

ONQ Lab Test T-002

Final execution report

Mass, Direction and Historical Polity Breakdown

Registered evidence category: Invalid / Uninterpretable.

The confirmatory execution ended at the mandatory feasibility gate. The eligible primary sample contains 543 state-decade landmarks and 53 distinct qualifying breakdown onsets, below the registered minima of 600 and 80. Sections 17, 23 Step 1 and 25 require this classification and stop. No model was fitted and no exploratory analysis was performed.

What this means

T-002 does not establish whether the registered Institutional Direction Proxy improves prediction beyond Material Mass Proxy. The frozen data and eligibility rules did not supply the minimum sample and event information required by this test. This result neither supports nor contradicts the narrow PrF Mass/Direction claim. It also provides no test of the broader PrF ontology. Thresholds, proxies and exclusions were not relaxed to obtain a result.

Primary feasibility and numerical results

Requirement	Observed	Registered minimum	Status
Eligible predictor-complete landmarks	543	600	Fail
Distinct analytical polities/lineages	98	60	Pass
Distinct qualifying breakdown onsets	53	80	Fail
Polities with qualifying onsets	46	30	Pass
Onsets in each grouped-CV test fold	Not assessed	10 per fold	Not reached
Usable temporal origins	Not assessed	3	Not reached
Distinct temporal-test onsets	Not assessed	30	Not reached

Events were counted uniquely across overlapping 50-year landmark windows. Positive landmark rows were not substituted for distinct events. The saved primary counts were independently reconciled to the constructed panel and annual onset audit after the interrupted run resumed.

Held-out log loss, relative improvement, high-Mass probability contrasts, coefficients, Brier scores, PR-AUC, AUROC, calibration statistics and bootstrap intervals are not available because the preregistered feasibility stop precedes model fitting. There are no missing model results being withheld.

Inclusion and exclusion flow

Primary 50-year processing stage	Removed	Remaining
Zero-ending candidate predictor landmarks	-	2,673
Predictor history, membership or code mapping ineligible	1,397	1,276
Active SF episode at landmark	19	1,257
Incomplete follow-up without a qualifying onset	714	543
Primary analytic sample	-	543

Predictor candidates are not the final analysis sample. The provisional pre-outcome count of 1,275 was superseded by 1,276 during pre-outcome code review: the Polity coverage start was removed as an additional ten-year requirement, while the registered CHANGE=99 history reset was retained. This occurred before SF access. The final locked candidate panel contains 1,276 eligible candidates.

The outcome-construction script also creates horizon-specific derived rows. No secondary predictive analysis or robustness verdict was computed. The raw flow JSON's 25-year included count is before its subsequent restriction to the primary sample; it is not a separate confirmatory sample claim.

Registration and authoritative specification

The authoritative specification is the original T-002 Preregistration v1.0 together with the approved September 23, 2026 temporal log-loss update.

Registration: <https://osf.io/gafmy/>

DOI: 10.17605/OSF.IO/GAFMY.

Original OSF date_registered: 2026-09-20T19:53:04.409969, preserved verbatim as returned without a timezone suffix.

Update reference: https://api.osf.io/v2/schema_responses/6ab42f9a9f394b37f0e97a12/

Update date_submitted: 2026-09-23T20:03:43.308980. Public approval action: 2026-09-23T20:03:47.420188Z. The update changes response field 344-77 only.

The update specifies equal-polity averaging of temporal losses within each usable origin, followed by equal-origin averaging, separately for baseline and model. Relative improvement is calculated from those pooled losses. This applies to IDP, O* and E*. The rule was verified and preserved, but temporal models were not reached.

Frozen Markdown SHA-256:

fa1303478361ff6fe2aaf465fd5af15ac756489e7204101b4d71209a4de8f982

Frozen local PDF SHA-256:

c119271f936dd6854fc4005337bd13c550e74b2497c6d029f03eebde47028848

The OSF-attached PDF has SHA-256

b0476e6b4dc433ad1bd99dcca145200c99c79b403ab71e1af40578e4d1b6fe11. Its bytes differ, but its extracted text matches the local frozen PDF across all 22 pages. The OSF Markdown hash matches the original local Markdown. The original freeze package was not modified.

Audit, pre-outcome gate and access boundary

The repeated data-free verdict audit passed against the original specification plus the approved update. It checked the numerical contradiction rules, exhaustive category precedence, qualitative-term definitions, temporal pooling, robustness roles and diagnostic veto. Source and runtime feasibility remained separate execution checks.

The pre-outcome gate recorded PASS at 2026-09-23T20:17:30.641099+00:00. The outcome-access boundary was explicitly crossed at 2026-09-23T20:17:30.941450+00:00. The archived log records source downloads and hashes, header/schema checks, predictor construction and entity-crosswalk construction before that boundary.

Predictor extraction used an explicit V-Dem column allowlist. No SF field entered MMP, IDP, O*, E*, MMP6 or the leave-one-component-out predictor calculations. Raw predictors were retained without full-sample standardisation. No training-set standardisation was needed because model fitting was not reached.

