmark of 46 countries using an unbalanced panel of 48 Sub-Saharan African countries (2000–2023) without imputed values. Smoothed bootstrap bias correction was then applied (1,000 replications) which founded structural correlates using truncated-bootstrap regression, and cross-checked the frontier using scale-efficiency, jackknife, governance-PCA, stochastic-frontier, matched-year, and two-way fixed-effects analyzes. The average UHC service coverage was 40.33. The mean country-specific corrected radial output-expansion potential was 21.8%, and the bias-corrected efficiency was 0.834 as opposed to a traditional 0.889. Efficiency varied not by income ($p=0.488$) but by financing mode (external 0.884, government 0.863, private/out-of-pocket 0.784; $p=0.005$). Lower inefficiency was linked to government health spending as a percentage of the government budget (-0.152 ; 95% bootstrap CI -0.330 to -0.008). The stochastic-frontier cross-check yielded a mean efficiency of 0.858 with significant concordance with DEA rankings ($p=0.651$; $p<0.001$), and efficiency was stable between matched 2018 and 2022 countries (Wilcoxon $p=0.318$; Spearman $\rho=0.752$). In conclusion, the estimated efficiency opportunity is larger after bias correction. Efficient systems are distinguished from inefficient ones by institutional capacity, operating scale, and financing architecture rather than just revenue. UHC service coverage could be increased without solely depending on new funding by lowering point-of-service payments, bolstering domestic pooled financing, enhancing public budget allocation, and investing in health-data systems.

Keywords: Bootstrap Data Envelopment Analysis; Health Financing; Health-System Efficiency; Universal Health Coverage; Institutional Quality; Sub-Saharan Africa

How to Cite: Armah, E., Osei-Anim, R., Agbeko-Dawuso, L., Blay, A. A., & Aboagye, J. S. (2026). Closing the Universal Health Coverage Gap through Efficiency: A Bias-Corrected Bootstrap DEA Analysis of Health-System Performance in Sub-Saharan Africa. *African Journal of Health, Nursing and Midwifery*, 9(4), 62–80. <https://doi.org/10.52589/AJHNM-MQOEFR>

INTRODUCTION

Universal health coverage (UHC) is the organizing health-system goal of SDG 3.8, in which people should have access to required quality health services without facing financial hardship. The worldwide advancement is falling short. Since 2015, progress at the level of service coverage has stalled, while catastrophic out-of-pocket expenditure (OOPE) has been increasing (World Health Organization & World Bank, 2023). This is particularly difficult in Sub Saharan Africa, where health systems are constrained in fiscal space, continuing staff shortages, high disease burden and are dependent on household and external funds.

In environments with limited resources, efficiency is as crucial a policy choice as the amount of money spent. Efficiency questions how much can be achieved from current resources as much as possible. At the system level, the DEA is the most commonly used frontier method in comparative health-system efficiency analysis, and system-level reviews also highlight sensitivity to choice of input-output, data quality and the validity of second-stage inference (Mbau et al., 2023; Dlouhy & Havlik, 2024). African studies indicate that significant efficiencies are possible, and meta-analysis indicates that efficiencies can be unlocked to free up capacity for expansion of services (Nabyonga-Orem et al., 2023).

Financing systems are important because the method by which funds are raised and aggregated can impact access and production incentives. Households face financial risk when paying out-of-pocket (OOP) and this can be a disincentive for using health services; payment methods that provide financial protection include prepayment and risk pooling (Saksena et al., 2014; Ifeagwu et al., 2021). Governance also plays a role: Institutions are responsible for whether public and external resources are converted into staff, medicines, service capacity and finally coverage. Structural efficiency analysis of financing and institutional quality as structural characteristics and not peripheral controls, is therefore needed in the wider African health agenda to inform the health agenda (Agyepong et al., 2017).

Four issues remain unresolved for Sub-Saharan Africa despite recent UHC-efficiency studies offering significant benchmarks: finite-sample bias in deterministic DEA scores; the distinction between short-term within-country change and persistent cross-country structural relationships; selective missingness in workforce data; and the stability of frontier conclusions across alternative technologies and stochastic methods. The usefulness of bootstrap DEA for a 30-country cross-section was shown by earlier research, such as Arhin et al. (2023), around 2015. Later global and African studies expanded the evidence for UHC-efficiency. With a 48-country, 2000-2023 data architecture, a nearly region-wide 2022 frontier, explicit missingness diagnostics, scale-efficiency decomposition, truncated-bootstrap second-stage inference, two-way and lagged fixed effects, matched-year analysis, leave-one-country-out frontier stability, governance-PCA robustness, and a stochastic-frontier cross-check, the current study expands on previous research.

Consequently, the study asks five questions in a row. First, how representative and thorough is the data used to determine the overall efficiency of the health system in the region? Second, after accounting for finite-sample bias, how well do SSA health systems translate labor and spending resources into UHC service coverage in 2022? Third, are efficiency levels systematically different by financing architecture, subregion, or income group? Fourth, which institutional, financial, and socioeconomic characteristics are associated with bias-corrected inefficiency, and do these associations persist over time across countries? Fifth, how resilient are the findings to matched-year comparisons, governance measurement, jackknife deletion, stochastic-frontier estimation, and alternative DEA technologies? This sequence distinguishes

between measurement, frontier estimation, structural association, longitudinal evidence, and robustness rather than treating them as a single, undifferentiated exercise.

LITERATURE REVIEW

Universal Health Coverage, Service Coverage and Protection from Financial Risk

UHC is a package of access to basic services plus protection from financial hardship. These dimensions are conceptually related, but are distinct from each other in practice. Service coverage indicators measure whether a population receives the interventions they need while financial protection indicators measure exposure to catastrophic or impoverishing health payment. The two dimensions are related, as can be seen in the example of a system expanding service use but also putting poor households in situations where they must pay for services they can't afford, or where poor households are not using services but it seems like they are because they are not paying for them. (Saksena et al., 2014). The threshold and welfare definition are also sensitive to measurement of financial protection. Recent evidence also shows that out-of-pocket payments for primary-care services can generate substantial financial hardship in low- and lower-middle-income settings (Bolongaita et al., 2023; Wagstaff et al., 2020). Methodological work further shows that estimates of catastrophic expenditure and SDG 3.8.2 financial protection are sensitive to indicator definitions and measurement choices (Hsu et al., 2018; Guo & Grépin, 2026). Grépin et al. (2020) show the limitations of the catastrophic-health-expenditure indicator for tracking progress towards UHC. This means that the present study does not impose financial hardship observations onto the DEA output when they are sparse. The main frontier employs the consistently available UHC service coverage index, and data on financial hardship are included as a secondary descriptive characteristic and coverage is reported openly.

Health-System Efficiency and Frontier Methods

DEA is a measure of relative efficiency for similar decision making units that does not assume a particular parametric production function. Its value in comparison to other health systems is that it has multiple inputs and multiple outputs, and the estimated frontier that is sample dependent and deterministic. However, simulation evidence in low resource settings warns that efficiency may be confused with measurement error or noise (Di Giorgio et al., 2016). With this in mind, all efficiency estimates should be treated as relative efficiency measures rather than causal or welfare measures. The difference between constant and variable returns to scale is a real difference. There are instances where a country operating at different size of the system can be compared relative to a country operating at a similar size of the system, which is useful in a region from small island countries to very large countries – such as situations where variable returns to scale (VRS) apply. UHC analysis is also appropriate as an output approach because the policy question is how much more coverage can be achieved within current resource envelopes, whereas input orientation asks how much resource use can be reduced while holding outputs constant. Systematic-review evidence likewise shows that improving efficiency in health spending depends on how available resources are converted into services and health-system outputs (Brieki et al., 2024).

The estimated frontier's deterministic nature and sensitivity to extreme observations is a major inferential drawback of traditional DEA. Because of this, finite samples often overestimate efficiency, especially when the frontier is supported by a small number of decision-making units. Bootstrap techniques allow for bias correction and uncertainty intervals by treating the frontier as estimated rather than fixed. Because they respect the bounded support and

dependence created by the first-stage frontier, truncated-bootstrap procedures are better than naive regressions on generated efficiency scores in second-stage work (Simar & Wilson, 1998, 2007; Badunenko & Tauchmann, 2019). Because it distinguishes between symmetric noise and one-sided inefficiency under explicit distributional assumptions, a stochastic frontier offers a complementary parametric check (Aigner et al., 1977; Di Giorgio et al., 2016).

Evidence From African Health Systems

There is a long history of the existence of significant differences in technical efficiency among African national health systems. Kirigia et al. (2007) reported large differences in productivity and efficiency within the continental health systems. Related evidence from Sub-Saharan Africa also documents substantial cross-country variation in health-system efficiency and productivity (Ibrahim et al., 2018). A second systematic review found that much of the research on the efficiency of the health system at the system-level has focused on wealthier countries, and that governance, macroeconomic conditions, demographic characteristics and health-system configurations repeatedly proved to be important drivers (Mbau et al., 2023). Nabyonga-Orem et al. (2023) also find significant efficiency gaps that could be realised in the WHO African Region. Recent UHC-focused activities focus the policy discussions. Jordi et al. (2020) found that there is very wide variation in the efficiency of moving towards UHC between different income levels of countries. Using UHC as an output, Eze et al. (2024) noted inefficiencies across the sample, with high levels of inefficiency globally. In a dynamic DEA application, Şener (2025) found that the efficiency of the SSA increased at some periods and decreased at the final period. This literature indicates that there is no necessary link between resource abundance and productive conversion to coverage, and that timing and institutional context is relevant.

Research Gap and Contribution

Therefore, the lack of another DEA application is not the only reason for the empirical gap. It is the requirement to integrate statistically rigorous frontier inference with a transparent, regionally wide data structure. Compared to UHC, financing, mortality, and macroeconomic indicators, workforce and hospital-bed series are significantly less complete; forcing all variables into a balanced panel would necessitate significant imputation and produce false precision. This study uses a layered design to address that issue: an annual efficiency panel only where the core inputs are jointly observed, a 2022 frontier with 46 countries that have the entire baseline input bundle, and a full 48-country descriptive panel. By using the smoothed bootstrap bias-corrected frontier as the headline estimator, breaking down the efficiency of CRS and VRS, testing missing-data selection, estimating structural correlates using truncated bootstrap, separating cross-sectional structure from within-country fixed-effects evidence, and putting the results through jackknife, PCA-governance, matched-year, specification, and stochastic-frontier checks, it then outperforms a deterministic benchmark. Since financial-hardship observations only cover 13.5% of potential country-years, financial protection is treated separately rather than being forced into the production frontier. As a result, the estimated frontier measures efficiency in producing SDG 3.8.1 service coverage rather than the complete two-dimensional UHC construct.

DATA AND METHODS

Study Design And Sample

1,152 possible country-year observations are produced by the analysis, which covers 48 Sub-Saharan African nations observed over a 24-year period (2000-2023). It is divided into three layers: (i) a comprehensive panel that describes resources, UHC coverage, financing, governance, and outcomes; (ii) a workforce-augmented DEA benchmark for 2022, where 46 out of 48 countries have the full baseline input bundle; and (iii) an annual efficiency panel that is only created for years when at least 15 countries have jointly observed baseline inputs. At every analytical step, this design makes use of the largest observed sample without filling in any gaps. The WHO-World Bank UHC service coverage index (SDG indicator 3.8.1) is the main result. Because population-facing financial-hardship observations only account for 13.5% of the panel, financial-protection data is analyzed independently to avoid confusing SDG 3.8.2 financial protection with SDG 3.8.1 service coverage. The two countries omitted from the 2022 benchmark because they did not meet the complete-case input requirements are Somalia and Sudan.

The full descriptive frame comprises 48 countries $\times$ 24 calendar years = 1,152 potential country-years. The 2022 output-oriented VRS frontier is a separate cross-section of 46 complete cases; it is not the denominator for full-panel descriptive means. The 16 annual DEA frontiers comprise 403 observed country-years meeting the input/output completeness screen and the stated minimum peer-count rule. The contemporaneous fixed-effects models use 233 observations from 43 countries after complete-case covariate filtering, whereas the alternative lagged specification uses 390 observations from 47 countries. The latter is not a simple time shift of the same common sample. The exact retained year list, the two excluded 2022 countries, and a variable-by-variable attrition ledger cannot be independently recovered from the manuscript alone and must be checked against the archived analysis dataset before a fully reproducible sample flow can be claimed.

