

The Temporal Correspondence Limitation of Evidence-Based Policy Making

A Preliminary Discussion Towards an Astrophysics-Inspired State-Aware Epistemology of Governance

Wanhong Huang

huangwanhong@serendip.ngo

Abstract

Evidence-based policy making assesses evidence for its reliability, for the credibility of its causal identification, and for its transportability to a target population. This paper identifies a further requirement that the established apparatus of assessment leaves unstated: the system state that a body of evidence represents must correspond to the system state obtaining at the moment of decision. Evidence is generated from a system at one time and intervention acts on that system at a later time, so a body of evidence may be reliable, well identified and appropriately transported, and still describe a condition the system has left. The requirement is called temporal correspondence, and it is distinguished from the classical problem of induction: granting the uniformity of nature leaves it in force, since a stable law applied to a superseded state yields a decision addressed to no existing condition.

The paper develops the requirement by comparison with astrophysics, an observational science whose object is accessible only through signals that departed it in the past, and which has accordingly built its epistemology on the assumption that observation is historical. Fifteen elements of that discipline's practice are examined in turn and tested for transfer to governance, in three families: the reconstruction of a past state from a present signal, the constraints imposed by the observing situation, and the self-discipline of the observing community. The pattern of what fails to transfer is treated as the object of the enquiry, on the ground that it characterises governance as an observational situation.

Six contributions are claimed. The paper states the requirement and establishes its independence from the problem of induction. It shows by inspection of the instruments, in place of inference from their consequences, that the dominant apparatus of evidence assessment contains no corresponding category, and it accounts for the omission. It classifies the ways in which the practices of a mature observational science fail to transfer, and it identifies the limit beyond which the comparison supplies nothing. It derives a set of requirements on the handling of evidence, each of which costs discipline in place of apparatus. And it allocates the burden of justification where the requirement fails and a course would be irreversible. The inspiration the comparison affords is accordingly a discipline of observation and report, and no technique for recovering the present condition of a governed system from lagged observation of it.

Keywords: evidence-based policy making; temporal correspondence; observation lag; state estimation; epistemology of governance; astrophysical observation.

Notices

Draft. This paper is circulated for comment as a working draft. Corrections and objections are welcome and may be sent to the author at huangwanhong@serendip.ngo.

Licence. © 2026 Wanhong Huang. This work is made available under the Creative Commons Attribution-NonCommercial 4.0 International Licence. The work may be copied, distributed and adapted for non-commercial purposes with attribution to the author. The licence terms are available at <https://creativecommons.org/licenses/by-nc/4.0/>.

Use of generative AI. Claude and ChatGPT were used in the preparation of this paper, for discussion of its argument, for surveying prior formulations, and for drafting assistance. The paper reports no simulation and no empirical result of its own; its arguments are discursive throughout, and the two tables summarise the approaches examined and the transfer analysis conducted in the body. Every claim has been reviewed and decided by the author. The bibliography has been checked against publisher, journal and library records, and where sources disagree on the details of a citation the entry records the disagreement in place of asserting one reading.

Suggested citation. Huang, Wanhong. “The Temporal Correspondence Limitation of Evidence-Based Policy Making: A Preliminary Discussion Towards an Astrophysics-Inspired State-Aware Epistemology of Governance.” Working draft, 2026.

1 Introduction

A body of evidence assembled to inform a decision describes a system as that system stood when the evidence was generated. The decision it informs takes effect later, and the system continues to move in the interval. This paper takes that interval as its object. It argues that correspondence between the state a body of evidence represents and the state obtaining at the moment of decision is a requirement on the use of evidence in governance, that the apparatus by which evidence-based policy making assesses evidence leaves the requirement unstated, and that the requirement fails in ways that the established assessment categories cannot register. The paper's method is comparative: it examines an observational science whose object is available only through signals that departed it in the past, catalogues the elements of practice by which that science works under the condition, and tests each element for transfer to governance, treating the failures of transfer as findings.

The requirement can be stated in one line. Let $x(t)$ denote the condition of a governed system, t_e the time at which a body of evidence was generated, and t_d the time at which a decision informed by that evidence takes effect. Evidence-based policy making assesses whether the evidence supports a claim about $x(t_e)$. Governance acts on $x(t_d)$. The two coincide only where the system has been still. Three forms of divergence are available: the state may have moved, $x(t_e) \neq x(t_d)$; the dynamics may have changed, so that relations estimated under one regime govern behaviour under another; and the relational structure connecting the system's components may have been reorganised, so that quantities measured at t_e stand in different relations at t_d . The first is a matter of degree, the second is the structural problem identified in macroeconomics by Lucas [31], and the third is the condition under which the categories of measurement themselves come loose.

The limitation is distinguishable from the classical epistemological objection to induction, and the distinction carries weight. Hume's problem concerns whether regularities observed in the past license expectations about the future [22]. Temporal correspondence concerns something else, and survives the concession of everything Hume's problem places in doubt. Grant that nature is uniform. Grant that the causal relation established by a well-conducted trial holds at the moment of decision exactly as it held during the trial. The decision is still addressed to a state the system has left, and a stable law applied to a superseded state issues in an intervention aimed at no existing condition. The correspondence failure is therefore reachable from inside an empiricist position, using only claims about measurement latency and system dynamics that any empiricist already licenses. This paper treats that internal character as the limitation's principal argumentative property.

The absence of the requirement from the assessment apparatus is a matter of record, open to settlement by inspection, and Section 2 settles it by examining the instruments. The reason for the absence is instructive. The evidence hierarchies that evidence-based policy making inherited from clinical epidemiology were developed for a setting in which the interval between measurement and action is short and the system under study is held as nearly fixed as the design permits. Under those conditions temporal correspondence holds by construction and needs no name. Carrying the hierarchies into governance carries the silence with them, and

the silence becomes consequential in a setting where the interval is long, the system moves on its own, and the intervention itself is one of the things moving it.

Astrophysics is the comparison case, and the choice requires a word of justification, since a comparison across such a distance invites the charge of metaphor. The justification is that astrophysics shares the structural condition exactly. Its object is accessible only through signals that left the object in the past; the delay is a property of the observing situation and survives every improvement in instruments; and the discipline has responded by building an epistemology in which the historical character of observation is assumed at the outset. The comparison is used here for a specific purpose. Astrophysics supplies a worked inventory of what a science does when it accepts that its observations are historical, and that inventory can be put to governance as a set of questions. For each element of practice, what does the element presuppose, does governance satisfy the presupposition, and what remains if it does not. This paper asks those questions of fifteen elements and reports the answers, including the negative ones. An element that transfers badly is retained in the exposition and its failure analysed, because the pattern of failure is what characterises governance as an observational situation.

The paper accordingly makes no methodological proposal. It recommends the construction of no estimation system, and it should be read as declining that route explicitly. The reason is internal to the critique the limitation belongs to. Where the response to lagged evidence is a more elaborate apparatus for inferring the present state, the resulting estimate is a further symbolic object standing between the decision and the world, produced by machinery whose assumptions are inspectable by fewer parties than the evidence was. A limitation whose remedy deepens the conditions its companion limitations describe is not remedied. The paper's consequences are therefore stated as conditions on the admissibility of intervention and as an allocation of the burden of justification.

Two results are argued. The first is a partition. The reconstruction of an unobserved state from observed signals requires an observability condition, and where the condition fails no accumulation of data recovers the state, since the failure is structural and not statistical. The state space of a governed system therefore divides into a region in which estimation is legitimate and a region in which it is undefined, and the division is a condition on method, with the force of a constraint and none of the character of a counsel of modesty. The second concerns the escape route. Astrophysics converts its lookback problem into an advantage by substituting an ensemble for a time series: objects at many distances are many epochs of one physics, observed together. The substitution requires that the instances share a law. Governance has one system and no such law, and its only available ensemble is comparative, across jurisdictions or populations, which exchanges the temporal mismatch for the contextual mismatch that Cartwright and Hardie [4] analyse. The two mismatches trade against one another, and the choice of which to bear is a choice that is ordinarily neither made nor stated.

The paper proceeds in seven parts. Part I states the requirement and establishes its absence from the assessment apparatus, situates it among three limitations of evidence-based policy making already argued in the literature and in the author's prior work, and reviews prior formulations that approach the same territory from adjacent directions. Part II describes the

comparison case and fixes the structure of the transfer test. Parts III to V run the test over the three families of astrophysical practice. Part VI reports the results: the disanalogy of reflexivity, the trade between the ensemble and the time series, the classification of transfer failures, and the connection between the grammatical tense in which evidence is reported and the displacement of responsibility onto it. Part VII draws the consequences for governance and states the paper's limits.

2 The Evidence-Quality Apparatus and Its Domains of Assessment

This section establishes by inspection that the instruments through which evidence-based policy making assesses evidence contain no category for temporal correspondence. Its objective is evidential: the claim that a requirement is unstated is a claim about documents, and the section examines the documents. It proceeds in three steps. It sets out the structure of the dominant assessment instrument and enumerates the grounds on which that instrument permits the certainty of a body of evidence to be lowered. It examines the neighbouring category that comes closest to the requirement and states the boundary of that category. It then turns to an operating national programme, where the instruments of assessment are administrative in origin, and finds the same absence in a different vocabulary.

2.1 The Structure of Certainty Assessment

The Grading of Recommendations, Assessment, Development and Evaluations approach is the most widely adopted system for rating the certainty of a body of evidence, and its categories have been carried well beyond clinical medicine into the assessment of evidence for policy. Under that approach a body of evidence receives an initial rating determined by study design and is then lowered on the strength of assessments in five domains: risk of bias, inconsistency, indirectness, imprecision, and publication bias [5]. Three further considerations permit an initial rating to be raised. The result is reported as one of four levels of certainty, and the assessments are recorded with explicit reasons in a summary table.

The five domains are exhaustive in the operative sense: an assessor working within the approach lowers certainty on these grounds and records the reason under one of these headings. It is therefore possible to ask directly whether a body of evidence describing a system that has since changed can be downgraded, and under which heading. Risk of bias concerns the conduct of the studies. Inconsistency concerns disagreement among their results. Imprecision concerns the width of the interval around the estimate. Publication bias concerns what has been withheld from the record. None of the four is available for the case in view, in which every study was well conducted, the results agree, the interval is narrow, nothing has been withheld, and the system has moved.

2.2 The Boundary of Indirectness

Indirectness is the domain that comes closest, and its scope is worth stating precisely because the closeness is instructive. The domain covers the distance between the evidence in hand and

the question being asked, and it is elaborated in terms of the population studied, the intervention delivered, the comparator against which it was tested, and the outcome measured [17]. An assessor invoking indirectness is recording that the studies were conducted on people, treatments, comparisons or endpoints differing from those at issue in the decision.

The elaboration is spatial and populational throughout. It anticipates that evidence generated in one setting will be applied in another and asks how far apart the settings are. What it does not anticipate is that the setting of application may be the setting of generation at a later moment, and that the two may have come apart through the passage of time alone. A population may be the same population, an intervention the same intervention, a comparator and an outcome identically specified, and the system in which they are now embedded may nevertheless be a different system. The domain has a category for the distance between here and elsewhere and no category for the distance between then and now.

The gap can be stated as a claim about the apparatus.

Claim 2.1. (Absence of a temporal domain). Under the dominant instrument for assessing the certainty of a body of evidence, no ground is available for lowering certainty on the sole basis that the interval between the generation of the evidence and the moment of decision is long, or that the system described has moved materially within that interval. The nearest available ground concerns distance between settings and is elaborated without reference to elapsed time.

Two qualifications belong with the claim. Practitioners are aware that evidence ages, and the awareness is documented: an early study of clinical practice guidelines estimated that about half were outdated within some six years, with about a tenth outdated within four, and observed that the rate of deterioration varies strongly by topic [41]. A later survival analysis of individual recommendations put the same phenomenon on a firmer empirical footing [34]. The response to this awareness has taken the form of periodic revision schedules. A schedule is a procedural remedy applied to a body of evidence as a whole, and Claim 2.1 concerns something narrower and more consequential: whether a particular body of evidence, offered in support of a particular decision at a particular moment, can be assessed for its correspondence to the state then obtaining. The schedule answers when to look again. It does not answer whether what is in hand supports the decision now being taken.

2.3 Administrative Instruments of an Operating Programme

Assessment instruments developed in clinical epidemiology are one route by which the requirement might have been registered. A second route runs through the administrative machinery of governments that have institutionalised the approach, and it is worth examining because its vocabulary is independent.

Japan supplies a developed instance. A programme of evidence-based policy making was established following the final report of the Statistical Reform Promotion Council in 2017, supported by a promotion committee under the administrative reform apparatus and by units within the Cabinet Office and the Ministry of Internal Affairs and Communications [18, 3]. Its working instruments are the logic model, the administrative project review sheet, and the

policy dashboard. The logic model articulates the chain from activity through output to intended outcome. The review sheet records the articulation and the measurement against it. The dashboard renders the relevant indicators visible.

Each instrument carries a temporal structure and none carries a temporal correspondence check. A logic model states the mechanism by which an activity is expected to produce an outcome and is assessed for the coherence of that statement. The assessment asks whether the chain is plausible and whether its links are evidenced. It does not ask when the state was measured that the chain was designed against, nor how far that state may have moved by the time the activity begins. An indicator on a dashboard is presented as a current value, and its reference period, where recorded at all, is recorded as a property of the statistic, carrying no qualification on the inference the dashboard supports. The review sheet records measurement against targets set earlier, and the interval between the setting and the measurement is treated as the period over which effects accumulate, never as an interval over which the premise of the target may have expired.