Entity-crosswalk audit

Matching uses COW code and calendar year, with nonunique matches excluded and recorded. Country names were not used to repair matches. CHANGE=99 starts a new analytical segment with the registered history requirement. The annual Polity schema does not provide explicit predecessor/successor code columns, so no cross-code continuity links were invented. Exits without a mechanically established link were censored under Sections 5.4 and 5.8.

The package preserves the code-year crosswalk, duplicate-mapping audits, candidate eligibility table, annual outcome audit and final analytic rows. These permit inspection of the implemented continuity and censoring procedure without retrospective case selection.

Required analysis status

Analysis	Execution status
Primary baseline and IDP models	Not run: mandatory feasibility stop
Five-fold grouped CV	Not run: mandatory feasibility stop
Expanding-window temporal validation	Not run: mandatory feasibility stop
O* and E* construct discrimination	Not run: mandatory feasibility stop
Full-refit polity-cluster bootstrap	0 of 2,000 attempts: mandatory feasibility stop
25-year and 100-year sensitivity models	Not run: mandatory feasibility stop
MMP6, level-only and change-only models	Not run: mandatory feasibility stop
Four leave-one-component-out models	Not run: mandatory feasibility stop
Pre-1900 exclusion and regional blocking models	Not run: mandatory feasibility stop
Exploratory analyses	None

Implementation and provenance exceptions

There were no intentional methodological changes. One post-boundary parsing correction is recorded in full, with both code hashes and the original script preserved. The initial outcome parser expected an explicit SF=0 value and treated blank cells as missing, leaving an empty result and raising an error before feasibility results were produced. The source file uses SF=1 or blank. The manual documents blank section-4 transition records when no change is recorded (page 15), and Section 4.12 defines the SF=1 failure flag. The corrected parser implements the presence/absence flag. No model estimates or performance comparisons informed this correction.

The earlier pre-outcome eligibility correction is described in the flow section. Neither correction changed a registered predictor, threshold, horizon, model or evidence rule. The original and corrected SF parser and the execution chronology are preserved; a claim of zero implementation corrections would be inaccurate.

Official Polity5 Excel and SPSS downloads returned HTTP 406. The analysis uses a preserved p5v2018.xls file from Zenodo record 15304940. Its published MD5 was verified and its SHA-256 recorded. The worksheet and required fields match the named version. Byte identity to the currently inaccessible publisher download could not be independently checked. Other dataset archives were obtained from the named publishers.

Final package verification found two nonmatching source archives at the resumed checkpoint. Fresh NMC and V-Dem publisher downloads matched the original pre-outcome SHA-256 hashes exactly and were restored. The cause of the checkpoint byte changes is unknown. Nonmatching copies and the recovery record are preserved; the locked predictor and analysis outputs were unchanged. No analysis was rerun or selected as part of this integrity repair.

Source provenance and SHA-256 hashes

COW NMC v7.0

File: NMCv7.zip.

Source: <https://correlatesofwar.org/wp-content/uploads/NMCv7.zip>

SHA-256: 07b7396cbf0a06a3a57d97dfb56e316c5720b6feeab06db4578bcb6447be51a9

COW State System Membership v2024 (intervals)

File: States2024.zip.

Source: <https://correlatesofwar.org/wp-content/uploads/States2024.zip>

SHA-256: c11689e8c0aa70ede855f6b605c4dd71c7b04122b71d5827dab5ba06b32e2015

COW State System Membership v2024 (annual)

File: System2024.zip.

Source: <https://correlatesofwar.org/wp-content/uploads/System2024.zip>

SHA-256: de7cae05d9908cff3606df185518cda380dbdc4c4f24739e72b0b13c1cb4a3d4

V-Dem v16

File: V-Dem-CY-FullOthers-v16_csv.zip.

Source: https://v-dem.net/media/datasets/V-Dem-CY-FullOthers-v16_csv.zip

SHA-256: 4fcd51c81923f11c23f01a2d43918712aefffc6f04947e4cfc415c9c2f345cf7

Polity5 v2018 (preserved copy)

File: p5v2018.xls.

Source: <https://zenodo.org/api/records/15304940/files/p5v2018.xls/content>

SHA-256: 051818bccdb6740493fe83da1d792e02d4d2a34eabb43f864687ea8710cd6927

Reproduction and preserved artifacts

The execution archive includes the original freeze package, OSF update and approval metadata, source archives and manuals, source hashes, predictor and crosswalk code, both versions of the outcome parser, candidate and analytic data, inclusion/exclusion flow, feasibility result, verdict record, analysis-status reports, runtime versions, implementation log and this report.

Reproduction instructions are in execution/README.md. Reproduce in a separate copy to preserve the original timestamped log. The delivered code reproduces the executed feasibility-stop branch; downstream model-fitting code is not presented as completed work. No model fitting is needed to reproduce the registered category.

The final verdict is fixed by the mandatory feasibility failure. No additional exploratory analysis was performed after freezing it.