Table 1: Analytical sample flow and observation counts

Stage	Coverage/period	Retained sample	Selection rule / Interpretation
Descriptive frame	2000–2023	48 countries; 1,152 country-years	Country-year frame; missingness retained, without imputation
2022 VRS DEA	2022	46 countries; 46 observations	Complete 3 inputs plus UHC SCI; two exclusions must be verified in dataset
Annual VRS DEA	16 eligible years (exact years unverified)	403 country-years	≥ 15 countries with joint baseline inputs and SCI in a year
Contemporaneous two-way FE	Eligible annual DEA years	233 observations; 43 countries	Complete-case covariates in common contemporaneous model
Lagged specification	Eligible annual DEA years	390 observations; 47 countries	Lagged model; different availability/model sample, requiring an attrition log

Note: Sample counts are those reported in the manuscript, not newly re-estimated. Exact eligible calendar years, two excluded 2022 countries and detailed exclusion reasons require the estimation dataset.

Data Sources and Variable Definitions

World Development Indicators (WDI), Worldwide Governance Indicators (WGI), WHO UHC and health-expenditure series, and associated World Bank health indicators are all used in the analytical dataset. Physicians per 1,000 people, nurses and midwives per 1,000 people, and current health spending per capita in purchasing-power-parity dollars are the baseline production inputs. Because hospital beds per 1,000 people have too little coverage for a required baseline input, they only enter a sensitivity model. The result is the UHC service coverage index; sensitivity analysis uses the RMNCH, infectious-disease, non-communicable-disease, and service-capacity/access sub-indices. Log GDP per capita (PPP), gross secondary enrollment, institutional quality as determined by averaging the six WGI dimensions, out-of-pocket spending share, government health spending as a percentage of total government spending, external health funding as a percentage of current health expenditure, and urban population share are examples of contextual variables. Contextual health outcomes include life expectancy, maternal mortality, and under-five mortality. Table 2 summarizes the availability and coverage of the key analytical variables across the 48-country panel.

Table 2: Data availability and coverage across the 48-country panel, 2000-2023

Variable	Observed N	Coverage
Current health expenditure per capita, PPP	1,112	96.5%
Physicians per 1,000	556	48.3%
Nurses/midwives per 1,000	538	46.7%
Hospital beds per 1,000	317	27.5%
UHC service coverage	1,152	100.0%
Financial hardship (%)	156	13.5%
Out-of-pocket share	1,112	96.5%
Government health spending (% govt exp.)	1,112	96.5%
External health funding (% CHE)	1,109	96.3%
GDP per capita, PPP	1,124	97.6%
Secondary enrolment	651	56.5%
Institutional quality	1,093	94.9%

Note: Coverage is based on 1,152 possible country-year observations. All values are observed; no missing values are imputed.

Figure 1 illustrates the observed country-year coverage of the key analytical variables.

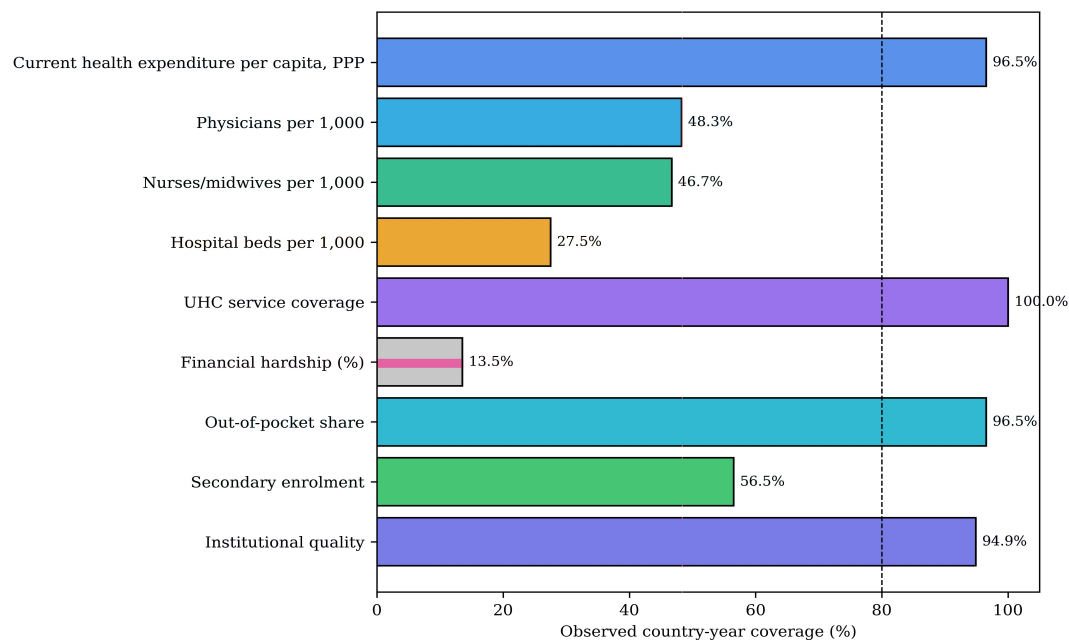


Figure 1: Observed country-year coverage of key analytical variables.

Source: Authors' calculations from the dataset.

Data Integrity, Missingness, and Selection Diagnostics

There were no interpolated, forward-filled, back-filled, adjacent-year observations, or artificially generated missing values. Workforce-complete and workforce-incomplete country-years are compared prior to frontier estimation because missingness may be selective. According to the diagnostics, UHC coverage, GDP per capita, urbanization, and institutional quality are all much higher in complete-workforce observations. The 2022 benchmark, which includes 46 out of 48 countries, is far less vulnerable to this selection issue than the annual efficiency panel, which is positively selected toward better-measured systems. These variations are not the result of missingness; rather, they are descriptive selection diagnostics. Table 3 reports the selection diagnostics comparing workforce-complete and workforce-incomplete country-years.

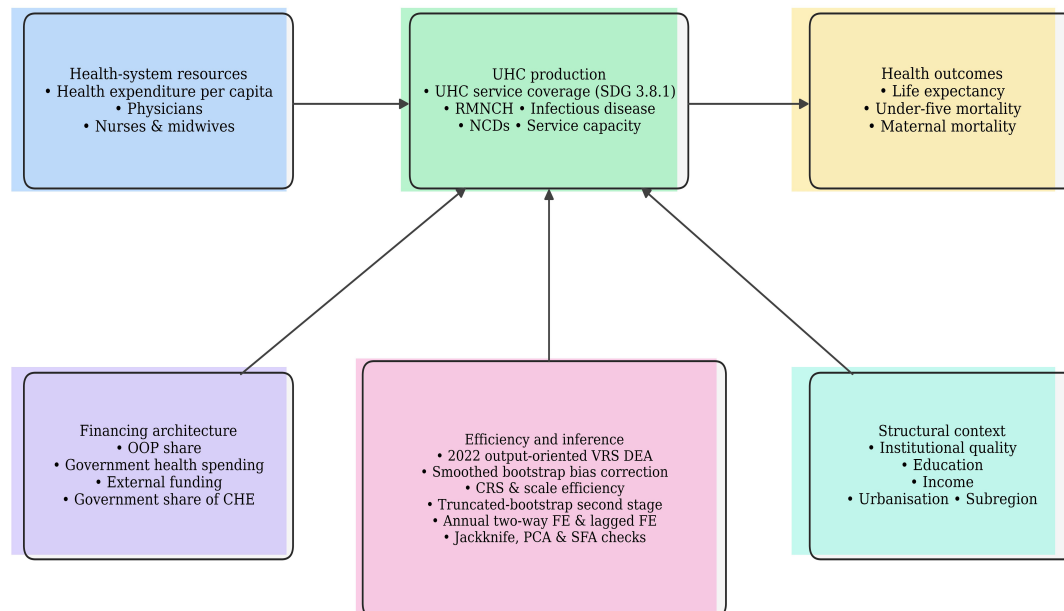
Table 3: Selection diagnostic: workforce-complete versus workforce-incomplete country-years

Variable	Complete Mean	Incomplete Mean	Standardized Mean Difference	Welch P-Value
UHC service coverage	43.784	37.962	0.418	<0.001***
GDP per capita, PPP	5306.102	4264.682	0.189	0.002***
Urban share	41.024	38.661	0.145	0.016**
Institutional quality	-0.546	-0.798	0.397	<0.001***

Note: Welch tests compare observed group means. Standardized mean differences describe the magnitude of imbalance.

Figure 2 summarizes the analytical framework linking health-system resources, UHC production, efficiency estimation, financing architecture, and structural context.

Analytical framework for bias-corrected health-system efficiency and UHC in Sub-Saharan Africa



Financial protection (SDG 3.8.2) is analysed separately because financial-hardship observations cover only 13.5% of country-years.

Figure 2: Analytical framework for bias-corrected health-system efficiency and UHC in Sub-Saharan Africa.

Source: Authors' conceptualization.

Note: Financial protection is monitored separately because financial-hardship data cover only 13.5% of potential country-years.

Output-Oriented Variable>Returns-To-Scale DEA

The output-oriented variable-returns-to-scale (VRS) DEA is estimated by the 2022 benchmark. Because the policy question is how much more UHC service coverage can be produced from the observed expenditure and workforce bundle, output orientation is appropriate. The vast disparities between national health systems are taken into account by VRS. Let x represent the input vector, y the output, ϕ the radial output-expansion factor, λ the peer weights, and θ the reported technical-efficiency score for country 0. The BCC/VRS program is:

The 2022 benchmark offers substantially more complete workforce coverage than most other years and is also after the acute 2020–2021 COVID-19 period. Data completeness is the defensible primary selection rationale; post-pandemic relevance is contextual, not evidence that 2022 was prospectively selected. The 2022 comparison is cross-sectional and therefore does not itself measure recovery or temporal productivity change.

$$\max_{\phi, \lambda_j} \phi \quad \text{s.t.} \quad \sum_j \lambda_j y_j \geq \phi y_0, \sum_j \lambda_j x_j \leq x_0, \sum_j \lambda_j = 1, \lambda_j \geq 0 \quad (1)$$

$$\theta = \frac{1}{\phi}, \quad 0 < \theta \leq 1 \quad (2)$$

A health system that operates on the estimated variable-returns-to-scale (VRS) frontier has a technical-efficiency score of one. The radial output-expansion potential is computed as $100(\phi - 1)$, where $\phi = 1/\theta$, and the simple efficiency shortfall is defined as $100(1 - \theta)$ for each decision-making unit. These metrics capture related but different ideas and should not be used interchangeably. While the former gives the feasible proportional increase in the output index at unchanged observed inputs, while the latter measures the distance of the efficiency score below one. Furthermore, scale efficiency is calculated as the ratio of CRS efficiency to VRS efficiency, and constant-returns-to-scale (CRS) efficiency is estimated. This breakdown makes it possible to distinguish between inefficiencies caused by operating at a non-optimal scale and inefficiencies that are solely technical (Charnes et al., 1978; Banker et al., 1984).

Smoothed Bootstrap Bias Correction

Conventional DEA tends to overestimate efficiency in finite samples and places the frontier on extreme observations. Therefore, the headline frontier estimates finite-sample bias and confidence intervals using a smoothed bootstrap with 1,000 replications. Every pseudo-sample is created around the estimated efficiency distribution, taking into account the projected production frontier and the unit interval. If the traditional score is θ and its bootstrap counterpart is θ^* , bias correction is:

To reproduce the bootstrap fully, the archived implementation should report random seed, smoothing kernel and bandwidth, resampling algorithm, boundary handling, confidence-interval construction and numerical tolerances, in addition to the 1,000 replications already documented. These settings and the original package versions are not available in the supplied materials and are not assumed here.

$$\widehat{\text{bias}}(\theta_i) = \frac{1}{B} \sum_{b=1}^B \theta_{ib}^* - \theta_i, \quad \hat{\theta}_i = \theta_i - \widehat{\text{bias}}(\theta_i), \quad \hat{\phi}_i = \frac{1}{\hat{\theta}_i} \quad (3)$$

For substantive interpretation, the corrected reciprocal expansion factor, bootstrap confidence intervals, and bias-corrected efficiency are employed. Additionally, leave-one-country-out re-estimation examines whether the estimated frontier is unduly dependent on any one nation (Simar & Wilson, 1998; Badunenko & Tauchmann, 2019).