The point generalises beyond the instance. Administrative evidence systems index data by reference period as a matter of statistical practice, and the indexing is real. What is absent is the inferential step that would convert the index into a qualification: a statement that the support the evidence lends to a claim about the present is a function of the interval and of the system's movement within it. The index is present in the data and absent from the reasoning.

3 Established Limitations of Evidence-Based Policy Making and Their Grounds

This section situates the requirement stated in Section 2 within a set of three limitations of evidence-based policy making that have been argued elsewhere, and establishes the property that distinguishes the fourth from the other three. Its objective is placement: each of the three is stated in the form its proponents give it, with attention to the kind of premise it requires, and the section closes by identifying what each premise costs an interlocutor who declines it. The method is comparative in a narrow sense. The three limitations are compared along one axis, the character of the ground on which each rests, and the comparison prepares the argument of Section 4 that the fourth occupies a position none of the three occupies.

3.1 The Inferential Gap between Evidence and Prescription

The first limitation is the oldest and concerns the passage from statements of fact to statements of what ought to be done. Evidence establishes what has been observed. A policy is a prescription. The passage between them requires a premise that no quantity of observation supplies, and the requirement holds however good the observation is [22]. A body of evidence establishing that an intervention produced an effect establishes nothing about whether the effect is to be sought.

The limitation is well understood by proponents of evidence-based policy making and is ordinarily met by a division of labour: evidence settles matters of fact and the political

process settles matters of value. The division is coherent as far as it goes. What it concedes is that the authority of evidence is bounded at the point where the prescription is issued, so that the phrase describing policy as based on evidence overstates the relation. The premise the limitation requires is a claim about the logical relation between two classes of statement, and an interlocutor declining it must hold that a prescription follows from a description alone.

3.2 The Remainder Left by Quantification

The second limitation concerns what is lost in the operation by which a condition becomes a measurement. To bring a human situation into an evidence base is to select from it the features an instrument registers, and the selection is not neutral with respect to what the situation is. Suffering enters as an index, participation as a rate, the quality of a relation as a score. The remainder is unavailable to the assessment and therefore unavailable to the decision, and its unavailability is systematic, since it follows from the very operation that made the assessment possible.

The premise this limitation requires is substantive: that features of human situations exist which resist exhaustive representation by the instruments in use, and that their resistance is a property of the features and not a temporary limitation of the instruments. An interlocutor who holds that any relevant feature is in principle measurable declines the premise coherently, and the dispute between the two positions is old and unresolved. The limitation is therefore argumentatively expensive: it can be pressed only against an opponent who has already conceded a claim about the limits of measurement.

3.3 The Displacement of Decision onto a Symbolic Authority

The third limitation concerns the position that an appeal to evidence occupies in the structure of a decision. In its administrative form the observation is familiar. Where a decision is contested and its author accountable, the invocation of evidence relocates the ground of the decision from the person deciding to a procedure, and a decision defended by the statement that it followed the evidence attributes the decision to an authority that occupies no office and can be asked no questions. The structure is recognisable in ordinary practice, where referral to a committee, a framework or a standard operates so that the outcome is attributable to a process and to no participant in it.

The administrative form leaves a question unanswered, and the question is what gives the limitation its force. A relocation of this kind succeeds only where something is available to receive what is relocated, and the appeal to evidence therefore presupposes that evidence occupies a position of authority that no participant in the decision occupies. The account of that position taken here is the psychoanalytic one.

A decision requires a warrant it cannot supply from within itself, and the place from which such a warrant is taken to issue is what Lacan describes as the locus of the Other: a position, internal to the symbolic order and occupied by no person, from which utterances are taken to receive their guarantee [28]. The position is structural, which is to say that it persists across the successive occupants installed in it, and the history of political authority may be read as

a sequence of such installations, from revelation through natural law and the sovereign will to History. What matters for the present argument is that the position is constant and its occupant is not, so that the vacancy left by a departing occupant is a vacancy requiring to be filled.

Two routes by which a contemporary governing arrangement might supply the warrant from within itself are unavailable to it. Genuine collective deliberation, in which a warrant issues from an organised constituency, is impeded by the atomisation of the constituencies that would conduct it. Delegation to a class of expert decision-makers is impeded on two counts at once, since the bounded rationality of any such class undermines the epistemic claim and the absence of a mandate undermines the democratic one. The place is accordingly empty and the pressure to fill it is a pressure the arrangement cannot resolve internally, which is the condition Habermas describes as a crisis of legitimation [20].

Evidence is installed in that place. What follows for responsibility is the consequence stated in the administrative form above, and it now has a reason: the party deciding is not evading accountability by a manoeuvre available to it, but occupying a position in which the warrant for its act is structurally located elsewhere. This is why the deflection is so stable, why it is performed by parties who are acting in good faith, and why exhortations to take responsibility have so little purchase against it.

The premise this limitation requires is correspondingly substantial. It is a structural account of where the authority of a decision is located, and an interlocutor who holds that the appeal to evidence disciplines decisions more than it shelters them may decline the account and confine himself to the administrative observation, which is weaker and is contestable by observation of practice. The limitation is therefore available in two strengths, and the stronger requires a premise drawn from outside the epistemology it criticises. Like the second limitation, it concerns the uses to which evidence is put and leaves untouched the question of what the evidence supports.

3.4 The Character of the Grounds

The three limitations rest on grounds of three kinds. The first rests on a claim about logical relations between classes of statement. The second rests on a substantive claim about the limits of representation. The third rests on a structural account of where the authority of a decision is located, and admits a weaker form resting on observation of administrative practice. Each is defensible, and each requires its proponent to secure a premise that an interlocutor committed to the sufficiency of empirical method may decline. The first is conceded readily and absorbed by a division of labour. The second and third are argued from positions outside that commitment, and an opponent who holds the commitment may hold his ground.

This is the position the fourth limitation occupies differently, and the difference is the subject of Section 4. The proposition that a measurement takes time to produce is a proposition about measurement. The proposition that a system continues to change while the measurement is being produced is a proposition about the system. Both are of exactly the kind that empirical method licenses and neither requires a premise drawn from outside it. An interlocu-

tor committed to the sufficiency of empirical method is therefore committed to the materials from which the fourth limitation is constructed, and the limitation is pressed against him with his own resources.

4 The Temporal Correspondence Requirement

This section states the requirement formally and fixes the vocabulary that the remainder of the paper uses. Its objective is to make the limitation precise enough to be tested against particular practices, and to establish two properties of it that the later argument relies on: that it is independent of the classical problem of induction, and that it is constructible from premises an empiricist already holds. The section proceeds by definition and then by argument. It fixes the temporal indices attaching to an item of evidence, defines correspondence in a form that is relative to the decision at hand, distinguishes three ways in which correspondence fails, and then argues the independence and internality properties. It closes by stating what the requirement does not assert, since a limitation of this kind is easily inflated into a general scepticism it does not support.

4.1 Temporal Indices on an Item of Evidence

An item of evidence carries several distinct times, and the practice of reporting a single date conceals the distinction.

Definition 4.1. (Indices of an evidence item). Let S be a governed system whose condition at time t is $x(t)$. An item of evidence e concerning S carries a *reference time* t_e , the time of the condition it represents; a *release time* $t_r \geq t_e$, the time at which it becomes available to a party deciding; and, where it has been revised, a *vintage*, the release at which its present values were fixed. A decision informed by e is taken at $t_d \geq t_r$ and takes effect over an interval beginning at $t_d + \delta$.

The interval $t_r - t_e$ is the production interval, comprising collection, processing, verification and publication. The interval $t_d - t_r$ is the deliberative interval. The interval δ is the implementation lag. The quantity that matters for the present argument spans the first two.

Definition 4.2. (Governance lookback time). The *governance lookback time* of an item of evidence relative to a decision is $\tau_G = t_d - t_e$, the interval between the condition the evidence represents and the moment at which the decision it informs is taken. Where the effects of the decision are at issue, the relevant interval extends to $t_d + \delta - t_e$.

The vintage index deserves emphasis because it is established practice in one field and absent in most others. Macroeconomic statistics are revised after release, and the analysis of policy conducted on revised figures differs from the analysis available to the party who decided. Croushore and Stark [8] constructed a data set of successive vintages precisely to test whether published results survive the substitution, and the exercise established that the release index carries content. An evidence base without vintage indices cannot reconstruct what any past decision was taken on.

4.2 Correspondence Stated Relative to a Decision

A definition of correspondence in terms of the distance between $x(t_e)$ and $x(t_d)$ would require a metric on the state space, and governed systems seldom supply one. The definition below avoids the requirement by fixing correspondence relative to the decision the evidence is offered to support.

Definition 4.3. (Temporal correspondence). Let a decision be a choice among alternatives $a_1, \dots, a_k$, and let E be a body of evidence offered in support of that choice. *Temporal correspondence holds* for E relative to the decision when the movement of S over the interval $[t_e, t_d]$ is insufficient to alter which alternative E supports. Correspondence *fails* when the movement is sufficient to alter it.

Three consequences follow from the relativisation and are worth stating. Correspondence is a property of a body of evidence together with a decision, so the same evidence may correspond for one choice and fail for another; a coarse choice between acting and declining to act tolerates a larger τ_G than a choice among calibrated intensities. Correspondence admits no assessment from the evidence alone, since the assessment requires a claim about how the system moves, and that claim is a further empirical commitment. And correspondence is not a matter of the evidence being wrong. A body of evidence that fails correspondence remains an accurate description of $x(t_e)$; what has lapsed is its bearing on the choice.

4.3 Modes of Correspondence Failure

Correspondence fails in three structurally distinct ways, and the distinctions organise Parts III to V.

Definition 4.4. (Modes of correspondence failure). Correspondence fails *in state* where $x(t_e) \neq x(t_d)$ with the dynamics and the descriptive categories unchanged. It fails *in dynamics* where the relations governing the system's evolution at t_d differ from those obtaining at t_e , so that a relation estimated under the earlier regime misdescribes the later one. It fails *in articulation* where the categories by which S was described at t_e stand in altered relations at t_d , so that a quantity retaining its name has changed what it measures.

Failure in state is the mode most readily acknowledged and the least severe, since it is a matter of degree and is in principle correctable by more recent measurement. Failure in dynamics is the structural problem identified by Lucas [31] for econometric policy evaluation: relations estimated under one policy regime alter when the regime alters, so the estimated relation supplies no guide to behaviour under the contemplated policy. The mode is general beyond economics wherever the governed parties adjust to the rule. Failure in articulation is the least discussed and the most damaging to longitudinal reasoning, since it defeats comparison itself; where a definition of employment, of poverty or of a diagnosis has been revised between two measurements, the series presenting them together presents two quantities under one name. Section 18 treats this mode directly.

4.4 Independence from the Problem of Induction

The requirement is distinct from the classical objection to induction, and the distinction is established by concession.

Proposition 4.5. (Independence of the correspondence failure). Correspondence failure is not a consequence of the problem of induction, and the concession of everything that problem places in doubt leaves the failure in force.

The argument proceeds by granting the strongest available position to an opponent. Grant that nature is uniform, so that regularities observed hold generally. Grant further that the causal relation established by a well-conducted study at t_e obtains at t_d in exactly the form established. Neither concession bears on where the system stands. A stable relation between an intervention and an outcome describes what follows from the intervention given the condition on which it acts, and where that condition has changed the relation issues in a different outcome, or in no outcome, without ceasing to hold. The uniformity of nature is a premise about laws. Correspondence is a requirement on states. Securing the former settles nothing about the latter, and the failure therefore survives the concession.

The converse holds as well and is worth recording. Where a system is genuinely still, the problem of induction remains in full force and correspondence is satisfied. The two conditions vary independently, which is what establishes them as distinct requirements on the use of evidence.

4.5 Constructibility from Empirical Premises

The property that gives the requirement its argumentative reach is that it can be assembled from materials an empiricist already accepts.

Claim 4.6. (Internal constructibility). The correspondence failure is derivable from two propositions of a kind that empirical method licenses: that the production of a measurement occupies an interval, and that a system under observation continues to change during that interval. An interlocutor committed to the sufficiency of empirical method is committed to both, and therefore to the failure.

Both propositions are themselves matters of observation. The production interval of an official statistic is a documented administrative fact. The movement of a governed system across an interval is measurable by the same instruments that measured its earlier condition. Neither proposition requires a claim about the limits of representation, about what quantification omits, or about what institutions do with an epistemic authority. The limitation therefore occupies the position identified at the close of Section 3: it is pressed against a commitment to empirical method using the resources that commitment supplies, and an opponent who declines it must deny either that measurement takes time or that systems move.

4.6 Heterogeneity of the Evidence Base

Definitions 4.1 and 4.2 attach indices to a single item. A decision is ordinarily informed by a body of items, and the indices differ across them.

Definition 4.7. (Temporal heterogeneity). An evidence base $E = \{e_1, \dots, e_n\}$ is *temporally heterogeneous* relative to a decision when the reference times $t_{e,1}, \dots, t_{e,n}$ differ. Its lookback profile is the collection $\{\tau_{G,i}\}$ of item-wise lookback times.

The heterogeneity has consequences beyond the aggregation of a set of delays into an average, and Section 13 develops them. It is recorded here because it defeats a natural simplification. Where the profile is heterogeneous, no single state of the system is described by the base as a whole, and the object the base appears to describe, the condition of the system at a moment, is assembled from fragments belonging to different moments and exists in no source.

