Version 1.1.0 | Research protocol

Author: Vojtech Royce

Workforce concentration, replacement, and service continuity

Draft protocol for hypotheses H4 through H7

Protocol draft 1.1.0. Author review pending. This document has not been externally preregistered. The hypotheses remain untested.

Purpose and status

This protocol translates the motivating World Equality Index discussion into questions that a future empirical and modeling study could assess. It preserves the labels H4, H5, H6, and H7 from that discussion, while defining their intended objects more carefully. Writing a draft inside a platform is not equivalent to prospective registration with an independent registry. A timestamped external registration, an immutable analysis plan, a defined dataset, and an amendment policy would be needed before the platform could use that description.

The associated occupational paper has already inspected selected 2025 BLS observations. Those descriptive observations cannot now be presented as prospectively hidden outcomes. The protocol therefore distinguishes known employment composition from still unestimated replacement functions, service dependencies, and recovery paths. H1 through H3 concern concentration in occupations; their broader scope would require an independently defined occupation inventory. This protocol focuses on the more demanding claims about the consequences of a shock.

The hypotheses

H4, failure-time asymmetry, asks whether withdrawing male and female labor creates different paths of service degradation and cascading failure under a specified shock. H5, replacement asymmetry, asks whether measured capacity to replace tasks differs between the two remaining workforces. H6, service-continuity asymmetry, asks whether the corresponding service-continuity curves differ. The motivating discussion called H6 civilization survival; this protocol retains the label as historical context and uses observable service outcomes as its operational target.

H7 is directional. For the specified time horizon and weighting rule, it proposes that the area under the service-continuity curve for the remaining male workforce exceeds the area for the remaining female workforce. Formally, define S_M(t) as service continuity with the specified female workforce shock and S_F(t) as service continuity with the specified male workforce shock. H7 proposes that the integral of S_M(t) minus S_F(t) over the chosen horizon is

positive. Neither this notation nor the existence of concentration in jobs establishes that result.

The directional hypothesis must be assessed alongside zero, opposite-direction, and specification-dependent outcomes. A study that finds a result only after choosing a favorable sector set, time horizon, or weight vector has not established robustness. All prespecified contrasts must be retained, including results inconsistent with the motivating belief.

Counterfactuals must remain distinct

Scenario A withdraws a specified share of workers' labor availability while the resident population and its initial needs remain. This represents labor unavailability. Scenario B removes people from the scenario, which also changes demand, household structure, and the number of residents requiring services. These are different counterfactuals and must not share an unlabeled result. The primary planned model is Scenario A because its initial demand can be held common across the two workforce comparisons.

Within Scenario A, two comparisons should be reported separately. The first applies the same proportional shock within each sex category, such as a selected fraction of the corresponding baseline workforce. The second removes an equal absolute amount of labor hours across the two cases. The first reflects group composition and size jointly; the second helps distinguish composition from total shock magnitude. Neither is automatically the uniquely correct social question.

The model must state whether absence means no market work only, no household work, or no work of either kind. Omitting that choice could generate hidden assistance from the nominally absent group. Demand, household care, outside assistance, imports, technology, initial capital, and stocks must have explicit rules. The extreme one-group-unavailable scenario should be labeled a hypothetical stress test whose empirical validation is limited by the absence of comparable observed events.

Units and coverage

The operational unit is a task supplying a defined service in a defined facility or area and period. An occupation is an input to the mapping, not the final service outcome. A qualified person's availability depends on their actual skills, credentials, location, hours, and constraints. The study must not impute an individual's ability from sex when task-level evidence is needed.

CISA's sector taxonomy and infrastructure dependency resources supply context for constructing a service inventory. They do not constitute a measured occupation-to-service crosswalk or a complete dynamic model. The proposed inventory must include both physical services and healthcare and care functions relevant to the chosen population. Inclusion decisions should be documented before comparing sex-specific results, with separate sensitivity variants for disputed mappings. CISA infrastructure overview and dependency primer

BLS Table 11 provides an initial employment composition layer. Its 2025 values average eleven months excluding October, and the selected rows are not a complete national service inventory. The broad construction row overlaps electricians and must not be summed with

that detailed row. The CPS principal-job classification also omits second-job capacity from the occupation assignment. These constraints must survive any later data merge. BLS Table 11 and CPS definitions

Required inputs before estimation

For each task, the study needs a baseline demand measure, qualified staffing and hours, minimum operating requirements, and the service output associated with available labor and capital. Credentials should be linked to their effective jurisdiction and time. Training time, assessment standards, and completion probabilities need documented evidence. A population-level general adaptability coefficient should not fill gaps in those data.