Truncated-Bootstrap Second Stage

Which structural traits are linked to bias-corrected inefficiency in 2022 is the question of the second stage. Instead of using naive Tobit regression or ordinary least squares, a left-truncated normal model is employed because the reciprocal inefficiency factor is bounded below at one. The reciprocal of bias-corrected technical efficiency is the dependent variable:

The estimation is described as a left-truncated normal second-stage model of the reciprocal corrected score, with truncation at one, standardized non-intercept regressors, separate financing-share specifications, and 500 parametric bootstrap replications. The source materials do not disclose optimization options, convergence tolerances, bootstrap resampling seed or whether a specific Simar–Wilson algorithm was implemented identically. Inference should not be described as a verified software-level replication until the code and version information are supplied.

$$\hat{h}_i = \frac{1}{\hat{\theta}_i} = \beta_0 + \mathbf{z}_i' \boldsymbol{\beta} + \varepsilon_i, \quad \hat{h}_i \geq 1, \quad \varepsilon_i \sim N(0, \sigma^2) \quad (4)$$

Non-intercept covariates were standardized, meaning that the coefficients reflect a one-standard-deviation change. A positive sign implies greater inefficiency, and a negative sign means the opposite. Four financing indicators, including the share of OOP, government health budget as a share of the government budget, external health funding, and government health budget as a share of current health expenditure, were introduced as separate models to avoid mechanical collinearity between financing shares. A 500-replication parametric bootstrap was employed. The institutional quality indicators were re-estimated using the first principal component extracted from the six WGI dimensions as a robustness check (Simar & Wilson, 2007; Badunenko & Tauchmann, 2019).

Longitudinal Two-Way Fixed-Effects Analysis

An annual VRS DEA frontier is estimated for each year that at least 15 countries are jointly observed on the three baseline inputs and UHC output. The 16 eligible years are 2004 and 2008–2022, producing 403 country-year efficiency observations. The common complete second-stage panel is estimated using country and year fixed effects and country-clustered standard errors in order to differentiate persistent cross-country structure from within-country change:

The minimum of 15 countries is the screening rule reported in the analysis; whether it was chosen before inspecting completeness is not documented. With three inputs and one output, a commonly used DEA dimensionality guideline is $\max\{m \times s, 3(m+s)\} = \max\{3, 12\} = 12$ decision-making units, so 15 exceeds this heuristic threshold but does not guarantee adequate discriminatory power. Results are consequently treated as conditional on the selected reference sets. Only years that meet both the complete input-output requirement and the minimum-peer threshold enter annual frontier calculations. The precise years and annual country counts require confirmation from the underlying processed panel; they should not be inferred from the total of 403.

$$\theta_{it} = \mathbf{x}_{it}'\boldsymbol{\beta} + \mu_i + \tau_t + \varepsilon_{it} \quad (5)$$

There are 233 observations from 43 countries in the contemporaneous common sample. 390 observations from 47 countries are included in a robustness specification with a one-year lag. Because financing and governance change slowly and may be influenced by health-system performance, these models are interpreted as within-country associations rather than causal estimates.

Computational Reproducibility: The submitted manuscript describes DEA orientation and returns to scale, 1,000 smoothed frontier-bootstrap replications, 500 second-stage bootstrap replications, six standardized governance inputs for PCA, country/year fixed effects with country-clustered inference, and a half-normal Cobb–Douglas stochastic frontier. However, it does not identify software and library versions, random seeds, smoothing bandwidths, solver settings, FE covariance implementation or stochastic-frontier optimizer. Accordingly, these details are not invented here; the archived code/environment is needed to make the estimations executable and to confirm exact attrition counts.

Cross-Method, Matched-Year, and Sensitivity Checks

There are six dimensions to robustness. First, the significance of operating scale is evaluated by CRS and scale efficiency. Second, alternative DEA input-output specifications test sensitivity to hospital beds, life expectancy, multi-output dimensionality, and workforce composition. Third, frontier stability is evaluated using leave-one-country-out jackknife

estimates. Fourth, the OOP model is re-estimated and principal components are used to summarize the six WGI indicators. Fifth, the Wilcoxon signed-rank test and Spearman rank correlation are used in a matched 40-country comparison of 2018 and 2022. Sixth, a noise-tolerant parametric cross-check is provided by a half-normal Cobb-Douglas stochastic frontier:

$$\ln y_i = \beta_0 + \beta_1 \ln(\text{CHE}_i) + \beta_2 \ln(\text{Physicians}_i) + \beta_3 \ln(\text{Nurses}_i) + v_i - u_i \quad (6)$$

where u is non-negative inefficiency and v is symmetric noise. Because health expenditure is measured in current international PPP dollars, a Malmquist productivity index is purposefully not reported; intertemporal productivity decomposition would be complicated by price and PPP fluctuations in the absence of a constant-price health-expenditure series. As a result, period comparisons continue to be descriptive, dividing COVID (2020–2021), pre-COVID (2000–2019), and post-COVID (2022–2023) windows.

RESULTS

Data Integrity, Coverage, and Selection Diagnostics

There are 1,152 country-year observations in the panel that are free of imputation and duplicate country-years. Physicians are observed in 48.3% of country-years, nurses and midwives in 46.7%, hospital beds in 27.5%, and financial hardship in just 13.5%. UHC service coverage and core outcome series are complete. Workforce-complete observations have higher UHC coverage (43.78 versus 37.96; standardized mean difference 0.418; $p < 0.001$), higher GDP per capita (US\$5,306 versus US\$4,265; $p = 0.002$), more urbanization (41.02% versus 38.66%; $p = 0.016$), and better institutional quality (-0.546 versus -0.798; $p < 0.001$). Because it keeps 46 of 48 countries and is significantly less selected than the annual panel, the 2022 benchmark is therefore favored for cross-sectional inference.

Descriptive Profile of Health Systems

The mean UHC service coverage index across all 1,152 country-years is 40.33 (SD 14.20), with a range of 13–81. Of the four pillars of universal health coverage, NCD coverage is greatest (60.30), followed by RMNCH (49.85), service capacity/access (35.09) and coverage of infectious disease (31.29). The mean current health expenditure per person is \$222.37 in PPP terms, physician density is 0.266 per 1,000 and nurse/midwife density is 1.231 per 1,000. The mean out-of-pocket spending is 39.84 percent of current health expenditure, government expenditure on health is 6.79 percent of the government budget, external health funding is 21.79 percent of current health expenditure and the institutional quality is -0.691 on the standardized WGI scale. Table 4 reports the descriptive statistics for health-system inputs, UHC coverage, financing, health outcomes, and contextual variables. The mean 40.327 (approximately 40.33) and standard deviation 14.199 refer to the 1,152 available country-year observations over 2000–2023, not to the 46-country 2022 efficiency sample. The UHC service coverage index (SCI) is a composite indicator on a 0–100 scale, not the percentage of the population receiving care. It also excludes financial-risk protection (SDG 3.8.2), which the study describes separately.

Table 4: Descriptive statistics of health-system inputs, UHC, financing, outcomes, and contextual variables

Variable	N	Mean	SD	Min	Median	Max
Current health	1,112	222.374	272.190	6.106	116.117	1565.915

Variable	N	Mean	SD	Min	Median	Max
expenditure per capita, PPP						
Physicians per 1,000	556	0.266	0.475	0.012	0.116	6.598
Nurses/midwives per 1,000	538	1.231	1.239	0.088	0.790	7.365
Hospital beds per 1,000	317	1.391	1.084	0.150	0.850	4.810
UHC service coverage	1,152	40.327	14.199	13.000	39.000	81.000
RMNCH coverage	1,152	49.847	14.511	11.000	48.000	78.000
Infectious-disease coverage	1,152	31.285	21.928	2.000	29.000	99.000
NCD coverage	1,152	60.300	8.353	18.000	61.000	77.000
Service capacity/access	1,152	35.089	18.044	9.000	31.000	96.000
Financial hardship (%)	156	37.375	14.176	7.000	38.350	77.450
Life expectancy	1,152	58.973	7.007	14.665	59.333	77.237
Under-five mortality	1,152	88.108	47.200	11.600	80.300	489.300
Maternal mortality	1,152	479.547	311.429	40.000	438.000	1662.000
Out-of-pocket share	1,112	39.842	20.805	2.993	39.723	85.726
Government health spending (% govt exp.)	1,112	6.790	3.481	0.734	6.240	32.571
External health funding (% CHE)	1,109	21.790	17.319	0.000	18.133	78.406
Government health spending (% CHE)	1,112	31.533	17.320	3.368	28.453	86.682
Government health spending (% GDP)	1,112	1.610	1.245	0.062	1.220	6.754
GDP per capita, PPP	1,124	4687.180	5540.725	433.064	2564.196	34219.061
Urban population share	1,152	39.621	16.293	8.044	38.611	91.638
Secondary enrolment	651	45.640	24.443	3.348	40.264	112.311
Institutional quality	1,093	-0.691	0.645	-2.190	-0.728	0.908

Conventional And Bias-Corrected 2022 Frontiers

With 14 countries on the frontier, the traditional output-oriented VRS frontier for 46 countries has a mean efficiency of 0.889 (SD 0.112; median 0.907). By moving all conventionally efficient systems below unity and lowering mean efficiency to 0.834 (SD 0.096; median 0.861), bootstrap correction changes the policy arithmetic. The average country-specific corrected radial output-expansion potential is 21.8%, and the simple mean efficiency shortfall increases from 11.1% under the conventional score to 16.6% under the corrected score. Since the latter is calculated using the reciprocal expansion factors rather than $1-\theta$, it is the pertinent output-oriented expansion measure. Suboptimal scale is a significant additional source of performance loss, as evidenced by the mean CRS efficiency of only 0.513 and the mean scale efficiency of

0.576. Table 5 presents the country-level conventional VRS, bias-corrected, CRS, and scale-efficiency estimates for 2022.

For each country i , corrected output-expansion potential is $g_i = 100 \times (1/\theta_{i,BC} - 1)$, where $\theta_{i,BC}$ is that country's bias-corrected output-oriented efficiency score. The reported 21.8% is the arithmetic mean of these country-specific g_i values across the 46 frontier countries. It must not be calculated as $100 \times (1/0.834 - 1)$, since a nonlinear transformation of the mean is generally not the mean of the individual transformations. By contrast, the simple shortfall is $100 \times (1 - \text{mean } \theta_{i,BC}) = 16.6\%$; this is not the radial output gain. Both describe hypothetical frontier opportunities, not realized increases in UHC or a share of people served.

Table 5: Conventional, bias-corrected, CRS, and scale-efficiency estimates for SSA health systems, 2022

Country	UHC	VRS TE	Bias- corrected TE	95% bootstrap interval	CRS TE	Scale Efficiency	Rank	Financing mode
Eswatini	70	1.000	0.947	[0.906, 0.986]	0.170	0.170	1	Government
Senegal	49	1.000	0.946	[0.893, 0.986]	0.661	0.661	2	Private/OO P
Gambia, The	52	0.997	0.941	[0.888, 0.984]	0.860	0.863	3	Government
South Africa	75	1.000	0.931	[0.874, 0.987]	0.112	0.112	4	Government
Mauritius	75	1.000	0.930	[0.877, 0.986]	0.107	0.107	5	Private/OO P
Seychelles	79	1.000	0.924	[0.860, 0.987]	0.101	0.101	6	Government
Rwanda	58	1.000	0.923	[0.856, 0.988]	0.636	0.636	7	External
Ghana	57	0.960	0.919	[0.882, 0.948]	0.423	0.441	8	Government
Zimbabwe	59	1.000	0.916	[0.865, 0.987]	0.793	0.793	9	External
Cabo Verde	71	1.000	0.913	[0.858, 0.986]	0.231	0.231	10	Government
Sao Tome and Principe	60	0.942	0.905	[0.878, 0.930]	0.303	0.321	11	External
Kenya	57	0.934	0.897	[0.870, 0.922]	0.382	0.409	12	Government
Burundi	48	1.000	0.896	[0.795, 0.985]	1.000	1.000	13	External
Malawi	53	1.000	0.894	[0.796, 0.987]	0.920	0.920	14	External