4.7 Restrictions on the Scope of the Requirement

Four restrictions bound the claim, and each is required to keep the limitation from expanding into a general scepticism it cannot support.

The requirement asserts no general staleness. Correspondence is satisfied for a great many decisions, since a great many governed systems move slowly relative to the intervals at issue. The claim is conditional, and the conditions under which it fires are the subject of Sections 10 and 16.

The requirement supplies no argument against the use of evidence. A decision taken without evidence stands in a worse relation to the system's condition, since it corresponds to no state at all. The requirement bears on how the support that evidence lends should be characterised and on what follows where the support has lapsed.

The requirement licenses no ranking of evidence types. Nothing here establishes that observational evidence corresponds better than experimental evidence, or the reverse. The lookback profile of a body of evidence is a fact about its production and is largely independent of its design.

The requirement is stated for a system whose movement is independent of the decision the evidence informs. Where the parties within a governed system anticipate an intervention, part of their response occurs before it is adopted and is recorded at reference times preceding the decision, so that the interval between a cause and its recorded effect is not of fixed sign. Section 23 develops the consequence, which qualifies Definition 4.3 in a direction the present section does not accommodate.

The requirement supplies no measure. Definition 4.3 fixes correspondence relative to a decision and delivers no quantity, and the paper introduces none. Section 10 develops a notion of an interval beyond which support lapses, and states there the conditions under which that interval can be bounded and the wide class of cases in which it cannot.

5 Relation to Prior Work

This section places the requirement of Section 4 among formulations that approach the same territory from adjacent directions, and identifies what each owns. Its objective is to concede priority where priority is owed and to isolate what remains unclaimed. The order runs from

the formulation nearest to the present one, through treatments of the spatial counterpart of the problem, then through fields that have confronted lag in their own practice, and closes with work on the political conditions of evidence use and on decision under deep uncertainty. Each subsection states what the formulation establishes and where its boundary falls. The section ends with a statement of the residue.

5.1 Temporal Validity and the Decay of Knowledge

The nearest formulation is Munger's [35], and it holds a substantial part of the present ground. Munger names the problem temporal validity, treats it as a species of external validity, and argues that the accumulation of internally valid causal estimates cannot deliver generalisable knowledge where the world changes at a rate comparable to the rate of accumulation. He introduces a rate parameter for the change of the world, observes that all published knowledge comes from the past while all action lies in the future, and concludes that assumption-free external validity is unavailable when the target context is future. The argument is anchored in Hume and pressed as a case for reallocating methodological effort.

Everything in that formulation is conceded. Two features fix its boundary and open the space the present paper occupies. Its object is the causal effect estimate, and the question is whether an estimate transports to a target; the present object is the condition of the system, and the question is whether any claim about that condition is supported at the moment of decision. And its remedy runs toward descriptive knowledge, on the ground that description is cheaper and decays more slowly than causal estimation. The requirement stated in Section 4 bears directly on that remedy. Descriptive knowledge carries its own production interval, its own vintage, and its own lookback profile, so a retreat from causal estimation toward description relocates the difficulty and leaves it in place. The present paper may therefore be read as taking up Munger's argument at the point where his recommendation begins.

Two further differences are structural. Munger's rate parameter is a single quantity attaching to a field of inquiry, where Definition 4.7 requires a profile of item-wise intervals within one evidence base. And the conclusion drawn there concerns the allocation of research effort, where the present conclusions concern the admissibility of intervention.

The same territory has been entered from development economics under the description of temporal external validity, where the observation is that a causal effect estimated in one year embeds the realisation of aggregate conditions in that year, so that a confidence interval computed from a single-year study understates the uncertainty attaching to any later application [39]. The term is therefore occupied and is used here as the ancestor's term.

5.2 The Spatial Counterpart and Its Formal Treatment

The problem of carrying a finding from the setting that produced it to the setting in which it will be applied is the subject of a developed literature, and the present requirement is its temporal counterpart.

Cartwright and Hardie [4] argue that instruments ranking methods by the trustworthiness of the evidence they produce leave unaddressed the question of what is required for a

finding to hold where it is to be used, and that the support factors present in the original setting must be present in the target for the finding to carry. Deaton and Cartwright [9] press the point against the standing of randomised trials as a general warrant. Formal treatments of the same problem, developed within the graphical approach to causation, state conditions under which an effect may be transported across populations [1], and within the potential-outcomes framework state conditions under which an estimate generalises to a target context [11]. Reiss [38] argues that external validity is defined in terms of internal validity in a way that misconstrues the practical question.

The boundary of this family is the same throughout. Its target is elsewhere, and its variables of interest are the features by which one setting differs from another. Where the target is the same setting at a later moment, the machinery has no purchase, since the population, the intervention, the comparator and the outcome may be specified identically at both times. The retreat that Egami and Hartman propose, from generalising the magnitude of an effect to generalising only its sign, is instructive here: it is a principled response to the unavailability of the required data about the target, and the present paper reaches a structurally similar retreat by a different route in Section 10.

5.3 Lag in the Analysis of Economic Policy

Economics has confronted the interval directly and has a settled vocabulary for it. The distinction among the interval before a change is recognised, the interval before a decision is taken, the interval before the decision is implemented, and the interval before its effects appear is standard, and Friedman [14] argued from the length and variability of these intervals that activist stabilisation policy may destabilise, since a measure calibrated to a condition may take effect on its opposite. The argument is the present requirement stated for one domain and drawn toward one conclusion.

Two things follow for the present paper. The claim that the interval has gone unnoticed would be false, and the paper makes no such claim. What has gone unregistered is narrower and concerns the instruments of assessment examined in Section 2: a field that reasons carefully about lags in its own subject matter may still assess a body of evidence by criteria containing no place for them. The vocabulary developed in economics is used here, and the term vintage is taken from that literature [8] in preference to a coinage.

5.4 Delay in Cybernetic and System-Dynamics Treatments of Governance

A further tradition treats government as an information-processing arrangement and makes delay one of its governing parameters. Deutsch [10] analysed political systems in terms of the intervals and gains characterising their response to information, and argued that the performance of a governing arrangement depends on these parameters at least as much as on the intentions of those within it. Work in system dynamics established that delays between action and consequence produce characteristic failures in systems operated by people who reason about them without accounting for the delay [13, 42].

The boundary of this tradition is its object. Its concern is the behaviour of the governing

arrangement, and delay enters as a parameter of that behaviour. The present concern is the epistemic standing of a body of evidence offered in support of a particular decision. The two are compatible and the second is presupposed by the first, since a claim about how a system responds to delayed information presumes that the information is delayed and that the delay bears on what the information supports.

5.5 Obsolescence in Clinical Practice

Clinical epidemiology has produced the most direct empirical measurements of the phenomenon. Shekelle et al. [41] estimated the rate at which practice guidelines cease to be valid and found the rate strongly dependent on the subject matter. A later survival analysis of individual recommendations [34] treated validity as a duration and estimated its distribution. Both were discussed in Section 2.

Their boundary was stated there and is restated for the record. The response the field has drawn from these findings is a schedule of revision applied to a body of guidance as a whole. The present requirement concerns a particular body of evidence offered for a particular decision, and is a question about support at a moment, which a schedule does not address. The empirical work nonetheless supplies something the present paper needs: an existence proof that the duration of validity is measurable in at least one domain, which Section 10 takes up.

5.6 The Political Conditions of Evidence Use

Parkhurst [36] distinguishes the technical bias by which evidence is distorted in its use from the issue bias by which the selection of what to evidence forecloses normative debate, and argues for institutional arrangements governing evidence use in place of exhortation to use more of it. He locates one source of difficulty in the inheritance of assessment hierarchies from clinical medicine, where the near-universality of the underlying mechanisms supports a generality that policy settings lack.

That diagnosis and the present one are complementary and address different objects. Parkhurst's subject is the use to which a body of evidence is put and the interests that shape its selection and interpretation. The present subject is a property of the evidence itself, which holds whatever the interests of the parties. The two meet at one point, developed in Section 26: where the support a body of evidence lends has lapsed and the report of it retains the grammatical form of a present-tense claim, the conditions Parkhurst describes operate with less resistance.

5.7 Decision under Deep Uncertainty and Adaptive Planning

A body of work addresses decision where the parties lack agreement on the system model, on the distributions over its inputs, or on the outcomes that matter [33]. Robust decision making [30] seeks strategies performing acceptably across many plausible futures in place of optimising against a predicted one. Dynamic adaptive policy pathways [19] designs a plan as

a sequence of options with conditions, specifying in advance the observable signals at which the plan switches from one pathway to another.

The last of these is the nearest ancestor of the conclusions drawn in Part VII, and its boundary is precise. The pathways approach handles uncertainty about the future by pre-specifying signals whose observation triggers adaptation. It presupposes that the signals are observable when they occur and that their observation is timely enough for the switch to be made. The present requirement bears on that presupposition: a signal is observed through the same apparatus that produces the rest of the evidence base, with the same production interval, and a trigger evaluated on a lagged signal fires late by the length of the lag. The family handles uncertainty about futures, and the present paper concerns uncertainty about the present.

5.8 The Residue Left by Prior Formulations

The formulations surveyed cover a great deal of the territory, and the claim of this paper is confined to what they leave.

Claim 5.1. (The unclaimed residue). Where a body of evidence is internally valid, where the causal relation it establishes is stable and holds at the moment of decision, where the target population, the intervention, the comparator and the outcome are specified identically to those of the studies, and where the context of application is the context of production, the evidence may still fail correspondence in the sense of Definition 4.3, and no assessment instrument surveyed here supplies a ground on which the failure is recorded.

The class is not marginal. It contains every decision taken about a system that has moved since it was measured, which is to say every decision taken about a system in motion. The temporal validity literature registers the difficulty for causal estimates transported to a future target and directs the response toward research strategy. The external validity literature registers it for settings separated in space. Policy economics and the cybernetic tradition register it as a parameter of governing behaviour. The clinical work measures its rate and answers it with a schedule of revision. The politics of evidence concerns the interests shaping evidence use. The deep uncertainty family addresses uncertainty about what is to come. What remains unaddressed is the standing of a particular body of evidence in support of a particular decision at the moment the decision is taken, and it is that standing that Parts III to V examine by comparison with a science whose observations are historical by construction.

6 Observation Under Lookback in Astrophysical Practice

This section describes the comparison case and justifies the comparison. Its objective is to establish that astrophysics stands in the same structural relation to its object as governance stands to a governed system, to set out what that discipline has done about the relation, and to state in advance the respects in which the two situations differ, so that the transfer tests of Parts III to V run against a boundary fixed beforehand. The section proceeds by describing the observational condition, identifying the concession the discipline has made in response to

it, enumerating the elements of practice that the remainder of the paper examines, and setting out the correspondences and the disanalogies.

6.1 The Observational Condition

Astrophysical objects are accessible through radiation and through other signals that propagate at finite speed. An object at distance d is observed as it stood approximately d/c earlier, and the interval is called the lookback time. The interval is a property of the observing situation and survives every improvement in instruments: a better telescope collects more of the signal that has arrived and creates no channel by which a signal might arrive sooner. An observer therefore has access to a hypersurface through the history of the universe, the past light cone of the observation event, on which objects at different distances appear at different epochs.

Three features of this condition are worth isolating, since they are the features that recur in governance.

The delay is structural. It follows from the geometry of the observing situation and from a constant of nature, and no expenditure removes it.

The observed collection is heterochronous. An image of the sky assembles objects observed at many different epochs into a single frame, and the frame therefore depicts no single moment. The apparent simultaneity of the display is an artefact of the display.

The object continues to evolve during the interval. A galaxy observed as it stood eight billion years ago has undergone eight billion years of evolution that the observation does not record, and inferences about its present condition are inferences and not observations.

6.2 The Concession of the Present Tense

The response of the discipline to this condition is worth stating as a matter of language before it is stated as a matter of method, because the linguistic form is the more consequential of the two.

Astrophysics has given up the present tense about its objects. A well-formed astrophysical statement carries an epoch: a quantity is reported at a redshift, a property is attributed to a population at a stated cosmic time, and a claim about how an object stands now is marked as an extrapolation and defended as such. The discipline does not report that a distant quasar is accreting at a given rate. It reports that the quasar was accreting at that rate at the epoch observed, and any claim about the present is carried by a model that is named.

The concession is a loss, and the discipline absorbed it. What it purchased is that every inference from observation to present condition became visible as an inference, which is to say that it acquired an assumption, a defender, and a place where an objection can be lodged. Section 26 returns to this point and argues that it is the element of astrophysical practice with the largest bearing on governance, since evidence in governance is ordinarily reported in a present tense that its production interval does not support.

6.3 The Elements Examined

The methods and observing disciplines by which astrophysics works under the lookback condition are examined individually in Parts III to V, in three families.

The first family concerns the reconstruction of a past state from a present signal, and contains inverse inference from the received signal to a distribution over source conditions; forward modelling, in which an assumed state is evolved and the resulting predicted observation compared with the actual one; the propagation of uncertainty alongside the propagation of state; data assimilation, in which a running model is corrected by arriving observations; and retrodiction with consistency checking, in which evidence from separated epochs is required to cohere through the dynamics connecting them.

The second family concerns constraints imposed by the observing situation itself, and contains the past light cone and the heterochrony it produces; the horizon, beyond which observation is unavailable in principle; selection effects that grow with distance, so that the sample observed at depth is systematically unrepresentative; the cadence of a survey, which determines what classes of variation are detectable at all; the single realisation of the universe and the ergodic substitution by which cosmology partly escapes it; and standard rulers, features of known extent used to calibrate across epochs.