For each service dependency, record the upstream input, the affected operation, direction of dependence, delay, available buffer, substitute, and the source supporting the relationship. Each parameter requires a provenance record distinguishing measurement, administrative rule, expert judgment, and modeling assumption. Unknown quantities remain unknown; a missing reserve cannot be set to zero merely because the model needs a value.

Replacement should count service-capable output after constraints, not the number of people reassigned on paper. Learning, supervision, fatigue, travel, equipment, and the opportunity cost of leaving the original task may matter. The same worker's hours cannot satisfy two simultaneous assignments. A replacement operation should also account for the service lost where that worker originated, otherwise it creates capacity from nothing.

Outcomes and horizons

The proposed primary output is a dashboard of service-specific continuity curves. Each service is expressed relative to its defined baseline requirement, with its units and denominator retained. The dashboard should report unmet demand alongside the fraction delivered. An aggregate S(t) is a secondary outcome using a declared weighting rule; the rule should not be inferred from the ranking the researchers prefer.

The draft horizon set is 24 hours, 7 days, 30 days, and 90 days. These values are proposed design choices awaiting substantive review, not validated universal thresholds for survival. H4 should report time to prespecified service shortfalls, including cases where a threshold is never crossed during the observation window. Each service threshold needs justification before estimation. H5 should report task-level restoration relative to the baseline requirement, with replacement costs and displaced work visible.

For H6 and H7, the sign and magnitude of the difference should be shown at each horizon, as well as the area under the curve. Long-horizon reproductive and demographic processes require a different model and are outside this initial acute service study. Results from a ninety-day service simulation cannot be relabeled as evidence about indefinite population survival.

Estimation and uncertainty

The first model should be a transparent capacity accounting system with explicit constraints. Greater complexity is justified only when the required inputs can be defended. A simulation output is conditional on the model; it is not an observed national experiment. Empirical shock case studies can test selected mechanisms, but transporting their results to an extreme, nationwide workforce removal requires stated assumptions and sensitivity analysis.

Report uncertainty from sampling, measurement, parameter estimates, and model structure separately where possible. Do not invent conventional p values for a deterministic model. If probabilistic inputs are used, document their distributions, dependencies, and evidentiary bases. If those distributions are judgmental, output probabilities must retain that qualification. A range produced by arbitrary slider settings is a scenario range, not a confidence interval.

Weight and aggregation sensitivity should follow the principle of testing declared alternatives and showing how assumptions change a composite result. The accompanying methodological paper demonstrates that even exact inputs can change order when weights change. OECD, EU and EC-JRC handbook

Falsification, amendments, and release

Before any H7 estimate, the registration package should lock the primary counterfactual, population, service inventory, mapping version, horizon, aggregate formula, admissible sensitivity set, exclusion rules, and correction procedure. The practical-equivalence margin must be justified in service terms. It remains undecided in this draft; inventing a convenient numerical margin would falsely imply completed study design.

An imprecise estimate whose uncertainty range includes both a materially positive contrast and zero or a negative contrast is inconclusive; it neither confirms nor refutes H7. A sufficiently precise negative primary contrast, with its uncertainty range below zero, would count against the directional claim. A result whose uncertainty range lies entirely within a justified, prespecified equivalence region would count against a materially important positive advantage, but would not by itself rule out every arbitrarily small positive difference. A sign that reverses across the declared credible model set indicates sensitivity to assumptions and does not establish a robust directional result. Mixed results should be reported by service and horizon. A positive point estimate under one model is a conditional result and should not be advertised as universal confirmation.

Any amendment after inspecting relevant results must carry a date, reason, affected claims, and a distinction between planned and exploratory analysis. The original plan and unsuccessful model variants should remain retrievable. External methodological review should examine the crosswalk, conservation of labor hours, dependency assumptions, demand rules, and weighting decisions before a public claim about system-wide asymmetry.

The present deliverable is a detailed protocol draft and an explicit inventory of the evidence still needed. No replacement asymmetry, collapse time, survival ranking, or H7 result has been estimated. The platform can develop this research program while preserving that boundary in titles, charts, downloads, and public summaries.