Country	UHC	VRS TE	Bias- corrected TE	95% bootstrap interval	CRS TE	Scale Efficiency	Rank	Financing mode
Congo, Dem. Rep.	42	1.000	0.890	[0.769, 0.987]	1.000	1.000	15	External
Namibia	65	0.931	0.890	[0.848, 0.920]	0.125	0.134	16	Government
Central African Republic	36	0.957	0.890	[0.786, 0.944]	0.939	0.982	17	Private/OO P
Botswana	61	0.923	0.887	[0.858, 0.912]	0.115	0.124	18	Government
Niger	37	1.000	0.886	[0.729, 0.987]	1.000	1.000	19	Private/OO P
Madagasc ar	34	1.000	0.886	[0.723, 0.986]	1.000	1.000	20	External
Chad	24	1.000	0.885	[0.715, 0.987]	0.664	0.664	21	Private/OO P
Zambia	55	0.904	0.862	[0.830, 0.893]	0.423	0.467	22	Government
Lesotho	55	0.896	0.862	[0.836, 0.884]	0.297	0.332	23	Government
Mozambi que	49	0.896	0.861	[0.829, 0.885]	0.639	0.713	24	External
South Sudan	39	0.926	0.859	[0.775, 0.915]	0.905	0.977	25	External
Tanzania	47	0.910	0.853	[0.794, 0.899]	0.788	0.866	26	External
Uganda	52	0.893	0.845	[0.807, 0.881]	0.584	0.655	27	External
Guinea	43	0.886	0.841	[0.795, 0.872]	0.594	0.670	28	Private/OO P
Mali	41	0.889	0.839	[0.790, 0.876]	0.764	0.859	29	Private/OO P
Comoros	53	0.855	0.825	[0.801, 0.844]	0.310	0.362	30	Private/OO P
Liberia	50	0.859	0.822	[0.796, 0.848]	0.425	0.495	31	Private/OO P
Cameroon	47	0.853	0.814	[0.775, 0.843]	0.426	0.499	32	Private/OO P
Congo, Rep.	46	0.831	0.799	[0.770, 0.821]	0.562	0.676	33	Government
Cote d'Ivoire	46	0.829	0.794	[0.760, 0.819]	0.396	0.478	34	Government

Country	UHC	VRS TE	Bias-corrected TE	95% bootstrap interval	CRS TE	Scale Efficiency	Rank	Financing mode
Equatorial Guinea	49	0.824	0.786	[0.755, 0.814]	0.242	0.294	35	Private/OOP
Burkina Faso	48	0.823	0.784	[0.756, 0.812]	0.445	0.541	36	Private/OOP
Sierra Leone	45	0.799	0.763	[0.735, 0.788]	0.513	0.642	37	Private/OOP
Nigeria	47	0.777	0.752	[0.732, 0.768]	0.321	0.414	38	Private/OOP
Eritrea	40	0.793	0.742	[0.686, 0.782]	0.740	0.933	39	Private/OOP
Guinea-Bissau	43	0.744	0.716	[0.693, 0.734]	0.378	0.508	40	Private/OOP
Angola	45	0.738	0.710	[0.689, 0.728]	0.303	0.410	41	Government
Togo	39	0.745	0.697	[0.631, 0.736]	0.470	0.631	42	Private/OOP
Gabon	48	0.710	0.674	[0.646, 0.701]	0.162	0.229	43	Government
Benin	35	0.664	0.629	[0.591, 0.656]	0.547	0.824	44	Private/OOP
Ethiopia	31	0.632	0.593	[0.547, 0.624]	0.601	0.951	45	Private/OOP
Mauritania	35	0.568	0.544	[0.529, 0.559]	0.217	0.381	46	Private/OOP

Note: Baseline inputs are current health expenditure per capita (PPP), physicians, and nurses/midwives. Output is the UHC service coverage index. Rank is based on bias-corrected VRS efficiency.

Figure 3 displays the bias-corrected 2022 VRS DEA efficiency scores together with their 95% bootstrap intervals.

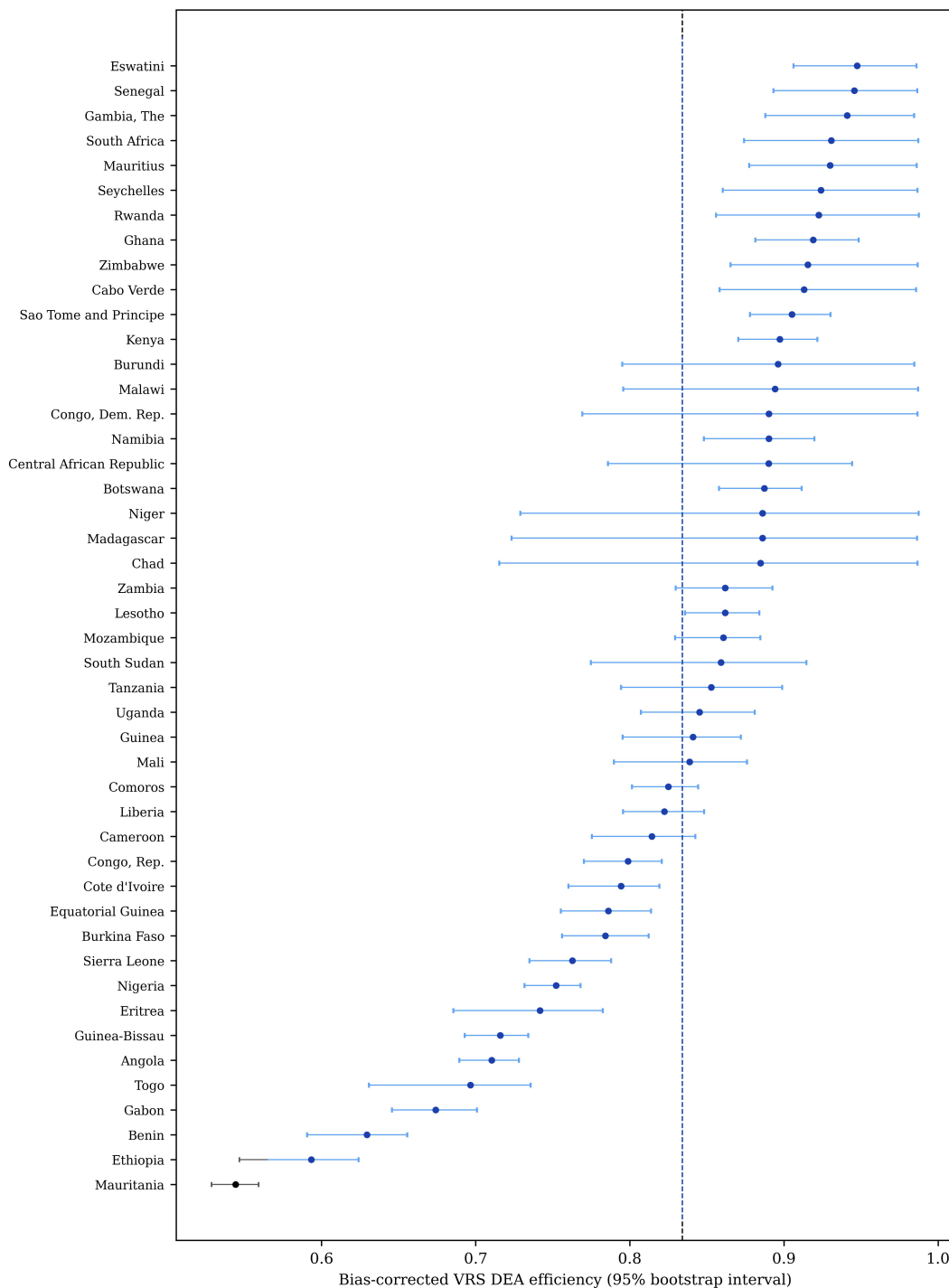


Figure 3: Bias-corrected 2022 VRS DEA efficiency scores and 95% bootstrap intervals.
Source: Authors' bootstrap calculations.

Heterogeneity By Income, Subregion, And Financing Mode

The most obvious difference in bias-corrected efficiency is found in financing architecture. With an average of 0.884 for externally financed systems, 0.863 for government-dominated systems, and 0.784 for private/OOP-dominated systems, the Kruskal-Wallis test rejects equality across dominant financing modes ($\chi^2=10.495$; $p=0.005$). At conventional levels, subregional heterogeneity is borderline ($\chi^2=7.624$; $p=0.054$; not significant at 5%), with West Africa having the lowest subregional heterogeneity (0.799) and Southern Africa having the

highest (0.905). The lack of significant differences between income groups ($p=0.488$) confirms that national income by itself is not a monotonic predictor of technical efficiency. Table 6 summarizes the Kruskal-Wallis comparisons of bias-corrected efficiency by income group, subregion, and dominant financing mode.

The descriptive financing-mode contrast is a Kruskal–Wallis association based on the largest observed 2022 financing share and should not be read as the causal effect of changing financing mechanism. The subregional comparison ($p = 0.054$) does not meet the conventional 5% significance threshold, although it can be described as suggestive if a 10% threshold was specified in advance. No such prespecification is documented here.

Table 6: Kruskal-Wallis comparisons of bias-corrected 2022 efficiency

Grouping	Category	N	Mean TE	Median TE	χ^2	p-value
Income group	High income	1	0.924	0.924	2.433	0.488
Income group	Low income	19	0.843	0.861	2.433	0.488
Income group	Lower middle income	19	0.827	0.853	2.433	0.488
Income group	Upper middle income	6	0.854	0.900	2.433	0.488
Subregion	Central Africa	9	0.817	0.814	7.624	0.054*
Subregion	East Africa	15	0.853	0.862	7.624	0.054*
Subregion	Southern Africa	6	0.905	0.903	7.624	0.054*
Subregion	West Africa	16	0.799	0.808	7.624	0.054*
Financing mode	External	11	0.884	0.890	10.495	0.005***
Financing mode	Government	15	0.863	0.890	10.495	0.005***
Financing mode	Private/OOP	20	0.784	0.800	10.495	0.005***

Note: Dominant financing mode is defined by the largest 2022 share among government financing, OOP expenditure, and external health funding.

Figure 4 illustrates the distribution of bias-corrected efficiency across the three dominant healthcare financing modes in 2022.

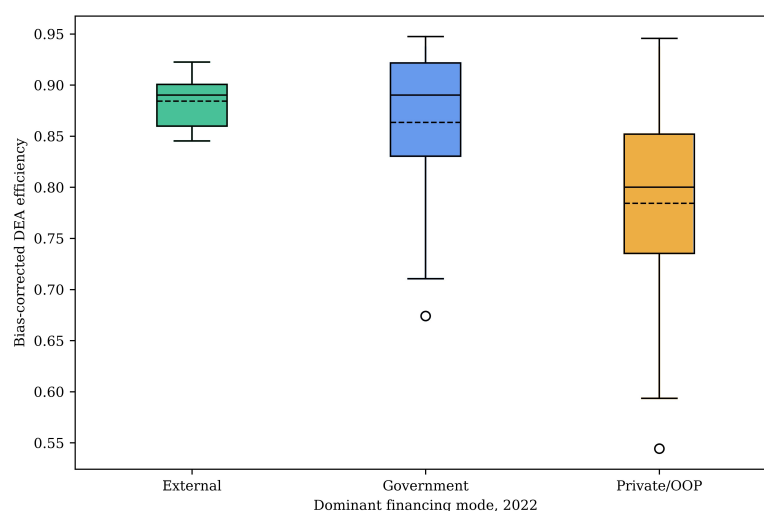
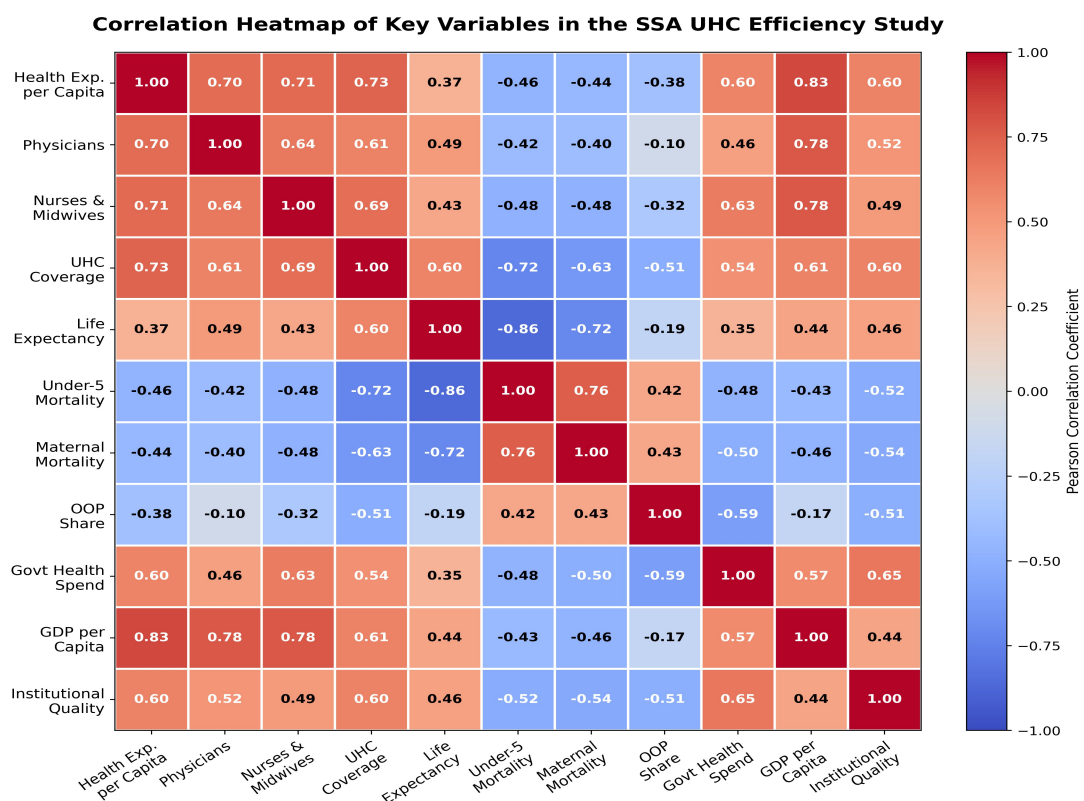


Figure 4: Bias-corrected efficiency by dominant healthcare financing mode, 2022.
Source: Authors' calculations.