The third family concerns the self-discipline of the observing community, and contains blind analysis, in which the result is concealed from the analyst until the analysis is fixed [27]; the recording of provenance and the re-reduction of archival data under improved calibration; the construction of mock observations, in which a simulated universe is passed through a simulated instrument and pipeline so that the measurement system is tested and not only the model [7]; and the calibration of distance ladders, together with the discipline's handling of persistent disagreement between independent measurement routes [44].

6.4 Notation for the Comparison

The examinations of Parts III to V use a common notation, fixed here so that the presupposition identified in each case may be stated as a condition on an object in it.

Let $x(t) \in \mathcal{X}$ denote the condition of the system under study and θ the parameters of its dynamics and of the process by which it is observed. Let

$$F_{\Delta} : \mathcal{X} \rightarrow \mathcal{X}, \quad x(t + \Delta) = F_{\Delta}(x(t); \theta), \quad (1)$$

denote the flow carrying a condition forward by an interval Δ , and let

$$h : \mathcal{X} \rightarrow \mathcal{Y}, \quad y = h(x; \theta) + \varepsilon, \quad (2)$$

denote the observation operator carrying a condition to a recorded quantity, with ε the error of the recording. In astrophysics the flow of Equation 1 is supplied by physical theory and the operator of Equation 2 by the physics of emission, propagation and instrumentation. In governance the flow is the unstated account of how the governed system evolves and the operator is the measurement system: the definitions, the sampling frame, the recording practice

Table 1: The six approaches by which astrophysics operates under the lookback condition, with the operation each performs and its governance counterpart. Sections 8 to 13 examine each in turn and determine the status of its presupposition.

Approach	Operation in astrophysics	Position in governance
Inverse inference	$y \mapsto p(x, \theta \mid y) \propto p(y \mid x, \theta) p(x, \theta)$	$E \mapsto p(x(t_e) \mid E)$, with h constituted administratively
Forward modelling	$x(t_0) \xrightarrow{F_\Delta} x(t_1)$, then $\hat{y} = h(x(t_1))$ compared with y	articulated mechanism $\mapsto$ predicted evidence stream, compared with the recorded one
Uncertainty propagation	$p(x(t_0)) \mapsto (F_\Delta)_\# p(x(t_0))$, with θ varied across an ensemble	$\text{Var}[x(t_d) \mid E]$ required, $\text{Var}[x(t_e) \mid E]$ reported
Data assimilation	$\hat{x}_t^- = F_\Delta(\hat{x}_{t-1}^+)$, $\nu_t = y_t - h(\hat{x}_t^-)$, $\hat{x}_t^+ = \hat{x}_t^- + K_t \nu_t$	running account of a condition from lagged indicator streams
Retrodiction and consistency	$F_{t_2-t_1}(\hat{x}(t_1))$ compared with $\hat{x}(t_2)$ from independent evidence	estimates $\{\hat{\theta}_k\}$ from studies at $\{t_{e,k}\}$, pooled as one quantity
Light-cone reasoning	observations lie on $\mathcal{L}^-(O)$, each carrying its own epoch	$E = \{e_i(t - \tau_{G,i})\}$, with the profile partly unrecorded

and the processing described in §2.3.

The temporal indices are those of Definition 4.1. An item of evidence e_i carries a reference time $t_{e,i}$ and a governance lookback time $\tau_{G,i} = t_d - t_{e,i}$, and an evidence base is written $E = \{e_i\}$ with the lookback profile $\{\tau_{G,i}\}$ of Definition 4.7.

Table 1 states the six approaches by which astrophysics works under the lookback condition, together with the operation each performs in this notation and the position each occupies in governance. The table is a summary of the material examined in Sections 8 to 13 and settles nothing by itself.

6.5 Correspondences and Disanalogies

The comparison is licensed by structural correspondence, and its limits are fixed by disanalogies that are stated here in advance of the sections that develop them.

Four correspondences hold. The object of knowledge is accessible only through signals with a finite production interval. The interval is a property of the observing situation and survives improvement in instruments, since the production of an official statistic requires collection, processing and verification whatever the technology. The collection of available observations is heterochronous, so that what appears as a description of a moment assembles material from many moments. And the object continues to evolve during the interval, so that a claim about its present condition is an inference carried by a model.

Five disanalogies hold, and each is developed at the point where it bears.

The lookback interval is known exactly in astrophysics and follows from a measured distance and a constant. In governance the interval is itself uncertain, differs across sources, and is frequently unrecorded. Sections 13 and 20 develop the consequence.

The dynamics connecting the observed epoch to the present are supplied in astrophysics by physical theory of established scope. Governance possesses no comparable law, and the

sections of Part III that presuppose one report failures for that reason.

The astrophysical observer does not act on the observed object. A governing party observes in order to act, and the acting alters what is subsequently observed. Section 23 treats this as the principal disanalogy.

The astrophysical object does not read its own description. A governed system contains parties who read the description and adjust to it, so that the description enters the dynamics it describes. This is developed together with the preceding point.

Astrophysics observes many instances of one physics at many epochs and converts its look-back condition into a means of studying evolution. Governance has one system. Section 24 argues that this is the deepest of the five and that it has a cost that has ordinarily gone unstated.

The disanalogies are not caveats attached to a comparison otherwise carried through. They are the substance of what the comparison yields, since the pattern of what fails to transfer is a characterisation of governance as an observational situation, and it is that characterisation the paper is after.

7 The Structure of the Transfer Test

This section fixes the procedure by which each element identified in Section 6 is examined, and the categories in which the results are reported. Its objective is to make the examination uniform, so that the fifteen sections following it are comparable and their results can be assembled into a single classification. The section states the template applied to each element, defines the four classes into which the outcomes fall, states the order in which the elements are presented and the reason the order of presentation differs from the order of the result, and closes with the restrictions on what the test establishes.

7.1 The Template Applied to Each Element

Each element is treated in five moves, carried in prose and in a fixed order.

The first move states the structure of the element as astrophysics employs it, at a level of detail sufficient to identify what the element does and no greater.

The second move identifies the presupposition on which the element depends. This is the load bearing move, and the discipline it requires is to name a condition that could fail. An element whose presupposition cannot be stated in a form admitting failure has been described and not analysed.

The third move identifies the governance counterpart: the practice, instrument or reasoning step that occupies the corresponding position in the assessment or use of evidence for policy.

The fourth move determines the status of the presupposition in governance, and supplies the reasons.

The fifth move states what survives. Where the presupposition holds, the element yields a requirement that governance can adopt. Where it fails, the element yields a condition under

which a claim about the present state is unsupported, and the paper records that condition as a restraint trigger in the sense developed in Section 28.

7.2 Classification of Outcomes

The outcomes of the fifth move fall into four classes, and the classes are distinguished by the reason for the outcome and not by its severity.

Definition 7.1. (Classes of transfer outcome). An element falls in class *i*, *precondition absent*, where its presupposition requires a quantity, a law or a structural property that governance does not possess, and where the deficiency is a fact about the governed system independent of the conduct of any party. It falls in class *ii*, *reflexivity*, where the presupposition requires that the observed system evolve independently of the observation and of the action the observation informs. It falls in class *iii*, *interested observer*, where the presupposition requires that the measurement apparatus, or the choice of what to measure, be settled independently of the party whose conduct is at issue. It falls in class *iv*, *transfers intact*, where the presupposition holds in governance and adoption of the element requires the alteration of practice with no new apparatus.

Three features of the classification are worth stating before it is applied.

The classes are not exclusive, and an element may fail in more than one way. Where this occurs the section states the primary and the secondary class and the table records both, since the combination is informative.

Class *iv* is defined by what adoption costs. An element transfers intact when what stands between governance and its adoption is the alteration of a practice, the keeping of a record, or the acceptance of a constraint on how an analysis is conducted. Elements requiring the construction of an estimation system fall outside the class whatever their merits, and the reason was stated in Section 1.

The classification is a result and not a scheme imposed in advance. The four classes were arrived at by conducting the examinations and observing that the failures fell into these kinds. They are presented here so that the reader may hold them while reading Parts III to V, and Section 25 assembles the assignments and draws the consequence.

7.3 Order of Presentation and Order of Result

The elements are presented in the order of the three families of Section 6: the reconstruction of a past state, the constraints of the observing situation, and the discipline of the observing community. That order follows the structure of astrophysical practice and places the most familiar material first.

The order of the result differs. Assembled by class, the elements sort so that the third family falls almost entirely in class *iv* and the first family contains the failures on which the argument turns. The argument therefore travels through Table 2 in Section 25 and not through the sequence of sections, and a reader following the argument alone may proceed from Section 6 to Section 25 and return.

7.4 Restrictions on the Reach of the Test

Four restrictions bound the results of Parts III to V.

The test establishes nothing about astrophysics. Every element is described as that discipline employs it, and a finding that an element fails to transfer is a finding about governance. No criticism of the astrophysical practice is intended or implied at any point.

The test establishes no completeness. The fifteen elements were selected as those bearing most directly on the historical character of observation, and the selection is defensible and not exhaustive. A further element may be added by anyone who identifies one, and the template of The template above supplies the form in which to treat it.

The test establishes no impossibility. A finding that an element fails on a presupposition governance does not currently satisfy leaves open that the presupposition may be satisfied in particular domains, and several sections identify such domains. What the finding excludes is the general adoption of the element as a method for governance at large.

The test yields no procedure. The results are conditions on the admissibility of a claim about the present state and on the allocation of the burden of justification. They compose into no algorithm and are administered by no office, and Section 29 states the cost of that refusal.

8 Inverse Inference and the Posterior Over Present States

This section examines the first element of the first family, the inference that runs from a received signal back to the condition of the source. Its objective is to identify the presupposition on which that inference rests in astrophysics, to locate the corresponding inference in the assessment of evidence for policy, and to determine the status of the presupposition there. The section follows the template fixed in Section 7 and concludes with an assignment under Definition 7.1. Its result is that the element fails principally because the apparatus converting the condition of a governed system into an observable quantity is settled by parties whose conduct that quantity is used to assess.

8.1 The Structure of the Inference in Astrophysics

An astrophysical observation delivers a signal: a flux, a spectrum, a light curve, a strain in a detector. The condition of the source is recovered by inverting the observation operator. What arrives is

$$y = h(x(t_{\text{obs}} - \tau); \theta) + \varepsilon, \quad (3)$$

and the inference delivers a distribution over conditions consistent with it,

$$p(x, \theta \mid y) \propto p(y \mid x, \theta) p(x, \theta), \quad (4)$$

in which the second factor is the prior, or equivalently the regularisation the inversion requires.

Two features of the practice matter here. The inference is ill-posed in the technical sense: distinct source conditions can produce indistinguishable signals, and small perturbations of

the signal can correspond to large excursions in the recovered condition. The practice therefore requires a regularisation, which in a Bayesian formulation is a prior, and the recovered distribution is a joint product of the data and that choice. And the deliverable is a distribution. A point estimate is reported as a summary of a distribution whose width is reported alongside it, and the width does substantive work in the inferences drawn.

8.2 Presuppositions of the Inversion

The inference presupposes two things about the forward operator and the regularisation.

The forward operator is fixed independently of the answer sought. Writing $\mathcal{A}$ for the set of parties to a dispute about x , the condition is that the operator of Equation 2 carries no dependence on any $a \in \mathcal{A}$. The physics of emission, absorption and propagation is established by work conducted for other purposes, and no party is in a position to adjust it in a favourable direction.

The regularisation is declared and defensible on grounds independent of the result. A prior in an astrophysical inversion is stated, its influence on the posterior is examined by varying it, and a result sensitive to the choice is reported as sensitive.

Both conditions admit failure. The first fails where the party whose conduct is at issue determines the mapping from condition to signal. The second fails where the regularisation is undeclared, so that its contribution to the result is carried without acknowledgement.

8.3 The Governance Counterpart

The corresponding inference in governance runs from indicators to the condition of a governed system. An unemployment rate, a recorded crime figure, an attainment score and an emissions inventory are signals, and the condition each is taken to represent is recovered by an inference that most reports leave implicit. The forward operator in the governance case is the measurement system: the definitions fixing what counts as an instance, the sampling frame, the recording practice of the officials who generate the entries, and the processing applied before publication.

The inference is ill-posed in the same technical sense. Distinct conditions of a governed system produce identical indicator values, since the operator standing in Equation 3 compresses a high-dimensional condition onto a scalar. And the recovery of the condition from the indicator requires assumptions playing exactly the role of a regularisation: assumptions about what has gone unrecorded, about the relation between the recorded quantity and the phenomenon named, and about the stability of recording practice across the period.

8.4 Status of the Presupposition in Governance

The first condition fails, and it fails structurally.

Claim 8.1. (The interested measurement operator). In governance the observation operator is constituted by administrative choices, and those choices are ordinarily made by, or under the

authority of, the parties whose conduct the resulting quantity is used to assess. The operator of Equation 2 is therefore indexed by a party, as in Equation 5,

$$e = h_a(x(t_e)) + \varepsilon, \quad a \in \mathcal{A}, \quad (5)$$

and the posterior of Equation 4 inherits that dependence. The ill-posedness of the governance inverse problem is accordingly constitutive, and no increase in the quantity of data corrects it, since the data are generated through h_a .

