

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Please do not complete any field with "not applicable" or n/a. Refer to the help text for what text to use if an item is not relevant to your study. For final submission: please carefully check your responses for accuracy; you will not be able to make changes later.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- | n/a | Confirmed |
|-------------------------------------|--|
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A description of all covariates tested |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
<i>Give P values as exact values whenever suitable.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Data analysis

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender

Not relevant to this study; no human participants data were collected.

Reporting on race, ethnicity, or other socially relevant groupings

Not relevant to this study; no human participants data were collected.

Population characteristics

Not relevant to this study; units of analysis are water infrastructure sites and revenue records.

Recruitment

Not relevant to this study; no recruitment of participants.

Ethics oversight

Not relevant to this study; no human participants or identifiable personal data.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences

Behavioural & social sciences

Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

Data exclusions

Replication

Randomization

Blinding

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description

Research sample

Sampling strategy

Data collection

Timing

Data exclusions

Non-participation

Randomization

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Observational study characterizing and predicting revenue from rural drinking water infrastructure across multiple countries.
Research sample	Water infrastructure sites (handpumps and piped schemes) with verified revenue records across 11 African countries
Sampling strategy	Convenience sample based on availability of independently verified revenue records from participating service providers
Data collection	Quarterly revenue records compiled through service provider administrative systems and verified under results-based contracting frameworks.
Timing and spatial scale	Multi-year revenue time series at infrastructure-site resolution across 11 countries (site-level geospatial attributes used for modeling).
Data exclusions	Outlier handling and data cleaning rules are described in the Methods; no experimental exclusions.
Reproducibility	All data and code are publicly available (Dryad: 10.5061/dryad.7pvmcvf7r; Zenodo: 10.5281/zenodo.18431343).
Randomization	Not relevant to this study; observational analysis with no random assignment.
Blinding	Not relevant to this study; observational analysis with no participant outcomes.

Did the study involve field work? Yes No

Field work, collection and transport

Field conditions	Not relevant to this study
Location	Not relevant to this study
Access & import/export	Not relevant to this study
Disturbance	Not relevant to this study

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern
<input checked="" type="checkbox"/>	<input type="checkbox"/> Plants

Methods

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Antibodies

Antibodies used	Not relevant to this study
Validation	Not relevant to this study

Eukaryotic cell lines

Policy information about [cell lines and Sex and Gender in Research](#)

Cell line source(s)	Not relevant to this study
Authentication	Not relevant to this study
Mycoplasma contamination	Not relevant to this study
Commonly misidentified lines (See ICLAC register)	Not relevant to this study

Palaeontology and Archaeology

Specimen provenance	Not relevant to this study
Specimen deposition	Not relevant to this study
Dating methods	Not relevant to this study
<input type="checkbox"/> Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.	
Ethics oversight	Not relevant to this study

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Animals and other research organisms

Policy information about [studies involving animals; ARRIVE guidelines](#) recommended for reporting animal research, and [Sex and Gender in Research](#)

Laboratory animals	Not relevant to this study
Wild animals	Not relevant to this study
Reporting on sex	Not relevant to this study
Field-collected samples	Not relevant to this study
Ethics oversight	Not relevant to this study

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about [clinical studies](#)

All manuscripts should comply with the ICMJE [guidelines for publication of clinical research](#) and a completed [CONSORT checklist](#) must be included with all submissions.

Clinical trial registration	Not relevant to this study
Study protocol	Not relevant to this study
Data collection	Not relevant to this study
Outcomes	Not relevant to this study

Dual use research of concern

Policy information about [dual use research of concern](#)

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

- | No | Yes |
|-------------------------------------|---|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Public health |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> National security |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Crops and/or livestock |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Ecosystems |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Any other significant area |

Experiments of concern

Does the work involve any of these experiments of concern:

- | No | Yes |
|-------------------------------------|--|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Demonstrate how to render a vaccine ineffective |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Confer resistance to therapeutically useful antibiotics or antiviral agents |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Enhance the virulence of a pathogen or render a nonpathogen virulent |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Increase transmissibility of a pathogen |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Alter the host range of a pathogen |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Enable evasion of diagnostic/detection modalities |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Enable the weaponization of a biological agent or toxin |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Any other potentially harmful combination of experiments and agents |

Plants

- | | |
|-----------------------|----------------------------|
| Seed stocks | Not relevant to this study |
| Novel plant genotypes | Not relevant to this study |
| Authentication | Not relevant to this study |

ChIP-seq

Data deposition

- Confirm that both raw and final processed data have been deposited in a public database such as [GEO](#).
- Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

- | | |
|--|--|
| Data access links
<i>May remain private before publication.</i> | Zenodo at DOI: 10.5281/zenodo.18431343 |
| Files in database submission | See DOI |
| Genome browser session
(e.g. UCSC) | Not relevant to this study |

Methodology

- | | |
|-------------------------|----------------------------|
| Replicates | Not relevant to this study |
| Sequencing depth | Not relevant to this study |
| Antibodies | Not relevant to this study |
| Peak calling parameters | Not relevant to this study |
| Data quality | Not relevant to this study |

Flow Cytometry

Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

Not relevant to this study

Instrument

Not relevant to this study

Software

Not relevant to this study

Cell population abundance

Not relevant to this study

Gating strategy

Not relevant to this study

- Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type

Not relevant to this study

Design specifications

Not relevant to this study

Behavioral performance measures

Not relevant to this study

Imaging type(s)

Not relevant to this study

Field strength

Not relevant to this study

Sequence & imaging parameters

Not relevant to this study

Area of acquisition

Not relevant to this study

Diffusion MRI

Used

Not used

Preprocessing

Preprocessing software

Not relevant to this study

Normalization

Not relevant to this study

Normalization template

Not relevant to this study

Noise and artifact removal

Not relevant to this study

Volume censoring

Not relevant to this study

Statistical modeling & inference

Model type and settings

Not relevant to this study

Effect(s) tested

Not relevant to this study

Specify type of analysis: Whole brain ROI-based Both

Statistic type for inference

Not relevant to this study

(See [Eklund et al. 2016](#))

Correction

Not relevant to this study

Models & analysis

n/a | Involved in the study

- Functional and/or effective connectivity
- Graph analysis
- Multivariate modeling or predictive analysis

Functional and/or effective connectivity

Not relevant to this study

Graph analysis

Not relevant to this study

Multivariate modeling and predictive analysis

Predictive modeling was used to estimate annual revenue levels from site and context covariates. A two-stage approach (classification into revenue bands followed by regression within bands) was evaluated using standard metrics (e.g., ROC-AUC/F1 for classification and R^2 /MAE/RMSE for regression) with cross-validation and leave-one-country-out tests to assess generalization. Full implementation details are provided in the Methods, and code is available at Zenodo (10.5281/zenodo.18431343).