Truncated-Bootstrap Correlates of 2022 Inefficiency

Lower inefficiency is indicated by negative coefficients in the truncated-bootstrap models, which model the reciprocal bias-corrected inefficiency factor using 44 countries. The most statistically reliable financing result is government health spending as a percentage of overall government spending, with a standardized coefficient of -0.152 and a 95% bootstrap interval of -0.330 to -0.008. Lower inefficiency in the external-funding specification is also linked to institutional quality (-0.142; 95% CI -0.343 to -0.018). While external funding and the government's share of CHE are negative but imprecise, the OOP coefficient is positive but narrowly includes zero in the baseline governance specification (0.107; 95% CI -0.004 to 0.234). These estimates do not imply causation, but they do support a structural financing-and-governance interpretation.

Before interpreting the multivariate second-stage estimates, the correlation structure among the principal health-system resource, universal health coverage, health-outcome, financing, socioeconomic, and institutional variables was examined. Figure 5 presents the Pearson correlation heatmap for the key variables used in the study. The heatmap provides a concise visual summary of the direction and magnitude of the pairwise relationships and helps identify potentially important patterns of association and possible multicollinearity before the truncated-bootstrap and fixed-effects results are considered.



Heatmap of the correlation structure among health-system inputs, UHC coverage, health outcomes, financing, income, and institutional quality.

Figure 5: Pearson correlation heatmap of key health-system, UHC, financing, socioeconomic, institutional, and health-outcome variables in Sub-Saharan Africa, 2000–2023.

Note: Cell values are pairwise Pearson correlation coefficients. Warmer shades indicate positive associations, while cooler shades indicate negative associations. Correlations are

calculated using available country-year observations; pairwise sample sizes may vary because the study retains observed missingness rather than imputing unavailable values. **Table 7 reports the truncated-normal bootstrap estimates of the structural and financing correlates of bias-corrected inefficiency.**

Table 7: Truncated-normal bootstrap regression of bias-corrected inefficiency

Variable	Model 1: OOP	Model 2: Govt budget share	Model 3: External funding	Model 4: Govt share of CHE
ln GDP per capita	0.009 [-0.144, 0.151]	0.044 [-0.101, 0.180]	-0.037 [-0.248, 0.136]	0.059 [-0.113, 0.221]
Institutional quality	-0.080 [-0.231, 0.047]	-0.080 [-0.218, 0.041]	-0.142 [-0.343, -0.018]	-0.094 [-0.285, 0.058]
Urban share	0.082 [-0.046, 0.234]	0.088 [-0.040, 0.245]	0.082 [-0.049, 0.240]	0.104 [-0.029, 0.255]
Out-of-pocket share	0.107 [-0.004, 0.234]			
Government health spending (% govt exp.)		-0.152 [-0.330, -0.008]		
External health funding			-0.086 [-0.287, 0.063]	
Government share of CHE				-0.115 [-0.281, 0.041]
Observations	44	44	44	44
Bootstrap replications	500	500	500	500

Note: Dependent variable is 1/bias-corrected TE. Covariates are standardized; positive coefficients indicate higher inefficiency. Bootstrap confidence intervals are based on 500 replications.

Institutional-Quality PCA Robustness

The baseline governance index is not influenced by arbitrary averaging, as evidenced by the fact that the first principal component of the six WGI dimensions explains 84.4% of their common variation and correlates 0.999 with the simple institutional-quality mean. Although this specification-sensitive result is interpreted cautiously, replacing the mean with PCA governance preserves the substantive pattern and makes the OOP coefficient positive and statistically distinguishable from zero (0.109; 95% CI 0.002 to 0.248). Table 8 reports the first principal-component loadings for the six WGI governance dimensions.

For reproducibility, the six WGI dimensions are summarized using the first principal component after standardization. The signs of component loadings must be aligned before interpreting higher PC1 scores as better governance. Original PCA preprocessing, missing-data treatment and software version should be recorded in the accompanying reproducibility files.

Table 8: First principal-component loadings for the six WGI governance dimensions

WGI dimension	PC1 loading
GE	0.415
CC	0.411
RQ	0.415
RL	0.435
PV	0.375
VA	0.396

Table 9 presents the PCA-governance robustness results for the OOP truncated-bootstrap specification.

Table 9: PCA-governance robustness of the OOP truncated-bootstrap model

Variable	Coefficient	95% bootstrap interval
ln GDP per capita	0.009	[-0.131, 0.132]
PCA institutional quality	-0.078	[-0.216, 0.071]
Urban share	0.080	[-0.045, 0.216]
Out-of-pocket share	0.109	[0.002, 0.248]

Note: The dependent variable is reciprocal bias-corrected efficiency. Positive coefficients indicate greater inefficiency.

Annual Efficiency Panel And Two-Way Fixed Effects

With at least 15 complete countries, the annual VRS DEA can be estimated in 2004 and 2008–2022, producing 403 efficiency observations. For contemporaneous FE models, 233 country-years from 43 countries make up the common covariate sample. None of the financing, institutional-quality, education, or GDP coefficients are statistically significant at conventional levels once country and year fixed effects are taken into account. This shows that the strongest signals function through enduring between-country structures rather than short-run annual variation, in contrast to the cross-sectional financing gradients. Table 10 reports the two-way fixed-effects estimates for the annual VRS DEA efficiency panel.

Table 10: Two-way fixed-effects models of annual VRS DEA efficiency

Variable	Model 1: OOP	Model 2: Govt budget share	Model 3: External funding	Model 4: Govt share of CHE
Secondary enrollment	0.0001 (0.0011)	0.0001 (0.0009)	0.0000 (0.0010)	0.0002 (0.0010)
ln GDP per capita	-0.0066	-0.0049	-0.0118	-0.0016

Variable	Model 1: OOP	Model 2: Govt budget share	Model 3: External funding	Model 4: Govt share of CHE
	(0.0758)	(0.0753)	(0.0817)	(0.0760)
Institutional quality	0.0148 (0.0767)	0.0049 (0.0746)	0.0204 (0.0784)	0.0151 (0.0789)
Out-of-pocket share	0.0008 (0.0014)			
Government health spending (% govt exp.)		-0.0061 (0.0042)		
External health funding			0.0001 (0.0009)	
Government share of CHE				-0.0011 (0.0010)
Country FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	233	233	233	233
Countries	43	43	43	43
Adjusted R-squared	0.673	0.680	0.671	0.674

Note: Country-clustered standard errors are reported. Financing variables are entered one at a time. These models identify within-country associations.

Figure 6 plots mean annual VRS DEA efficiency for the years in which at least 15 countries have complete baseline data.

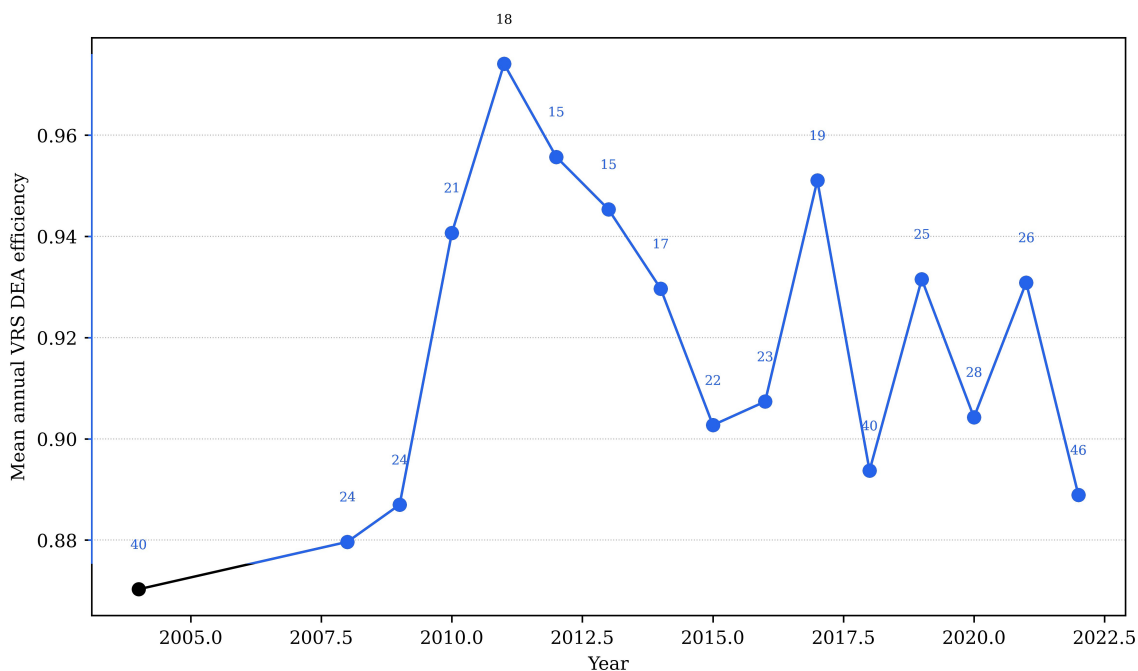


Figure 6. Mean annual VRS DEA efficiency for years with at least 15 complete countries.
Source: Authors' calculations. The composition of countries varies by year and the trend is therefore descriptive.

One-Year-Lagged Fixed-Effects Robustness

The common sample is increased to 390 observations from 47 countries by delaying the structural and financing variables by one year. The majority of lagged coefficients continue to be statistically insignificant. At the 10% level, government health spending as a percentage of the government budget is marginal (coefficient -0.0054; $p=0.069$), but its sign in the TE equation is the opposite of the cross-sectional truncated-bootstrap association with inefficiency. Therefore, rather than being a contradiction of the structural cross-section, the result is interpreted as evidence of weak and specification-dependent short-run dynamics. Table 11 presents the one-year-lagged two-way fixed-effects robustness results.

Table 11: One-year-lagged two-way fixed-effects robustness

Model	Lagged financing variable	Coefficient	Clustered SE	p-value	N	Countries
Model 1	L1 Out-of-pocket share	-0.0003	0.0010	0.778	390	47
Model 2	L1 Government health spending (% govt exp.)	-0.0054	0.0030	0.069*	390	47
Model 3	L1 External health funding	0.0007	0.0007	0.348	390	47
Model 4	L1 Government share of CHE	-0.0010	0.0008	0.206	390	47

Note: Dependent variable is annual VRS technical efficiency. Country and year fixed effects are included; standard errors are clustered by country.

Matched 2018-2022 Comparison

The mean efficiency for 40 countries in the complete baseline input-output bundle is observed in both 2018 and 2022. The mean value goes from 0.894 to 0.885; 14 countries improve, 16 worsen, and 10 countries remain the same to this displayed precision. The paired Wilcoxon test does not reject equal distributions ($p=0.318$), while the Spearman rank correlation is high ($\rho=0.752$; $p<0.001$). Thus, relative positions are persistent despite individual country movements. Figure 7 compares matched-country VRS DEA efficiency in 2018 and 2022.