Recorded crime supplies a clean instance. The quantity depends on the recording practice of the police, on the classification rules applied, and on the propensity of the public to report, and the police are simultaneously the apparatus generating the signal and a party assessed by it. An astrophysical parallel would require the emission physics of a source to depend on the preferences of the astronomer, and the parallel fails because no such dependence exists.

The general form of this failure is documented independently of the present argument. Campbell [2] stated that a quantitative social indicator becomes more subject to distorting pressure the more it is used for social decision-making, and that the distortion reaches the process the indicator was intended to monitor. Goodhart's observation [16] is the same phenomenon stated for monetary aggregates. Both are ordinarily read as warnings about incentives. Read against the present template they state something narrower and more structural: the forward operator is endogenous to the use made of its output, so the inversion is conducted with an operator whose form has been shaped by the interest of a party to the result.

The failure compounds with elapsed time in a way worth recording, since it bears on the correspondence requirement directly. The longer the interval τ_G , the more opportunity the measurement system has had to be adjusted, and the greater the chance that the operator in force at t_e differs from the operator a present reader assumes. This is the mode of correspondence failure named in articulation at Definition 4.4, and Section 18 treats it in its own right.

The second condition fails in a milder form and is remediable. Regularising assumptions in governance inference are ordinarily undeclared, and their contribution to the recovered condition is therefore carried without acknowledgement. Nothing in the structure of governance prevents their declaration, and the declaration is a matter of practice.

8.5 The Transferable Residue

Two things survive, and they belong to different classes.

The posterior form survives as a requirement and is available at low cost. Where an indicator is offered in support of a claim about a system condition, the claim admits statement as a range of conditions consistent with the indicator, and the width of that range is a reportable quantity. Section 28 attaches the burden of justification to the width, on the ground that a party proposing an irreversible course under a wide posterior is proposing to act on a condition it has not identified.

The declaration of the regularisation survives, and carries a consequence the astrophysical case does not have. A prior in a governance inversion encodes a judgement about what the

unrecorded portion of a phenomenon contains, and that judgement bears directly on which parties the indicator represents. The astrophysical discipline of stating the regulariser therefore migrates into governance as a requirement with a political character, since a declared prior is contestable by a party who holds a different view of what the record omits.

What fails to survive is the inversion itself as a warranted procedure. Where the operator is constituted by an interested party, the recovered distribution inherits that interest, and no increase in the quantity of data corrects it, since the data are generated through the same operator.

Assignment 8.2. (Class of the element). Inverse inference falls in class *iii*, interested observer, on the ground of Claim 8.1, with a secondary assignment in class *iv* for the two surviving requirements, the reporting of posterior width and the declaration of the regularisation.

9 Forward Modelling and the Predicted-Evidence Requirement

This section examines the element that runs in the opposite direction to Section 8, from an assumed condition through a dynamical model to a predicted observation, which is then confronted with the observation actually obtained. Its objective is to determine what governance retains from a practice whose central presupposition governance violates, and its result is divided: the element fails as a method of inference on the ground that the modelled system responds to the modeller, and it survives as a discipline in the form of a requirement that a proposed intervention state in advance what evidence it expects. The section follows the template fixed in Section 7.

9.1 The Structure of the Practice in Astrophysics

Forward modelling assumes a condition at some earlier epoch together with a dynamical law, evolves the condition to the epoch of observation, and generates from the evolved condition the observation an instrument would record,

$$x(t_0) \xrightarrow{F_\Delta} x(t_1) \xrightarrow{F_\Delta} x(t_2), \quad \hat{y}_k = h(x(t_k); \theta), \quad (6)$$

and the generated $\hat{y}_k$ is compared with the recorded y_k . The comparison bears on the assumed condition and the assumed law together.

The structure of the reasoning is worth stating in its contrast with a more familiar form. Reasoning that proceeds from an observation to a conclusion places the whole inferential weight on the step out of the data. Forward modelling places the weight on a confrontation between a predicted observation and an actual one, so that the model is required to generate a commitment before it is assessed. The commitment is the object of the test, and the practice extends to simulating the instrument and the analysis pipeline as well as the physical system, so that what is compared is a mock observation and not a mock world [7].

9.2 Presuppositions of Forward Modelling

Two conditions are required.

The dynamical law holds over the interval of the evolution and is independent of the modeller. The chain of Equation 6 carries the assumed condition forward under a flow whose form owes nothing to the fact that the modelling is being conducted, and a source does not alter its behaviour because a prediction has been made about it.

Degeneracy is characterisable. Distinct pairs of assumed condition and assumed parameters frequently generate indistinguishable predicted observations, so that the set

$$D(y) = \{(x_0, \theta) : h(F_\Delta(x_0; \theta)) = y\} \quad (7)$$

contains more than one element, and the practice requires that $D(y)$ be identified and reported, so that agreement between prediction and observation is read as membership of $D(y)$ and never as identification of a unique condition.

Both admit failure. The first fails where the modelled system responds to the model. The second fails where degeneracy goes unexamined and a single pair is carried forward as though the agreement had selected it.

9.3 The Governance Counterpart

The counterpart is the articulated model of how an intervention is expected to produce its outcome, and Japanese administrative practice supplies a developed instance in the logic model, which states the chain from activity through output to intended outcome and is recorded on the review sheet described in §2.3. Programme theory and realist evaluation occupy the same position in the evaluation literature, where the mechanism through which an intervention is expected to operate is articulated in advance and the evaluation is designed against it [37].

The correspondence is close and the difference is precise. The logic model states a mechanism and is assessed for the coherence of the statement. Forward modelling states a mechanism and derives from it a prediction about what the observing apparatus will record, at a stated time, under a stated measurement definition. The second is a commitment of a kind the first does not make.

9.4 Status of the Presupposition in Governance

The first condition fails, and the failure is the central structural fact about governance as an observational situation.

Claim 9.1. (Response of the system to the model). A governed system contains parties who read the model, anticipate the intervention it recommends, and adjust their conduct accordingly. The flow of Equation 1 therefore carries a dependence on the course adopted, as in Equation 8,

$$x(t + \Delta) = F_\Delta(x(t); \theta, a), \quad (8)$$

so that a flow estimated under a course a_0 misdescribes evolution under a contemplated

course a_1 . Forward modelling in governance evolves a condition under a law that the evolution itself has altered.

The claim is a restatement in the present vocabulary of the objection Lucas [31] raised against the use of estimated econometric relations for policy evaluation, and its reach extends past economics to any domain in which the governed are informed. It is the mode of correspondence failure named in dynamics at Definition 4.4, and Section 23 develops its consequences for the paper as a whole.

The second condition fails in a form that is remediable in principle and neglected in practice. Governance models are ordinarily carried forward in a single specification, and where a predicted outcome is observed the specification is treated as confirmed. The set of alternative mechanisms generating the same observable pattern, which is the governance counterpart of Equation 7, is seldom enumerated, and where it is enumerated the selection among its members is made on grounds that go unstated. The astrophysical discipline here is to report the degeneracy and to decline the identification, and nothing in the structure of governance prevents its adoption.

9.5 The Transferable Residue

The element yields a requirement whose cost is discipline, and the requirement is the most directly transferable result in Part III.

Requirement 9.2. (Predicted evidence). An intervention proposed on the strength of an articulated mechanism states, before adoption, what the evidence stream is expected to show at specified later times if the mechanism holds, under the measurement definitions then in force; and the comparison between the predicted and the recorded evidence is itself dated, since the recorded evidence arrives with its own production interval.

Three features of the requirement are worth stating. It converts a claim about an unobservable present condition into a claim about future observations, which is testable in a way the former is not. It costs no apparatus, since a party able to articulate a logic model is able to state what that model predicts the indicators will show. And the dating clause is load bearing: a comparison conducted at $t + \Delta$ against evidence whose reference time is $t + \Delta - \tau$ is a comparison against a condition earlier than the one the prediction addressed, and an undated comparison conceals the discrepancy.

The requirement carries a limit that must be stated with it. Under Claim 9.1 the confrontation between predicted and recorded evidence assesses a compound of the mechanism and the system's response to the intervention, and the two are not separated by the comparison. A prediction that succeeds may have succeeded because the mechanism held or because the parties adjusted in the anticipated direction. The requirement therefore disciplines the proponent without delivering the identification that forward modelling delivers in astrophysics.

Assignment 9.3. (Class of the element). Forward modelling falls in class II, reflexivity, on the ground of Claim 9.1, with a secondary assignment in class IV for Requirement 9.2 and for the reporting of degeneracy.

10 Uncertainty Propagation and the Evidence Horizon

This section examines the practice of carrying a distribution forward in place of a point, and the notion it yields for governance. Its objective is to establish that the intervals reported alongside policy evidence describe uncertainty about the reference condition and are presented as uncertainty about the decision-time condition, and to define from the astrophysical practice an interval beyond which a body of evidence supports no claim about the present. The section follows the template fixed in Section 7. Its result is a failure of the element as a computational method together with a surviving concept whose ordering properties hold where its magnitudes cannot be obtained.

10.1 The Structure of the Practice in Astrophysics

Where a condition is evolved from one epoch to another, astrophysical practice evolves a distribution over conditions. The distribution at the assumed epoch is carried through the dynamics as its push-forward,

$$p(x(t_0)) \mapsto p(x(t_1)) = (F_\Delta)_\# p(x(t_0)), \quad (9)$$

with the parameters θ varied across an ensemble so that uncertainty about the dynamics is carried alongside uncertainty about the condition, and the distribution at the later epoch is reported with the condition. Where the dynamics are nonlinear the distribution ordinarily broadens with the interval, and the broadening is computed.

Two elements of the practice bear on the transfer. Uncertainty in the model is carried alongside uncertainty in the state, by running ensembles across parameter values and across model variants, so that the reported spread reflects disagreement about the dynamics as well as ignorance about the condition. And the growth of the spread with the interval, obtained by iterating Equation 9, is itself a reported result, which permits a statement of the interval over which a conclusion remains supported.

10.2 Presuppositions of the Propagation

Three conditions are required.

A dynamical law is available through which to propagate.

The error structure is specified well enough that the propagated distribution means something, which requires that the uncertainty attaching to the initial condition and the uncertainty attaching to the dynamics both be represented.

The uncertainty in the model is representable by variation within a family. Ensembles explore disagreement inside a space of models the discipline can write down, and the practice is understood to say nothing about error lying outside that space.

All three admit failure. The first fails where no law is available. The second fails where the uncertainty about the dynamics is unquantified. The third fails where the space of candidate models is unbounded.

10.3 The Governance Counterpart

Policy evidence is ordinarily reported with an interval. A survey estimate carries a sampling interval, an econometric estimate carries a standard error, a projection carries a range across scenarios. The counterpart of astrophysical uncertainty propagation is the operation that would carry such an interval from the reference time of the evidence to the moment of decision.

The observation to be made about that counterpart is that the operation is absent, and its absence has a consequence that the form of reporting conceals.

Claim 10.1. (Misattribution of the reported interval). An interval reported with an item of policy evidence quantifies $\text{Var}[x(t_e) \mid E]$, the uncertainty about the condition at the reference time. The quantity bearing on the decision is $\text{Var}[x(t_d) \mid E]$, and nothing in ordinary reporting practice carries the first into the second. The interval is nevertheless read, and ordinarily presented, as the uncertainty attaching to a claim about the condition at t_d . The reported quantity and the quantity the reader requires are therefore different quantities bearing the same name.

The claim is a statement about a mismatch of reference and admits a direct illustration. A survey conducted over a reference month, published four months later, and used in a decision two months after that, reports a confidence interval computed from its sample. That interval narrows with sample size. The uncertainty attaching to the condition six months after the reference month does not narrow with sample size, since it is compounded of the sampling uncertainty and the unknown movement of the system across the interval. A large survey therefore reports a narrow interval about a condition that may be arbitrarily distant from the one at issue, and the narrowness of the reported interval carries a false assurance.

10.4 Status of the Presuppositions in Governance

The first presupposition fails generally. Governance possesses no dynamical law for the systems it governs comparable to the physics available for the propagation of an astrophysical condition, and the discussion of Section 11 treats the consequences of this absence in their strongest form.

The second fails in a specific and consequential way. Where a propagation is attempted, the uncertainty carried forward is ordinarily the uncertainty about the condition, and the uncertainty about the dynamics is represented by a small number of scenarios chosen for their policy salience. A distribution propagated under a model held with unexamined confidence understates the spread by an amount that grows with the interval, so the failure of this presupposition biases the result in the direction of false precision.

The third fails without qualification. The space of candidate models for a governed system is unbounded, and the movement of such a system across an interval includes changes of kind that no enumerated family anticipates. Section 15 and Section 18 identify two such changes that are systematic and therefore partly anticipable, and the residue beyond them remains.

10.5 The Transferable Residue

The element yields a concept and an ordering, and withholds a magnitude.

Definition 10.2. (Evidence horizon). For a body of evidence E and a decision over alternatives $\mathcal{A} = \{a_1, \dots, a_k\}$, the *evidence horizon* is

$$H(E, \mathcal{A}) = \sup \{ \Delta \geq 0 : p(x(t_e + \Delta) | E) \text{ separates the members of } \mathcal{A} \}, \quad (10)$$

the interval beyond the reference time after which the distribution over conditions consistent with that evidence has broadened past the resolution required to distinguish the alternatives among which the decision chooses. Within the horizon the evidence supports a choice. Beyond it the evidence remains an accurate description of the reference condition and supports no choice among the alternatives at issue.