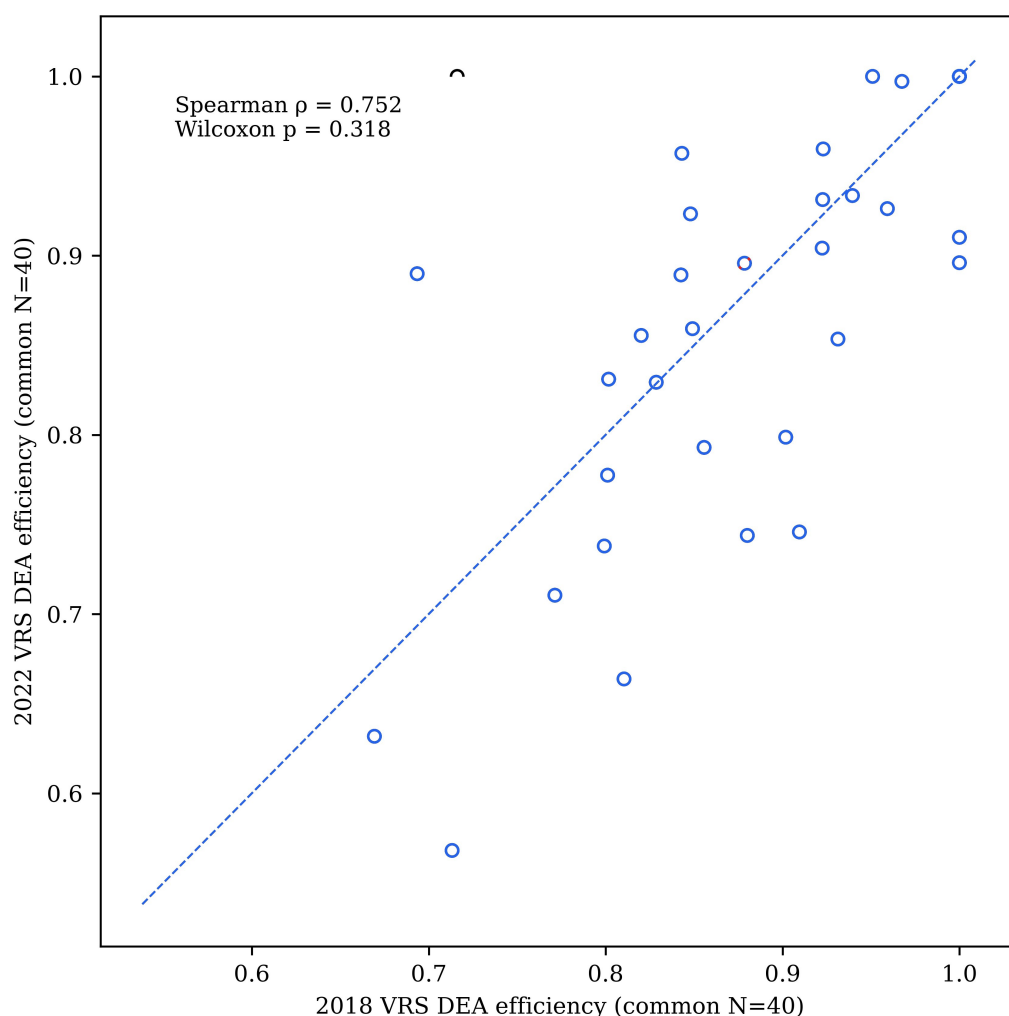


Figure 7: Matched-country VRS DEA efficiency in 2018 and 2022 (N=40).
Source: Authors' calculations.

Specification Sensitivity And Frontier Stability

The mean efficiency is 0.851 with expenditure plus physicians, 0.871 with expenditure plus nurses/midwives, and 0.819 with expenditure alone, indicating that the baseline frontier is substantively robust to reasonable simplifications of the input set. Because the dimensionality of the frontier reduces discriminatory power, adding multiple outputs significantly increases efficiency: the four-UHC-sub-index specification reaches 0.980 and places 30 of 46 countries on the frontier. Although the hospital-bed model reaches 0.993, it is not very informative because it only includes eight countries and seven frontier units. The workforce-reduced models have the highest baseline rank correlations, while the hospital-bed and CRS technologies have significantly lower correlations. With a median rank correlation of 1.000, a minimum of 0.969, and a maximum absolute efficiency change of 0.114, leave-one-country-out jackknife re-estimation is extremely stable. Table 12 summarizes the sensitivity of the DEA results to alternative input sets, output definitions, and frontier technologies.

The numerical differences across alternative input and multi-output DEA technologies are descriptive unless a formal paired test is explicitly reported. In particular, adding outputs can increase apparent DEA efficiency because additional dimensions make it easier for units to

appear efficient, while reducing the method’s ability to distinguish peers. These changes are not described as statistically significant without a corresponding test.

Table 12: DEA sensitivity to alternative technologies and output definitions

Specification	N	Mean	SD	Median	Efficient DMUs	Rank ρ vs baseline
Baseline: CHE + physicians + nurses/midwives -> UHC	46	0.889	0.112	0.907	14	1.000
CHE + physicians -> UHC	46	0.851	0.136	0.856	11	0.854
CHE + nurses/midwives -> UHC	46	0.871	0.116	0.887	12	0.913
CHE only -> UHC	48	0.819	0.132	0.816	6	0.711
3 inputs -> four UHC sub-indices	46	0.980	0.036	1.000	30	0.633
3 inputs -> UHC + life expectancy	46	0.956	0.047	0.960	18	0.684
4 inputs incl. hospital beds -> UHC	8	0.993	0.020	1.000	7	0.088
CRS baseline (same 46 countries)	46	0.513	0.280	0.457	4	0.231

Note: Rank ρ is the Spearman correlation with the baseline 2022 VRS ranking. Models with additional dimensions mechanically weaken DEA discrimination.

Figure 8 compares mean efficiency across the alternative 2022 DEA specifications.

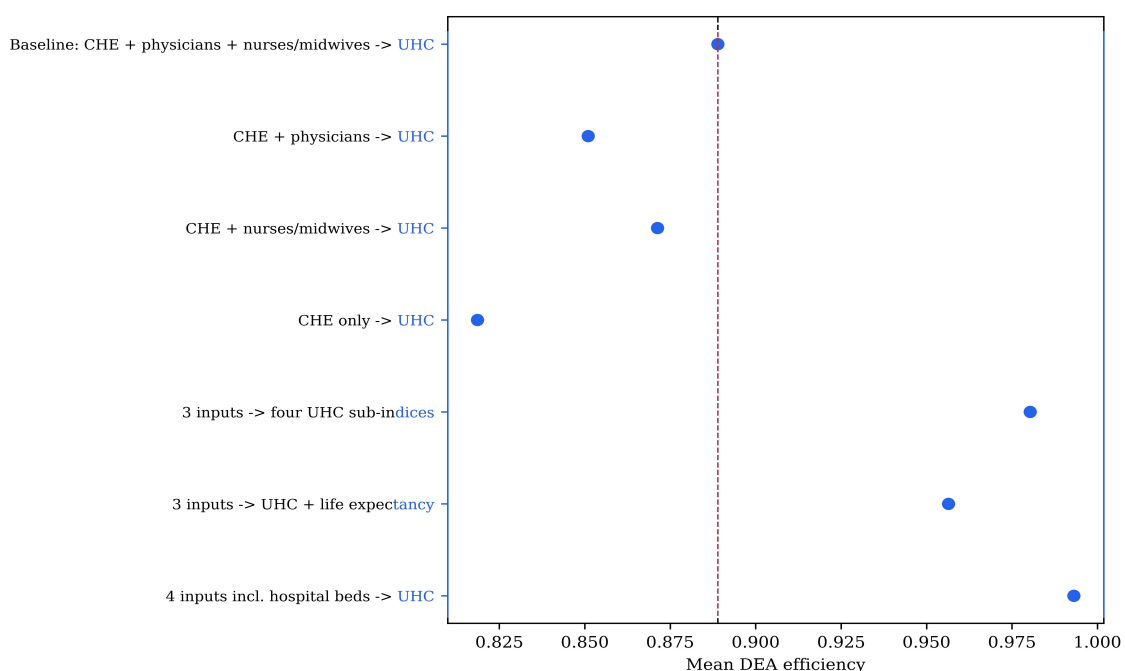


Figure 8: Mean efficiency across alternative 2022 DEA specifications.

Source: Authors' calculations.

Stochastic-Frontier Cross-Check

The 46-country cross-section used to estimate the half-normal Cobb-Douglas stochastic frontier yields mean technical efficiency of 0.858 and median efficiency of 0.883. One-sided inefficiency variation is significantly greater than symmetric noise under the parametric model, according to the estimated noise-to-inefficiency structure, which implies $\lambda=2.817$ with $\sigma_u=0.205$ and $\sigma_v=0.073$. The core efficiency ordering is supported by the DEA and SFA rankings' positive and significant concordance (Spearman $\rho=0.651$; $p<0.001$), which also confirms that the two approaches are not the same. Table 13 reports the half-normal Cobb-Douglas stochastic-frontier estimates used as a cross-method robustness check.

Table 13: Half-normal Cobb-Douglas stochastic-frontier cross-check

Statistic	Estimate
Constant	3.4294
$\ln(\text{CHE per capita, PPP})$	0.1087
$\ln(\text{Physicians per 1,000})$	-0.0020
$\ln(\text{Nurses/midwives per 1,000})$	0.1145
σ	0.2170
$\lambda = \sigma_u/\sigma_v$	2.8168
σ_u	0.2045
σ_v	0.0726
Mean SFA TE	0.8579
Median SFA TE	0.8834
DEA-SFA Spearman ρ	0.6508 ($p<0.001$)

Note: SFA assumes a Cobb-Douglas production function, normally distributed statistical noise, and half-normal non-negative inefficiency.

Figure 9 compares the bias-corrected DEA efficiency scores with the stochastic-frontier efficiency scores for 2022.

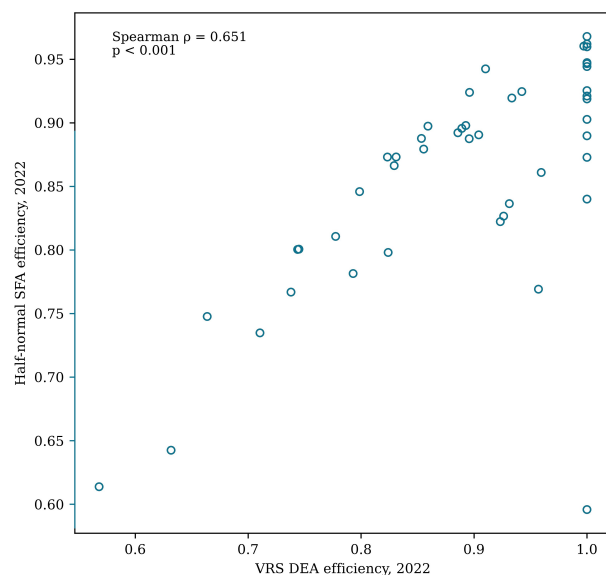


Figure 9: Bias-corrected DEA and stochastic-frontier efficiency scores, 2022.

Source: Authors' calculations.

Pre-Covid, Covid, and Post-COVID Comparison

Without a statistically significant change in efficiency, the period comparison reveals significant increases in coverage and resource mobilization. Health spending per capita increases from US\$203.94 to US\$292.58 and then US\$328.83 ($p<0.001$), while mean UHC service coverage increases from 38.61 prior to COVID to 48.13 in 2020–2021 and 49.67 in 2022–2023 ($p<0.001$). The OOP share decreases from 40.78% to roughly 35% ($p=0.004$). In contrast, DEA efficiency shifts from 0.913 pre-COVID to 0.917 during COVID and 0.889 in the 2022 post-COVID benchmark ($p=0.202$), while institutional quality and government health spending as a percentage of the budget do not significantly change over time. As a result, the area saw increased coverage both during and after the pandemic, but there was no discernible change in technical efficiency. Table 14 summarizes the pre-COVID, COVID, and post-COVID comparisons of coverage, financing, health outcomes, institutions, and efficiency. These period patterns should also be interpreted against the broader disruption to essential health services and disease-control systems documented across Africa during the COVID-19 pandemic (Inzaule et al., 2021).

Table 14: Pre-COVID, COVID, and post-COVID comparison of coverage, financing, outcomes, institutions, and efficiency

Indicator	Pre-COVID mean (N)	COVID mean (N)	Post-COVID mean (N)	Kruskal-Wallis p
UHC service coverage	38.614 (960)	48.125 (96)	49.667 (96)	<0.001***
Health expenditure per capita, PPP	203.940 (920)	292.579 (96)	328.830 (96)	<0.001***
Out-of-pocket share	40.780 (920)	35.538 (96)	35.156 (96)	0.004***
Government health spending (% govt exp.)	6.729 (920)	7.295 (96)	6.866 (96)	0.251
Life expectancy	58.128 (960)	62.579 (96)	63.817 (96)	<0.001***
Under-five mortality	93.142 (960)	62.950 (96)	62.925 (96)	<0.001***
Institutional quality	-0.688 (901)	-0.698 (96)	-0.708 (96)	0.908
Annual DEA efficiency	0.913 (303)	0.917 (54)	0.889 (46)	0.202

Note: For DEA efficiency, the post-COVID value is the 2022 benchmark because fewer than 15 countries have the full baseline input bundle in 2023.