The horizon of Equation 10 inherits the decision-relativity of Definition 4.3 and shares its advantages. A horizon is defined for a body of evidence together with a decision, so a coarse choice has a longer horizon than a fine one, and the same evidence may lie within the horizon for one question and beyond it for another.

Two results attach to the definition, and they differ in strength.

Claim 10.3. (Measurability in a bounded domain). The horizon is measurable where a domain records both the evidence and the subsequent determination of whether it remained adequate. Clinical guidance supplies the instance. The survival analysis of guideline recommendations conducted by Martínez García et al. [34] estimates a distribution of durations of validity, and the earlier estimate of Shekelle et al. [41] reports the strong dependence of that duration on subject matter. Their finding establishes that the horizon is an empirical quantity in at least one domain and supplies no method for domains lacking a comparable record.

Claim 10.4. (Ordering under sensitivity). Where the magnitude of the horizon is unavailable, an ordering remains available. Between two bodies of evidence with equal reference intervals, the one describing a system nearer a stability boundary has the shorter horizon, since the movement consistent with a given interval is greater there. Between two intervals for one system, the longer has the greater accumulated movement. The horizon therefore shortens with the interval and with the sensitivity of the system, and these two orderings hold without a model that would deliver a magnitude.

Claim 10.4 requires a restriction that is stated here and observed throughout. The ordering is qualitative, and no quantitative composition of the two factors is asserted. A formula expressing the horizon as a function of interval and sensitivity would require the dynamical model whose absence this section has established, and the appearance of such a formula in a framework that denies itself the model would be an equivocation. What the ordering supports is a comparative judgement, and Section 28 shows that a comparative judgement is sufficient for the allocation of a burden.

Assignment 10.5. (Class of the element). Uncertainty propagation falls in class i, precondition absent, on the ground that the dynamical law and the bounded model family it requires are unavailable, with a secondary assignment in class iv for the evidence horizon and for the

disclosure that Claim 10.1 requires, namely that a reported interval be labelled with the time whose condition it describes.

11 Data Assimilation and the Observability Condition

This section examines the element that most directly promises what governance appears to want, a running estimate of the condition of a system maintained by correcting a model with arriving observations. Its objective is to establish the condition on which that promise depends, to show that the condition is structural and admits no remedy by the accumulation of data, and to derive from it the partition that Part VII treats as the paper's principal result. The section follows the template fixed in Section 7. Its result is the strongest failure recorded in Part III together with a residue that is available at no cost in apparatus.

11.1 The Structure of the Practice in Astrophysics and Its Neighbours

Data assimilation maintains an estimate of the condition of a system by alternating two operations. A model advances the current estimate to the time of the next observation, and the arriving observation is then combined with that forecast, weighted by the relative confidence attaching to each,

$$\hat{x}_t^- = F_\Delta(\hat{x}_{t-1}^+), \quad \nu_t = y_t - h(\hat{x}_t^-), \quad \hat{x}_t^+ = \hat{x}_t^- + K_t \nu_t, \quad (11)$$

where ν_t is the innovation, the discrepancy between the observation anticipated and the one received, and K_t is the weight given to it. The cycle repeats [24, 25].

The practice belongs most centrally to numerical weather prediction and to the earth and space sciences adjacent to astrophysics, and it is the reason a forecast issued today incorporates observations arriving continuously. Its attraction for the present subject is evident: the cycle of Equation 11 is a procedure for holding an estimate of a present condition using observations that are individually partial and individually late.

11.2 Presuppositions of Assimilation

Three conditions are required, and the first is the one that matters.

The system is observable through the available measurements. Observability is a structural property relating the dynamics to the quantities measured: writing

$$\mathcal{O} = (h, h \circ F_\Delta, h \circ F_{2\Delta}, \dots) \quad (12)$$

for the map carrying a condition to the sequence of observations it would generate, the system is observable when $\mathcal{O}$ is injective. Where injectivity fails, the state decomposes as $\mathcal{X} = \mathcal{X}_{\text{obs}} \oplus \mathcal{N}$ with $\mathcal{N} = \ker \mathcal{O}$, and the output carries no information whatever about the component lying in $\mathcal{N}$. The failure is not a matter of noise or of sample size. An unobservable component of a condition remains unrecovered under any quantity of data collected through the same measurements, since the map of Equation 12 is by construction insensitive to it.

A dynamical model is available with which to advance the estimate between observations.

The relative confidence attaching to the forecast and to the observation is correctly specified. The weighting operation requires a statement of how much error to expect from each, and the behaviour of the cycle depends on that statement.

All three admit failure, and the third fails in a direction with a name.

11.3 The Governance Counterpart

The counterpart is any arrangement maintaining a running account of the condition of a governed system from streams of partial and lagged indicators. Epidemic surveillance supplies the closest genuine instance, and its status is literal: the estimation of infection incidence from case reports arriving with a delay distribution is a reconstruction of a present condition from lagged observation conducted for the purpose of governing [21, 29]. The dashboards described in §2.3 occupy the same position with none of the machinery, presenting lagged indicators in a display whose form asserts currency.

11.4 Status of the Presuppositions in Governance

The first presupposition fails for a substantial part of what governance would need to know, and the failure is the section's principal result.

Proposition 11.1. (The observability partition). Relative to a given set of measurements, the condition of a governed system divides as $\mathcal{X} = \mathcal{X}_{\text{obs}} \oplus \mathcal{N}$. Within $\mathcal{X}_{\text{obs}}$, estimation of the present condition from lagged observation is a well-posed problem whose difficulty is a matter of degree. Within $\mathcal{N}$ the problem is undefined, and for a component $x_{\mathcal{N}}$ lying there,

$$p(x_{\mathcal{N}} \mid y_{0:t}) = p(x_{\mathcal{N}}) \quad \text{for every } t, \quad (13)$$

so that the posterior equals the prior however long the observation continues. No increase in the frequency, the volume or the quality of observation alters this, since the deficiency lies in the relation between the measurements and the dynamics and never in the quantity of data.

Equation 13 converts a familiar counsel into a condition on method. The recommendation that governing parties be modest about what they know is an exhortation, which a party may accept in principle and disregard in practice. The partition states that for one region of the condition of a governed system the question of estimation does not arise, so that a claim about that region derives from something other than the evidence, and the claim can be identified as such.

The partition has two boundaries, and their difference is consequential.

Claim 11.2. (The inner and the outer boundary). The unobservable component has an inner boundary fixed by what no available instrument reaches, and an outer boundary fixed by the measurements a governing arrangement has chosen to institute. The region between the two is unobservable by a decision that was taken, and is recoverable by instituting a measurement. The outer boundary is therefore determined by the party whose measurement design Claim 8.1 identified as interested in the result.

The claim links the two failures of Part III. A party may enlarge the unobservable region by declining to measure, and a claim that a matter is unknowable may record a fact about the world or a fact about what was instituted. Distinguishing the two requires establishing what a measurement would have reached, which is an argument a party asserting unknowability may be required to make.

The second presupposition fails for the reasons recorded in Section 10. Governance possesses no dynamical model with which to advance an estimate between observations, and where a model is used it belongs to the class whose uncertainty Section 10 found unrepresentable.

The third presupposition fails in a direction with a documented pathology.

Claim 11.3. (Hardening of the estimate under understated model error). Where the error attaching to the forecast is specified as smaller than it is, the cycle weights the forecast above the arriving observation, the estimate departs from the system, and the departure is self-sustaining, since a confident estimate assigns little weight to the observations that would correct it. The engineering literature calls this divergence. Its institutional counterpart is an arrangement whose account of a situation has hardened, which receives discrepant indicators and accommodates them within the account it holds, and whose confidence in that account is a cause of its persistence.

The counterpart is worth stating carefully, since a metaphor drawn from a filter would be an ornament. What transfers is the structural relation: confidence in a model, expressed as the weight given to it against arriving evidence, is a parameter of the arrangement, and setting it too high produces a specific failure in which evidence arrives and changes nothing. The failure is diagnosable by the same means in both cases, and the means is the subject of the residue.

11.5 The Transferable Residue

The element yields one requirement, and it costs no apparatus because it discards the estimator and retains only the comparison.

Requirement 11.4. (Monitoring of the residual). An arrangement holding an account of a governed system records, for each arriving indicator, the discrepancy between the value the account anticipated and the value received, and retains the sequence of such discrepancies. A sequence of discrepancies of consistent sign, or of growing magnitude, is evidence that the account has departed from the system, and it is evidence of that without requiring the account to be good and without supplying a corrected estimate.

Three features recommend the requirement. It presupposes no dynamical model, since the anticipated value may come from the account a party already holds, and Requirement 9.2 has already obliged that party to state it. It presupposes no observability, since the discrepancy is computed on quantities that are by construction measured. And it yields a restraint trigger in place of an estimate: a persistent discrepancy establishes that the account is wrong without establishing what is true, and the appropriate response to knowing that one's account is wrong and having no better one is the reduction of intervention intensity and not its

redirection.

What fails to survive is the estimator. A governing arrangement is not in a position to maintain a corrected running estimate of the condition of the system it governs, and Proposition 11.1 establishes that for one region of that condition no such estimate exists to be maintained.

Assignment 11.5. (Class of the element). Data assimilation falls in class i, precondition absent, on the ground of Proposition 11.1, with a secondary assignment in class iii on the ground of Claim 11.2, since the outer boundary of the unobservable region is set by an interested party, and a further assignment in class iv for Requirement 11.4.

12 Retrodiction and Cross-Epoch Consistency

This section examines the last element of the first family, the practice of requiring evidence drawn from separated epochs to cohere through the dynamics connecting them. Its objective is to identify the governance practice occupying the corresponding position, which is the pooling of studies conducted at different times into a single estimate, and to state what the astrophysical discipline would require of that practice. The section follows the template fixed in Section 7. Its result differs in character from those of Sections 8 to 11: the element transfers, with a modification, and its transfer bears directly on an established procedure.

12.1 The Structure of the Practice in Astrophysics

Cosmology holds evidence from widely separated epochs. The microwave background records conditions in the early universe; galaxy surveys record the distribution of structure at later epochs and at the present. These bodies of evidence are not averaged. They are required to cohere: a condition inferred from the early evidence is evolved forward through the theory of structure formation and confronted with the condition inferred independently from the later evidence,

$$F_{t_2-t_1}(\hat{x}(t_1)) \stackrel{?}{=} \hat{x}(t_2), \quad (14)$$

within the uncertainties attaching to each. Disagreement indicts the model, the parameters or the assumptions, and the location of the fault becomes the research question.

The comparison of Equation 14 treats each epoch as a separate constraint on one evolving system. The evidence from an early epoch is evidence about that epoch and constrains the present only through a stated dynamics. Two bodies of evidence from different epochs that disagree about a quantity are not thereby in conflict, since the quantity may have changed, and establishing whether they conflict requires the propagation.

12.2 Presuppositions of the Consistency Check

Two conditions are required.

A dynamics is available through which evidence from one epoch can be projected onto another, so that agreement and disagreement are well defined.

The system is one system. The evidence from separated epochs concerns the same object at different stages, and the comparison is diachronic.

The first admits failure where no dynamics is available. The second admits failure where the bodies of evidence concern different objects, in which case their comparison is a comparison across cases and belongs to the problem of contextual transport treated in §5.2.

12.3 The Governance Counterpart

The counterpart is the synthesis of a body of studies into a single estimate, which is the central instrument by which evidence is prepared for policy use. A systematic review identifies studies meeting stated criteria and a meta-analysis combines their estimates, weighting each by its precision. Variation among the estimates is assessed under the heading of inconsistency, one of the five domains recorded in Section 2, and where the variation exceeds what sampling explains, a random-effects specification represents the excess as a distribution of true effects from which the studies are drawn.

The temporal structure of this operation is worth stating explicitly, since it is ordinarily left implicit. A synthesis holds estimates $\{\hat{\theta}_k\}$ produced by studies with reference times $\{t_{e,k}\}$ spanning sometimes two or three decades, and combines them as draws from a common distribution, which is to say under the assumption $\theta(t_{e,k}) = \theta$ for every k . The operation therefore treats the estimand as time-invariant, and the treatment is a substantive assumption carried by the procedure.

Claim 12.1. (Pooling across epochs under an invariance assumption). Meta-analytic synthesis combines studies conducted at different times as samples informative about one quantity, and thereby assumes that the quantity did not change over the period the studies span. Where the assumption holds, the synthesis gains precision. Where it fails, the synthesis produces an estimate of an average over a period in which the system moved, and reports that estimate as the current value of a parameter.

12.4 Status of the Presuppositions in Governance

The first presupposition fails, for the reasons established in Sections 10 and 11. No dynamics is available through which an estimate obtained in one decade may be projected onto another, so the astrophysical operation of propagating an early constraint forward and confronting it with a later one has no direct counterpart.

The second holds, and its holding is what makes this element transfer. Studies conducted at different times in one jurisdiction, on one population, of one intervention, concern one system at different stages. The comparison is diachronic in exactly the sense the astrophysical practice requires, and the material for a consistency check is therefore present even though the propagation is unavailable.

What follows is a modification of the astrophysical procedure in a direction the absence of dynamics permits. Where a dynamics is available, the consistency check is a confrontation between a propagated prediction and a later observation. Where none is available, the check

reduces to an examination of whether the estimates ordered by study date exhibit structure, and the examination is conducted without a model.

12.5 The Transferable Residue

The element yields a requirement, and the requirement bears on an established procedure.