DISCUSSION

Principal Findings

The magnitude and interpretation of the study's main finding are shown in the analysis above. First, traditional VRS DEA overstates efficiency; the mean country-specific radial output-expansion opportunity is 21.8%, and the mean drops from 0.889 to 0.834 after bootstrap correction. Second, operating scale is important. The significantly lower CRS mean and average scale efficiency demonstrate that the regional efficiency issue involves both scale mismatch and managerial resource conversion. Third, compared to national income, financing architecture is a more powerful discriminator. Private/OOP-dominated systems perform worse than external/government-dominated systems, but there is little difference between income groups. Fourth, institutional quality becomes particularly significant in the external-financing

model, and government health spending as a percentage of the government budget is the most obvious second-stage correlate of reduced inefficiency. Fifth, matched-year stability, strong rank persistence, and null within-country FE results suggest that the main differences are not short-term annual responses but rather structural and long-lasting.

Positioning Relative to The African and Global Efficiency Literature

The bias-corrected mean of 0.834 changes the implied efficiency dividend and is more in line with previous African evidence than the uncorrected benchmark of 0.889. While the current analysis covers 46 countries in the 2022 frontier and embeds that benchmark in a 48-country longitudinal data architecture, Arhin et al. (2023) reported a bias-corrected mean of 0.810 for 30 SSA countries around a 2015 median reference year. Although the output definition, time reference, input set, and data treatment vary, the direction is consistent—deterministic frontier scores are optimistic and significant efficiency gains are still possible—so numerical comparisons must be done with caution. The current contribution is more than just a more recent rank ordering; it incorporates cross-method robustness, scale decomposition, structural versus within-country separation, explicit selection diagnostics, and wider geographic coverage.

Bias Correction Changes The Policy Arithmetic

It is crucial to distinguish between the radial output-expansion potential and the simple efficiency shortfall. Although output-oriented DEA is parameterized by the reciprocal expansion factor, a traditional mean TE of 0.889 indicates an 11.1% shortfall when expressed as 1-TE. The average country-specific radial expansion opportunity is 21.8% following finite-sample correction. Therefore, planning based only on deterministic 1-TE would underestimate the amount of extra service coverage that could be produced if nations adopted the estimated best-practice technology. However, there is no assurance that all nations can attain sufficient UHC without more funding; rather, this is a relative technical-efficiency opportunity. Low-resource nations may have low absolute coverage even though they are near a VRS frontier.

Financing Architecture And Domestic Budget Allocation

The most obvious stratification that affects policy is the financing-mode gradient. The lowest mean bias-corrected efficiency is found in private/OOP-dominated systems, which is in line with evidence that point-of-service payments disrupt risk pooling, put households at risk of financial hardship, and deter essential utilization (Saksena et al., 2014; Ifeagwu et al., 2021; Rahman et al., 2022). With OOP spending accounting for an average of 39.84% of the panel's current health expenditures, the financing mechanism itself presents a plausible obstacle to turning available resources into coverage. A one-standard-deviation increase in government health spending as a percentage of the government budget is linked to lower inefficiency, and its bootstrap interval excludes zero, according to the truncated-bootstrap result, which supports the domestic-financing interpretation. This result links the efficiency agenda to the longstanding Abuja commitment, but it should be read as an association rather than a causal budget multiplier. This interpretation is consistent with broader evidence that progress toward UHC depends on future financing trajectories, sustained domestic resource mobilisation, and government health-spending commitments, while the relationship between health expenditure and macroeconomic performance remains important in African settings (Dieleman et al., 2018; Elovainio & Evans, 2017; Megbowon & Zerihun, 2025; Piabuo & Tieguhong, 2017).

External Funding and Governance Conditionality

Although externally financed systems have the highest mean efficiency, the truncated model's external-funding share is not statistically accurate. This distinction is crucial because, rather than being a mechanical advantage of more donor funding, the observed premium may reflect the institutional and programmatic environments in which external resources are deployed. The conversion of external resources into coverage is conditioned by absorptive capacity, procurement discipline, accountability, and implementation quality, according to a statistical correlation between institutional quality and reduced inefficiency in the external-funding specification. Therefore, the policy implication is sequencing rather than permanent dependence: while effective coverage expansion can be supported by external resources, durable UHC necessitates domestic pooling and institutional capacity that can maintain those gains (Agyepong et al., 2017; Karamagi et al., 2023).

Structural Persistence And The Meaning Of The Fixed-Effects Nulls

The cross-sectional structural evidence is not refuted by the lack of significant contemporaneous FE coefficients. The dimensions along which the greatest cross-sectional differences occur—financing architecture, institutional capacity, geography, and health-system organization—are eliminated by country fixed effects. This interpretation is supported by the matched 2018-2022 exercise, which shows high rank persistence but little change in mean efficiency. With only a marginal government-budget coefficient whose sign in the TE equation differs from the cross-sectional inefficiency model, the one-year-lagged models likewise continue to be weak. Therefore, the methodical conclusion is that variations in financing and governance are essentially long-term structural correlates rather than quickly shifting causal levers that can be seen in annual data.

Robustness And What The Alternative Frontiers Reveal

The robustness architecture makes it clear which results are technology-dependent and which are stable. While adding multiple outputs inflates efficiency and weakens discrimination—an expected dimensionality effect rather than proof that most systems are actually efficient—workforce-reduced DEA specifications maintain rankings fairly well. Because only eight countries are complete and seven become efficient, the hospital-bed model is particularly uninformative. The leave-one-country-out jackknife, on the other hand, demonstrates that the baseline frontier is not determined by a single observation. PCA governance amplifies the negative OOP signal and replicates the baseline institutional-quality construct. Lastly, although the two frontier technologies are not interchangeable, the stochastic frontier produces a mean TE of 0.858 and significant rank concordance with DEA, suggesting that the broad ordering survives a model that explicitly accommodates statistical noise (Di Giorgio et al., 2016).

Policy Implications

The findings have some important implications for policy in order to improve the health-system efficiency and achieve UHC in Sub-Saharan Africa. First, it was found that policymakers should focus on the bias-corrected radial output-expansion potential of 21.8% as opposed to the deterministic 11.1% efficiency shortfall for setting realistic efficiency-improvement targets. The former figure offers a more nuanced view of the additional service coverage that can be achieved with the same set of inputs. Second, the reduction of out-of-pocket (OOP) payments should remain a priority for health-financing reform in the region. Indeed, compared to the systems dominated by private/OOP financing, those featuring prepayment, risk pooling, and other forms of non-point-of-service financing instruments were associated with lower

inefficiency. Third, there is a need to focus on the quality and efficiency of the domestic health-budget allocation. Specifically, the government's share of the public budget proved to be the most statistically reliable second-stage correlate of reduced inefficiency. Nevertheless, it should be noted that improving the latter requires more than just increasing the former. Fourth, it was found that scale inefficiency, along with pure technical inefficiency, should be a key policy focus. The fact that health systems differed in terms of scale efficiency suggests that the organization of services, referral networks, decentralization policies, facility distribution patterns, and the scale/resource clustering principles in a given jurisdiction also matter. Fifth, external health financing should be leveraged strategically because, despite the relative efficiency of externally financed systems, donor dependency can undermine sustainability. In other words, external resources should be used to build domestic financing, governance, and pooled funding/response mechanisms that will be able to withstand the aid transition. Finally, there is a need to prioritize health-data infrastructure investments. Indeed, the analysis found that data on financial hardship coverage was available for only 13.5% of the panel, whereas workforce and hospital capacity data, remain, for now, incomplete. As such, there is a need to improve household-expenditure surveys, workforce databases, and hospital capacity/accountability reporting for UHC-related analytical, planning, monitoring, evaluation, and accountability needs.

Limitations and Interpretation Boundaries

Several limitations are worth noting despite the methodological advances outlined above. First, the analysis is based on a 2022 cross-section and therefore evaluates performance relative to 2022 technology and not causal welfare gains. Second, the work force is missing at random in the long panel, and so annual estimates should not be extrapolated casually to all countries and years within the study's scope of the SSA region. Bootstrap DEA partly addresses finite sample concerns but does not make DEA estimates robust to measurement error, omitted factors, or different definitions of a country's frontier output. Third, the second-stage cross-sectional truncated regression models are based on 44 countries and therefore report associations, not causal effects, of financial hardship on cross-country performance variation. Fourth, financial hardship information on hardship is too limited to estimate a two-dimensional UHC frontier for the entire region; instead, the analysis focuses on coverage as the primary outcome. Fifth, the SFA estimation requires assuming a Cobb-Douglas technology and half-normal inefficiency distribution. Finally, a Malmquist productivity index was not estimated because current international PPPs for health spending are not a reliable constant-price measure of expenditure for analyzing productivity growth between waves.

Interpretation also depends on the country classification and complete-case screens. The 2022 benchmark is relatively comprehensive, but missing workforce data select the annual frontier toward better-recorded health systems. The annual fixed-effects estimation set and the lagged specification differ in size, and their coefficients therefore cannot be compared as if calculated on identical country-years. Reproducibility further requires an explicit retained-year list, country exclusion list and software environment, none of which can be reconstructed securely from the text alone.

CONCLUSION

The UHC gap in Sub-Saharan Africa is partially caused by an efficiency issue, but an uncorrected deterministic frontier underestimates the severity of the issue. After bootstrap correction, the mean VRS efficiency across 46 countries in 2022 decreases from 0.889 to 0.834,

and the mean country-specific radial output-expansion potential is 21.8%. Additionally, the data demonstrates that efficiency is not just about revenue. Stronger systems are distinguished from weaker ones by financing architecture, institutional capacity, and operating scale: externally and government-financed systems perform better than private/OOP-dominated systems; higher public budget commitment to health is linked to lower inefficiency; and scale efficiency is significantly below unity. Simultaneously, the DEA-SFA concordance, matched-year stability, jackknife results, and null within-country FE estimates advise against straightforward causal interpretations. Therefore, the practical message is disciplined but consequential: SSA countries can increase service coverage by improving the conversion of current resources; however, sustained gains necessitate pooled domestic financing, lower point-of-service payment, capable institutions, appropriate system scale, and significantly improved workforce and health-finance data.

Ethics Approval and Consent to Participate

Not applicable. The study relied exclusively on publicly available, aggregated country-level secondary data and did not involve human participants, human tissue, animals, or personally identifiable information. Ethical approval and informed consent were therefore not required.

Consent for Publication

Not applicable. The manuscript does not contain any individual person's identifiable data, images, or other material requiring consent for publication.

Availability of Data and Materials

The data supporting the findings of this study are publicly available in the Figshare repository at <https://doi.org/10.6084/m9.figshare.33735043>. The repository provides the study data and associated supporting materials for verification and reproducibility of the analyses reported in the manuscript.

Competing Interests

The authors declare that they have no competing financial or non-financial interests that could have influenced the research, interpretation of the findings, authorship, or publication of this manuscript.

Funding

This research received no external funding. The authors conducted the study without financial support from any public, commercial, or not-for-profit funding agency.

Authors' Contributions

Author 1 contributed to the conceptualization, methodology, investigation, literature review, and writing of the original draft and participated in reviewing and editing the manuscript. Author 2, 4 contributed to the conceptualization and methodology, conducted the formal analysis, software implementation, data curation, validation, visualization, and interpretation of the results, and contributed to the original drafting, review and editing, and supervision of the study. Author 3, 5 contributed to the literature review, methodology, validation, interpretation of the findings, critical revision of the manuscript, and project administration. All authors read and approved the final manuscript and agree to be accountable for all aspects of the work.

Acknowledgments

The authors acknowledge the World Health Organization and the World Bank for making the health-system, universal health coverage, health-financing, development, and governance data used in this study publicly available.

REFERENCES

- Agyepong, I. A., Sewankambo, N. K., Binagwaho, A., Coll-Seck, A. M., Corrah, T., Ezeh, A., Fekadu, A., Kilonzo, N., Lamptey, P., Masiye, F., Mayosi, B. M., Mboup, S., Muyembe, J., Pate, M. A., Sidibe, M., Simons, B., Tlou, S., Gheorghe, A., Legido-Quigley, H., ... Piot, P. (2017). The path to longer and healthier lives for all Africans by 2030: The Lancet Commission on the future of health in sub-Saharan Africa. *The Lancet*, 390(10114), 2803–2859. [https://doi.org/10.1016/s0140-6736\(17\)31509-x](https://doi.org/10.1016/s0140-6736(17)31509-x)
- Aigner, D., Lovell, C. A. K., & Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, 6(1), 21–37. [https://doi.org/10.1016/0304-4076\(77\)90052-5](https://doi.org/10.1016/0304-4076(77)90052-5)
- Ambapour, S. (2015). Assessing the Relative Efficiency of Health Systems in Sub-Saharan Africa Using Data Envelopment Analysis. *American Journal of Operations Research*, 5(1), 30–37. <https://doi.org/10.4236/ajor.2015.51003>
- Amu, H., Dowou, R. K., Saah, F. I., Efunwole, J. A., Bain, L. E., & Tarkang, E. E. (2022). COVID-19 and Health Systems Functioning in Sub-Saharan Africa Using the “WHO Building Blocks”: The Challenges and Responses. *Frontiers in Public Health*, 10, 856397. <https://doi.org/10.3389/fpubh.2022.856397>
- Arhin, K., Oteng-Abayie, E. F., & Novignon, J. (2023). Assessing the efficiency of health systems in achieving the universal health coverage goal: evidence from Sub-Saharan Africa. *Health Economics Review*, 13(1), 25. <https://doi.org/10.1186/s13561-023-00433-y>
- Badunenko, O., & Tauchmann, H. (2019). Simar and Wilson two-stage efficiency analysis for Stata. *The Stata Journal*, 19(4), 950–988. <https://doi.org/10.1177/1536867X19893640>
- Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, 30(9), 1078–1092. <https://doi.org/10.1287/mnsc.30.9.1078>
- World Health Organization & World Bank. (2023). Tracking universal health coverage: 2023 global monitoring report. World Health Organization. <https://www.who.int/publications/i/item/9789240080379>
- Bolongaita, S., Lee, Y., Johansson, K. A., Haaland, Ø. A., Tolla, M. T., Lee, J., & Verguet, S. (2023). Financial hardship associated with catastrophic out-of-pocket spending tied to primary care services in low- and lower-middle-income countries: findings from a modeling study. *BMC Medicine*, 21(1), 356. <https://doi.org/10.1186/s12916-023-02957-w>
- Brieki, N., Alao, R., Wang, H., Erlangga, D., & Hanson, K. (2024). Improving the efficiency in spending for health: A systematic review of evidence. *LSHTM Research Online* (London School of Hygiene and Tropical Medicine), 2, 100008. <https://doi.org/10.1016/j.ssmhs.2024.100008>
- Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6), 429–444. [https://doi.org/10.1016/0377-2217\(78\)90138-8](https://doi.org/10.1016/0377-2217(78)90138-8)

- Dieleman, J. L., Sadat, N., Chang, A. Y., Fullman, N., Abbafati, C., Acharya, P., Adou, A. K., Kiadaliri, A., Alam, K., Alizadeh-Navaei, R., Alkerwi, A., Ammar, W., Antonio, C. A. T., Aremu, O., Asgedom, S. W., Atey, T. M., Ávila-Burgos, L., Ayer, R., Badali, H., ... Murray, C. J. L. (2018). Trends in future health financing and coverage: future health spending and universal health coverage in 188 countries, 2016–40. *The Lancet*, 391(10132), 1783–1798. [https://doi.org/10.1016/s0140-6736\(18\)30697-4](https://doi.org/10.1016/s0140-6736(18)30697-4)
- Dlouhý, M., & Havlík, P. (2024). Efficiency evaluation of 28 health systems by MCDA and DEA. *Health Economics Review*, 14(1), 59. <https://doi.org/10.1186/s13561-024-00538-y>
- Elovainio, R., & Evans, D. (2017). Raising more domestic money for health: prospects for low- and middle-income countries. *Health Economics Policy and Law*, 12(2), 139–157. <https://doi.org/10.1017/s1744133116000426>
- Şener, M. (2025). Can Access to Health Services and Universal Health Coverage Improve the Efficiency of Health Systems in Sub-Saharan African Countries? A Study Based on a Two-Stage Dynamic Data Envelopment Analysis (DEA) Model. *World Medical & Health Policy*, 17(4), 883–897. <https://doi.org/10.1002/wmh3.70047>
- Eze, P., Idemili, J. C., & Lawani, L. O. (2024). Evaluating health systems' efficiency towards universal health coverage: A data envelopment analysis. *INQUIRY The Journal of Health Care Organization Provision and Financing*, 61. <https://doi.org/10.1177/00469580241235759>
- Di Giorgio, L., Flaxman, A. D., Moses, M. W., Fullman, N., Hanlon, M., Conner, R. O., Wollum, A., & Murray, C. J. L. (2016). Efficiency of health care production in low-resource settings: A Monte-Carlo simulation to compare the performance of data envelopment analysis, stochastic distance functions, and an ensemble model. *PLOS ONE*, 11(1), e0147261. <https://doi.org/10.1371/journal.pone.0147261>
- Grépin, K. A., Irwin, B. R., & Trakinsky, B. S. (2020). On the Measurement of Financial Protection: An Assessment of the Usefulness of the Catastrophic Health Expenditure Indicator to Monitor Progress Towards Universal Health Coverage. *Health Systems & Reform*, 6(1). <https://doi.org/10.1080/23288604.2020.1744988>
- Guo, B., & Grépin, K. A. (2026). Measurement matters: implications of the newly revised Sustainable Development Goal 3.8.2 financial protection indicator for global monitoring. *The Lancet Global Health*, 14(4). [https://doi.org/10.1016/s2214-109x\(25\)00536-4](https://doi.org/10.1016/s2214-109x(25)00536-4)
- Hsu, J., Flores, G., Evans, D., Mills, A., & Hanson, K. (2018). Measuring financial protection against catastrophic health expenditures: methodological challenges for global monitoring. *International Journal for Equity in Health*, 17(1), 69. <https://doi.org/10.1186/s12939-018-0749-5>
- Ibrahim, M. D., Daneshvar, S., Hocaoglu, M., & Oluseye, O.-W. G. (2018). An Estimation of the Efficiency and Productivity of Healthcare Systems in Sub-Saharan Africa: Health-Centred Millennium Development Goal-Based Evidence. *Social Indicators Research*, 143(1), 371–389. <https://doi.org/10.1007/s11205-018-1969-1>
- Ifeagwu, S. C., Yang, J., Parkes-Ratanshi, R., & Brayne, C. (2021). Health financing for universal health coverage in Sub-Saharan Africa: a systematic review. *Global Health Research and Policy*, 6(1), 8. <https://doi.org/10.1186/s41256-021-00190-7>
- Inzaule, S., Ondoa, P., Loembé, M. M., Tebeje, Y. K., Ouma, A. E. O., & Nkengasong, J. N. (2021). COVID-19 and indirect health implications in Africa: Impact, mitigation measures, and lessons learned for improved disease control. *PLoS Medicine*, 18(6). <https://doi.org/10.1371/journal.pmed.1003666>
- Jordi, E., Pley, C., Jowett, M., Abou Jaoude, G. J., & Haghparast-Bidgoli, H. (2020). Assessing the efficiency of countries in making progress towards universal health coverage: A

- data envelopment analysis of 172 countries. *BMJ Global Health*, 5(10), e002992. <https://doi.org/10.1136/bmjgh-2020-002992>
- Karamagi, H., Njuguna, D., Kidane, S. N., Djossou, H., Kipruto, H., Seydi, A. B.-W., Nabyonga-Orem, J., Muhongerwa, D. K., Frimpong, K. A., & Nganda, B. (2023). Financing health system elements in Africa: A scoping review. *PLoS ONE*, 18(9). <https://doi.org/10.1371/journal.pone.0291371>
- Kirigia, J. M., Zere, E. Z., Greene, W. D., & Emrouznejad, A. (2007). Technical efficiency, efficiency change, technical progress and productivity growth in the national health systems of continental African countries. *Eastern Africa Social Science Research Review*, 23(2). <https://doi.org/10.1353/eas.2007.0008>
- Lal, A., Erondy, N., Heymann, D., Gitahi, G., & Yates, R. (2020). Fragmented health systems in COVID-19: rectifying the misalignment between global health security and universal health coverage. *The Lancet*, 397(10268), 61–67. [https://doi.org/10.1016/s0140-6736\(20\)32228-5](https://doi.org/10.1016/s0140-6736(20)32228-5)
- Mbau, R., Musiega, A., Nyawira, L., Tsofa, B., Mulwa, A., Molyneux, S., Maina, I., Jemutai, J., Normand, C., Hanson, K., & Barasa, E. (2022). Analysing the Efficiency of Health Systems: A Systematic Review of the Literature. *Applied Health Economics and Health Policy*, 21(2), 205–224. <https://doi.org/10.1007/s40258-022-00785-2>
- Megbowon, E. T., & Zerihun, M. F. (2025). Government Health Expenditure Gap Analysis: A Study on the Commitment to International Agreement and the Role of Economic and Institutional Factors. *Health Economics and Management Review*, 6(1), 95–109. <https://doi.org/10.61093/hem.2025.1-06>
- Nabyonga-Orem, J., Christmal, C., Addai, K. F., Mwinga, K., Aidam, K., Nachinab, G. T., Namuli, S., & Asamani, J. A. (2023). The state and significant drivers of health systems efficiency in Africa: A systematic review and meta-analysis. *Journal of Global Health*, 13, 4131. <https://doi.org/10.7189/jogh.13.04131>
- Piabuo, S. M., & Tieguhong, J. C. (2017). Health expenditure and economic growth - a review of the literature and an analysis between the economic community for central African states (CEMAC) and selected African countries. *Health Economics Review*, 7(1), 23. <https://doi.org/10.1186/s13561-017-0159-1>
- Rahman, T., Gasbarro, D., & Alam, K. (2022). Financial risk protection from out-of-pocket health spending in low- and middle-income countries: a scoping review of the literature. *Health Research Policy and Systems*, 20(1), 83. <https://doi.org/10.1186/s12961-022-00886-3>
- Saksena, P., Hsu, J., & Evans, D. (2014). Financial Risk Protection and Universal Health Coverage: Evidence and Measurement Challenges. *PLoS Medicine*, 11(9). <https://doi.org/10.1371/journal.pmed.1001701>
- Simar, L., & Wilson, P. W. (2007). Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of Econometrics*, 136(1), 31–64. <https://doi.org/10.1016/j.jeconom.2005.07.009>
- Simar, L., & Wilson, P. W. (1998). Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models. *Management Science*, 44(1), 49–61. <https://doi.org/10.1287/mnsc.44.1.49>
- Wagstaff, A., Eozenou, P., & Smits, M. (2020). Out-of-Pocket Expenditures on Health: A Global Stocktake. *The World Bank Research Observer*, 35(2), 123–157. <https://doi.org/10.1093/wbro/lkz009>

APPENDIX

List of Abbreviations

Abbreviation	Meaning
BCC	Banker-Charnes-Cooper
CC	Control of Corruption
CHE	Current Health Expenditure
CI	Confidence Interval
COVID-19	Coronavirus Disease 2019
CRS	Constant Returns to Scale
DEA	Data Envelopment Analysis
DMU / DMUs	Decision-Making Unit / Decision-Making Units
FE	Fixed Effects
GDP	Gross Domestic Product
GE	Government Effectiveness
L1	One-period lag (first lag)
ln	Natural logarithm
N	Number of observations / sample size
NCD	Non-Communicable Disease
OOP	Out-of-Pocket
OOPE	Out-of-Pocket Expenditure
PC1	First Principal Component
PCA	Principal Component Analysis
PPP	Purchasing Power Parity
PV	Political Stability and Absence of Violence/Terrorism
RMNCH	Reproductive, Maternal, Newborn and Child Health
RL	Rule of Law
RQ	Regulatory Quality
SD	Standard Deviation
SDG	Sustainable Development Goal
SE	Standard Error
SFA	Stochastic Frontier Analysis
SSA	Sub-Saharan Africa
TE	Technical Efficiency
UHC	Universal Health Coverage
US\$	United States dollar
VA	Voice and Accountability
VRS	Variable Returns to Scale
WDI	World Development Indicators
WGI	Worldwide Governance Indicators
WHO	World Health Organization