Requirement 12.2. (Temporal consistency in place of temporal pooling). A synthesis combining studies conducted over an extended period reports the estimates ordered by the reference time of each study, examines whether the ordered sequence exhibits trend, and treats a trend as a finding about the system in place of a nuisance to be absorbed into a variance component. Where trend is present, the synthesis reports the recent estimates separately from the pooled estimate, and the burden of showing that pooling is warranted lies with the party pooling.

The instruments for such an examination exist and require no development. Cumulative synthesis, in which the estimate is recomputed as each study is added in date order, displays the trajectory directly. Meta-regression on the year of the study tests it. What the requirement changes is the interpretation: these instruments are ordinarily deployed to detect a bias, and the requirement directs that a trend be entertained as a description of a moving system before it is assigned to bias.

The change of interpretation is where the paper's claim meets a substantial existing literature, and the meeting must be handled with care. Effect estimates in clinical research frequently decline as evidence accumulates, and highly cited studies reporting strong effects are often followed by studies reporting weaker ones or none [23]. The established readings of this pattern attribute it to features of the research process: selective publication, selection of favourable results, the greater precision of later and larger studies, and regression toward a true value from an initially exaggerated one. Those readings are supported and the present argument does not displace them.

Claim 12.3. (Underdetermination of the temporal trend). A trend in effect estimates ordered by study date admits two classes of explanation. The estimates may have converged on a stable quantity from an initially distorted position, in which case the trend is a fact about the research process. Or the quantity itself may have moved, in which case the trend is a fact about the system and the later estimates are correct for their own epoch while the earlier ones were correct for theirs. The two classes are not distinguished by the trend alone, and the established literature has developed instruments for the first while leaving the second largely unexamined.

The claim states a research problem and supplies no solution, which is the honest position. Three observations bear on how the problem might be approached, and each is offered as a direction. Distortions of the research process are expected to operate uniformly across subject matters, while genuine movement is expected to be concentrated in domains whose underlying systems change quickly, so a comparison across domains of differing rate carries information. A distorted initial estimate is expected to converge and remain, while a moving system may produce a trend that continues or reverses. And a change in a system's condition

is frequently traceable to a datable event, so the coincidence of an inflection in an estimate sequence with an identifiable change in the setting is evidence of the second class.

The applied consequence stands independently of the resolution of Claim 12.3, and it is worth stating in the strongest available form. A synthesis that pools across decades without examining the date ordering has assumed the answer to a question it has not asked. The examination costs nothing, since the reference dates are recorded and the instruments exist, and the assumption it would test is one on which the standing of the synthesis depends.

Assignment 12.4. (Class of the element). Retrodiction with consistency checking falls in class iv, transfers intact, with the modification recorded above, since the propagation is unavailable and the examination of the ordered sequence stands in its place. A secondary assignment in class i records the unavailability of the dynamics through which a full consistency check would run.

13 The Past Light Cone and the Temporal Heterogeneity of the Evidence Base

This section opens the second family, which concerns constraints imposed by the observing situation in place of methods applied to observations once obtained. Its subject is the structure of what an observer holds at a moment. Its objective is to establish that the evidence base available to a governing party has the form Definition 4.7 described, to identify what astrophysics does about the corresponding structure and why the remedy is unavailable here, and to derive the consequence that follows for the weighting of sources. The section follows the template fixed in Section 7. Its principal result is that a preference for recent evidence has a systematic direction, and that the direction runs against the variables carrying a system's approach to a transition.

13.1 The Structure of the Situation in Astrophysics

An observation event O has a past light cone $\mathcal{L}^-(O)$, and what is observable at that event lies on or within it. Objects at different distances intersect the cone at different epochs, so a survey assembles material from many epochs into one catalogue,

$$\{y_i = h(x_i(t - \tau_i))\}_i, \quad \tau_i \approx d_i/c, \quad (15)$$

and the catalogue depicts no moment.

Astrophysics handles this by carrying the epoch with the datum. A source is recorded with its redshift, from which the interval τ_i of Equation 15 follows, and no analysis treats two sources at different redshifts as contemporaneous. The heterochrony is fully represented in the data structure, and the discipline then exploits it: a catalogue spanning many epochs is a record of evolution, and cosmology's principal method of studying change over time is to observe many objects at many distances at once. Section 24 treats that exploitation and its conditions.

13.2 Presuppositions of the Handling

Two conditions are required.

The interval attaching to each datum is known. It follows from a measured distance and a constant of nature, is computed for every source, and is carried with the source through every subsequent analysis.

The intervals are commensurable. Two sources at different redshifts are located on a common time coordinate, so their separation in time is a determinate quantity and the collection is ordered.

Both admit failure. The first fails where the interval attaching to a datum is unrecorded or unknown. The second fails where the reference times of different sources are defined in ways that do not compose into a common ordering.

13.3 The Governance Counterpart

The evidence base of a governing party has the same structure and lacks the handling. A decision draws on national accounts referring to a quarter concluded some months earlier, a census conducted some years earlier, administrative returns of the preceding month, an academic literature describing fieldwork conducted years before publication, satellite observations of the preceding day, and transactional or platform data of the preceding minutes. Each item refers to a different moment, and their assembly into a briefing is an assembly of a catalogue whose entries lie at different depths.

The presentation ordinarily suppresses this. A dashboard displays indicators side by side, and the display asserts contemporaneity by its form. A briefing paragraph places a figure from one source next to a figure from another and joins them with a conjunction. The reader receives a description of a moment, and no moment is described.

13.4 Status of the Presuppositions in Governance

The first fails in practice. Reference periods are recorded in statistical documentation and are ordinarily discarded in transmission, so that by the time an indicator reaches a decision it travels as a value. Section 20 treats the remedy, which is available and inexpensive.

The second fails more deeply, and the failure is structural. Reference times in governance are defined by heterogeneous conventions: a period of accrual for a flow, an enumeration date for a stock, a fieldwork window for a survey, a date of recording for an administrative event, a date of publication for a study whose data collection is undated. These are not points on a common coordinate, and an evidence base assembled from them supports no determinate statement of how far apart its items are.

The consequence for the correspondence requirement is direct. Definition 4.2 defines the governance lookback time of an item, and where reference times are incommensurable the lookback profile of Definition 4.7 is partially undefined. A party cannot compute how stale its evidence base is, and can determine only that different parts of it are stale to different and partly unknown degrees.

13.5 The Transferable Residue

The handling does not transfer, and the situation yields a finding.

Claim 13.1. (Direction of the recency preference). Sources differ systematically in their production intervals, and the difference is associated with what they measure. Transactional, administrative and platform sources are fast and record quantities that are counted as they occur. Enumerations, cohort studies, ecological monitoring and periodic surveys are slow and record quantities that require a designed instrument to observe. A preference for recent evidence is therefore a preference among subject matters and not only among dates, and it selects against the second class.

The claim becomes consequential when joined to a feature of systems approaching a transition. The variables whose movement carries a system toward a stability boundary are ordinarily slow, and the fast variables record fluctuations about a condition the slow variables determine. So the sources that arrive quickly report the quantities that move quickly, and the quantities whose movement matters for whether the system is near a boundary are reported by the sources that arrive late.

Claim 13.2. (Blindness of a recency-weighted base). An evidence base weighted toward recency is weighted toward fast variables, and is therefore weighted away from the variables carrying a system's approach to a transition. Recency accordingly fails as an unqualified criterion of evidential quality, since improving a base by that criterion degrades it with respect to the class of question on which the cost of error is greatest.

The claim requires two restrictions. It concerns the composition of a base and not the quality of any item, and a recent item is better than a stale one where both measure the same quantity. And it presupposes the association between speed of production and character of subject matter, which is an empirical generalisation admitting exceptions and worth examining in particular evidence systems.

What survives as a requirement is modest and follows from the first presupposition alone: an evidence base states the reference time of each item and states it in a form permitting comparison, and a display presenting items of different reference times marks the difference. Section 20 develops this, and Section 16 treats a further consequence of the composition of a base that Claim 13.1 leaves untouched.

Assignment 13.3. (Class of the element). The past light cone falls in class I, precondition absent, on the ground that the intervals attaching to governance data are neither reliably known nor commensurable, with a secondary assignment in class IV for the recording and display of reference times.

14 The Horizon and the Limits of Observability in Principle

This section examines a feature of the observing situation distinct from the delay treated in Section 13, namely that some part of what exists lies outside the reach of observation as a matter of the situation's geometry. Its objective is to establish that governance possesses a

counterpart of this feature, to distinguish that counterpart from the lag with which it is easily conflated, and to give the partition of Proposition 11.1 its in-principle form. The section follows the template fixed in Section 7 and is the shortest in Part IV, since its content is a single distinction with a consequence. Its result falls in the class of elements that transfer without impediment.

14.1 The Structure of the Situation in Astrophysics

The past light cone bounds what an observer can have received. Regions of the universe lie outside it, and their exclusion follows from the geometry of the situation and from the finite speed of propagation. No improvement in instruments reaches them, since the improvement affects what is done with arriving signals and creates no signal that has not arrived.

Two features distinguish this from lookback. Lookback concerns material that has arrived and carries an epoch; a horizon concerns material that has not arrived at all. And lookback admits partial remedy by waiting, since a later observation reaches a later epoch of the same source, while a horizon recedes on its own schedule and is unaffected by the observer's diligence.

Astrophysics states its horizons explicitly and reasons about what lies beyond them by inference from theory, marking such reasoning as extrapolation of a different standing from the description of what has been observed.

14.2 Presuppositions of the Handling

One condition is required, and it is a condition on the observer.

The location of the horizon is determinable. An observer can say which material lies inside the cone and which outside, so that a claim about the exterior is identifiable as a claim about the exterior and is assessed accordingly.

The condition admits failure where the boundary is undetermined, in which case claims about the exterior circulate alongside descriptions of the interior and are received on the same footing.

14.3 The Governance Counterpart

The counterpart is the region identified by Proposition 11.1, on which the instituted measurements carry no information. Section 11 established the partition from the requirements of estimation. The present section supplies its in-principle form and distinguishes the two boundaries introduced in Claim 11.2.

Three kinds of content ordinarily occupy the region. There are quantities no instrument available to a governing arrangement reaches, of which the intentions and reasoning of parties are the standing example. There are quantities whose measurement would alter them, so that the measured value bears on a condition the measurement produced. And there are properties of a system's relational structure that admit no representation in the categories through which the system is recorded, of which the distance of a system from a stability boundary is the instance bearing most directly on the present argument.

14.4 Status of the Presupposition in Governance

The condition fails in practice and is satisfiable in principle, which is why the element transfers.

It fails in practice because the boundary is ordinarily undeclared. A briefing states what is known and passes over what the measurement system does not reach, and the passing over is silent. A reader receives an account whose interior and exterior are undistinguished, and a claim about the region beyond the measurements arrives with the same grammatical standing as a report of a measured quantity.

It is satisfiable in principle because the boundary is determinable by inspection of the measurement system. What a set of instituted measurements can and cannot bear on is a question about that set, answerable by examining it, and the answer requires no new observation.

14.5 The Transferable Residue

The element yields a requirement and a distinction, and both are available at the cost of declaration.

Requirement 14.1. (Declaration of the unobserved region). An account of the condition of a governed system offered in support of a decision states which of the quantities bearing on that decision the instituted measurements reach and which they do not, and marks claims concerning the second class as claims resting on grounds other than the evidence adduced.

Claim 14.2. (Standing of a claim beyond the boundary). A claim about a quantity outside the reach of the instituted measurements is not thereby false and is not thereby inadmissible. What follows from its position is that the evidence adduced supplies no support for it, so that its support lies in theory, in analogy, in the experience of the parties, or in interest, and the identification of which of these it rests on becomes possible once the position is declared.

The distinction between the two boundaries of Claim 11.2 acquires its practical form here. Where a quantity lies outside the reach of every available instrument, a declaration records a limit. Where it lies outside the reach of the instituted measurements while an available instrument would reach it, the declaration records a choice, and the choice is attributable to the party that made it. A governing arrangement that declares its unobserved region therefore exposes the composition of that region to examination, and the examination separates the two.

The requirement carries one consequence worth stating, since it bears against a party that adopts it. A declared boundary invites the enlargement of the measurement system, and Section 8 established that the measurement system is constituted by parties interested in the results it produces. The enlargement is therefore not neutral, and Section 16 identifies a specific direction in which such enlargements characteristically run.

Assignment 14.3. (Class of the element). The horizon falls in class iv, transfers intact, since the boundary is determinable by inspection of the measurement system and its declaration requires no apparatus. A secondary assignment in class iii records that the position of the outer boundary is fixed by an interested party, per Claim 11.2.

15 Selection at Depth and the Compounding of Age with Selectivity

This section examines a feature of observation at distance that has no counterpart in the categories of Section 2 and a close one in governance practice. Its objective is to establish that the selectivity of an evidence base varies with the age of its contents, and that the variation runs in a direction which compounds with the correspondence failure. The section follows the template fixed in Section 7. Its result is a claim that the two defects of an old body of evidence, its distance from the present condition and its unrepresentativeness of the period it describes, are associated and not independent.

15.1 The Structure of the Situation in Astrophysics

A survey complete to a limiting apparent brightness detects, at greater distance, only objects of greater intrinsic luminosity. Nearer the observer, objects of average and below-average luminosity enter the sample as well. The statistical properties of the detected sample therefore depend on distance, and the sample at depth is systematically unrepresentative of the population at that depth. The effect was described by Malmquist and is standard [32].

Its structure is what matters here. Depth in such a survey is lookback, so the same coordinate that measures how old an observation is measures how selected it is. The two properties are not independent features of a datum. They are one feature seen twice, and a correction for one is made in the same coordinate as a correction for the other.

Astrophysics corrects by characterising the selection. Where the luminosity distribution of the population and the detection threshold of the instrument are known, the selection function is computable, and the sample is reweighted or restricted to a volume within which detection is complete.

15.2 Presuppositions of the Correction

Two conditions are required.

The selection function is characterisable. What entered the sample and what failed to enter is determined by a threshold and a population distribution, both of which are available or estimable.

The selection operates on a stated dimension. Objects are excluded on a property that is identified, so that the exclusion is described and its consequence computed.

Both admit failure. The first fails where the process determining what entered the record is unknown. The second fails where exclusion proceeds on several properties at once, some of them unrecorded.

15.3 The Governance Counterpart

The counterpart is the dependence of an evidence base's composition on the age of its contents, and it operates through three channels.

What was recorded at a past time was determined by what the then-current measurement system counted, and measurement systems record what their designers considered consequential. An older body of evidence therefore represents an older judgement about what matters, and the categories absent from it are absent because they were not then salient.

What survives from a past time is a subset of what was recorded, and the subset is not random. Retention schedules, changes of custody, migrations of format and the ordinary attrition of archives remove material differentially, and material whose retention was advocated by an interested party survives at a higher rate.

What was published from a past time is a subset of what was found, and the selection operating on publication is documented and is one of the five domains recorded in Section 2. Its treatment there, however, is as a static property of a body of evidence, assessed once and recorded as a downgrade. Its variation with the age of the material is not among the considerations the domain directs an assessor to weigh.

15.4 Status of the Presuppositions in Governance

The first fails. The process determining what entered and remained in the historical record of a governed system is composite, is documented unevenly, and is in substantial part unrecoverable. There is no governance counterpart of a luminosity function and a detection threshold from which a selection function might be computed.

The second fails as well, and its failure is the more consequential. Exclusion from the historical record proceeds on several properties at once: on salience to the then-current measurement design, on the interests of parties advocating retention, on the accidents of custody and format, and on the publication practices of a period. These operate together, and the resulting composite is not a threshold on a stated dimension.

The claim that follows is the section's result.

Claim 15.1. (Association of age with selectivity). The selectivity of an evidence base varies with the age of its contents, and varies in the direction of greater selectivity for older material. An older body of evidence is therefore subject to two defects at once: its distance from the present condition of the system, and its unrepresentativeness of the period it purports to describe. The two are associated through the same coordinate, so a body of evidence that is stale is, for that reason, also more selected.

The association has a consequence for the operation examined in Section 12. A synthesis pooling studies across decades draws material whose selectivity varies systematically across the period pooled, and a trend in the pooled estimates ordered by date may therefore record the variation of the selection and not the movement of the system. Claim 12.3 stated the underdetermination of such a trend between the research process and the system's motion, and the present claim adds a specific mechanism to the first of those readings and identifies it as age-dependent. The two claims should be read together, and the direction they jointly indicate is that separating the readings requires evidence about the selection régime of each period and not about the estimates alone.

15.5 The Transferable Residue

The correction does not transfer. What transfers is the recognition that the defect exists and runs in a determinate direction, together with one requirement of a familiar kind.

Requirement 15.2. (Statement of the recording régime). Where a body of evidence spans an extended period, the account of it states what the measurement and publication régimes of each period recorded and omitted, so far as this is recoverable, and identifies the categories present in later material and absent from earlier material. A comparison across the period is conducted on the categories common to it, and a category present in one part alone is marked.

The requirement is weaker than the astrophysical correction and is what the failure of the two presuppositions permits. It supplies no reweighting and no computed adjustment. What it supplies is that a comparison across a period is conducted on material whose comparability has been examined, and that the examination is recorded where a reader may assess it.

One further observation belongs here, since it bears on the strength of the claim. The three channels of the governance counterpart differ in tractability. Publication selection is documented and partly correctable by established methods. Retention selection is examinable where archival practice is recorded. The selection operating through the categories of a past measurement system is the least tractable of the three, since recovering what a system omitted requires a description of the period from outside that system, and such descriptions are themselves subject to the same condition. The claim of this section is therefore strongest for the third channel and weakest for the first.

Assignment 15.3. (Class of the element). Selection at depth falls in class I, precondition absent, on the ground that the selection function is uncharacterisable and the exclusion proceeds on unstated composite properties, with a secondary assignment in class IV for Requirement 15.2.

16 Survey Cadence, Aliasing, and the Fabrication of Trend

This section examines the constraint that a schedule of observation places on what can be observed at all. Its objective is to establish that the release schedule of a statistical system fixes, in advance of any analysis, the class of phenomena that can appear in evidence, and that variation faster than the schedule resolves is registered as slow variation of a kind the system does not have. The section follows the template fixed in Section 7. Its result is the most directly demonstrable finding in Part IV, since the relevant limit is computable from the schedule alone, and it pairs with Claim 13.2 to give the composition of an evidence base a defect at each end.

16.1 The Structure of the Constraint in Astrophysics

A survey observes on a cadence: a sampling interval, a total baseline, and a pattern of gaps imposed by the seasons, the weather and the allocation of instruments. The cadence determines

what the survey can discover. A source varying on a timescale shorter than the sampling interval is recorded at scattered phases of its variation, and the variation is unavailable to the survey as variation.

The formal statement of the limit is standard. A signal sampled at regular intervals of length Δ carries information about a component of frequency f only where

$$f < \frac{1}{2\Delta}, \quad (16)$$

so that variation with periods shorter than 2Δ is indistinguishable from variation at a longer period that the sampling does resolve [40]. The shorter variation is registered, and is registered as something else. Astronomical practice treats the consequence explicitly, since irregular sampling produces a window function whose structure imprints spurious periodicities on the recovered spectrum, and the identification of a periodicity requires the window structure to be examined before the periodicity is claimed.

Survey design accordingly proceeds from a statement of the phenomena the survey is intended to detect, and the cadence is chosen against that statement. A survey that will detect events of a given duration is built to sample more finely than that duration, and its inability to detect faster events is a designed and recorded property.

16.2 Presuppositions of the Handling

Two conditions are required.

The cadence is set against a stated class of target phenomena, so that what the design excludes is excluded knowingly and the exclusion is recorded with the survey.

The sampling structure is known well enough that the limit and the aliasing pattern are computable, so an apparent variation can be tested against the possibility that it is an artefact of the schedule.

The first admits failure where the schedule is fixed by considerations unrelated to the timescales of the phenomena. The second admits failure where the schedule is unrecorded or irregular in ways that are undocumented.

16.3 The Governance Counterpart

The counterpart is the release schedule of a statistical system. National accounts appear quarterly, labour force statistics monthly or quarterly, an enumeration decennially, administrative returns on an annual cycle tied to a fiscal year, and inspection regimes on a cycle set by resources. The schedule is the sampling interval at which a governed system is observed, and it fixes the temporal resolution of everything a governing party can subsequently learn from that source.

The correspondence is exact and the practice differs at both presuppositions.

16.4 Status of the Presuppositions in Governance

The first fails generally. Release schedules are set by administrative convention, by the fiscal calendar, by the cost of collection and by the requirements of legislation, and they are seldom set against a stated hypothesis about the timescales on which the observed phenomena vary. A quantity varying on a scale of weeks and reported annually is reported annually because annual reporting is the practice, and the mismatch is not a designed exclusion because no design was conducted against the question.

The second, by contrast, is satisfied and unused. The release schedule is documented, so the limit follows from it by computation, and the pattern of gaps and revisions is recorded in the statistical documentation of every serious system. The information required to state what a series cannot show is available and is ordinarily left unstated.

Claim 16.1. (Determination of the observable class by the schedule). The release schedule of a source fixes, before any analysis, the class of phenomena that can appear in evidence drawn from it. A phenomenon whose characteristic timescale violates Equation 16 for that source cannot appear in it as itself. The class of policy problems that can be evidenced from a statistical system is therefore determined in part by the schedule of that system, and the determination is made by parties setting schedules for administrative reasons.

Claim 16.2. (Fabrication of slow structure from fast variation). Variation faster than a source resolves is registered by that source as variation at a longer period. A governing party observing an annual series of a quantity that varies on a scale of months may therefore observe a trend, a cycle or a level shift that the underlying quantity does not exhibit, and the artefact is produced by the schedule and not by any error of collection, processing or analysis. The appearance is indistinguishable, within the source, from the phenomenon it imitates.

The second claim bears directly on the correspondence requirement, and the bearing is worth drawing out. Section 4 treated correspondence failure as a matter of the system having moved between the reference time and the decision. The present claim identifies a case in which the source misrepresents the movement itself, so that a party reasoning carefully about lookback may still hold a picture of the system's trajectory that the schedule manufactured. Correction requires observation on a finer cadence and is unavailable within the source.

16.5 The Transferable Residue

The element yields a requirement whose satisfaction requires arithmetic and no apparatus, and it yields a joint finding with Section 13.

Requirement 16.3. (Statement of the resolution limit). A source used in evidence states the shortest timescale of variation it can represent, computed from its sampling interval, and an inference drawn about variation faster than that limit is marked as unsupported by the source. Where a series exhibits structure whose period is close to a multiple or a submultiple of the sampling interval, the possibility that the structure is an artefact of the schedule is examined before the structure is interpreted.

Claim 16.4. (Defect at each end of the composition). Claim 13.2 established that an evidence base weighted toward recency is weighted away from slow variables, since slow variables are measured by sources with long production intervals. Claim 16.1 establishes that sources with long sampling intervals cannot represent fast variation. Production interval and sampling interval are both set by administrative convention and are associated with one another, since the sources that are expensive to collect are collected infrequently and released late. The variables that carry a system toward a stability boundary are therefore observed by sources that are simultaneously the latest to arrive and the coarsest in resolution, and the two defects fall on the same variables.

Claim 16.4 states the compound and requires the same restriction as its components. The association between production interval and sampling interval is an empirical generalisation about statistical systems as they are ordinarily constituted, and it is contingent. A system could collect a slow variable frequently and release it quickly, and the observation that existing systems do otherwise is an observation about resource allocation and administrative practice. This is the point at which the finding becomes a claim about institutions and admits remedy by their redesign, and Section 28 takes it up in that form.

Assignment 16.5. (Class of the element). Survey cadence falls in class iv, transfers intact, since the resolution limit is computable from documented schedules and its statement requires no apparatus. A secondary assignment in class iii records that the schedules themselves are set by parties whose conduct the resulting series are used to assess, per Claim 8.1.

17 Single Realisation and the Unavailability of an Ergodic Substitute

This section examines the condition of holding one instance of the object studied. Its objective is to identify the substitution by which cosmology partly escapes that condition, to establish what the substitution requires, and to show that the governance counterpart of the substitution imports the difficulty treated in §5.2. The section is short and hands its result forward, since the consequence of the failure is the subject of Section 24. It follows the template fixed in Section 7.

17.1 The Structure of the Condition in Astrophysics

Cosmology studies one universe. A theory of the early universe predicts a distribution of possible outcomes, and one draw from that distribution is available for observation. The resulting uncertainty is irreducible by improvement in instruments, since it arises from the number of realisations and not from the quality of the measurement, and it is reported as a floor below which no observation of the accessible volume can reduce the uncertainty about a theoretical parameter.

The discipline escapes this condition in part by a substitution. Under the assumption that the universe is statistically homogeneous at sufficient scale, widely separated regions are treated as approximately independent samples from one distribution, so that averaging over

space substitutes for averaging over an unavailable ensemble of universes. The substitution is what converts a single realisation into a large sample, and it is an assumption with content: it can fail, its failure has observational signatures, and the scale at which it holds is itself an object of investigation.

17.2 Presuppositions of the Substitution

Two conditions are required.

The parts are drawn from a common distribution, so that the variation observed across them is the variation the theory predicts across realisations.

The parts are approximately independent, so that averaging across them reduces the uncertainty in the manner the count of parts suggests.

The first fails where the parts differ systematically. The second fails where they influence one another, in which case the effective number of independent samples is smaller than the count and the reduction in uncertainty is overstated.

17.3 The Governance Counterpart

Governance holds one history of the system it governs. The counterpart of the cosmological substitution is comparison across cases: jurisdictions, regions, organisations or populations, treated as instances from which a common regularity may be recovered. The substitution performs the same function, converting a single unrepeatable trajectory into a sample.

17.4 Status of the Presuppositions in Governance

Both fail, and they fail for reasons that the literature reviewed in §5.2 has established independently of the present argument.

The first fails because jurisdictions differ in the support factors on which an intervention's operation depends, which is the substance of the objection Cartwright and Hardie [4] raise against the transport of findings across settings. The variation observed across cases is therefore a compound of the variation a common regularity would produce and the variation produced by the differences among the settings, and the two are separated only by a theory of what the relevant differences are.

The second fails because the cases influence one another. Jurisdictions imitate one another's policies, are subject to common shocks, and are connected by flows that transmit conditions across their boundaries. The effective number of independent instances is accordingly smaller than the count, and frequently much smaller.

Claim 17.1. (Unavailability of the ergodic substitute). The substitution by which cosmology converts a single realisation into a sample requires that the parts be drawn from a common distribution and be approximately independent. Governance possesses no comparable substitution, since the cases available for comparison are structurally heterogeneous and mutually influencing. A governed system is therefore observed in one realisation, and the uncertainty

attaching to a claim about its condition contains a component that no accumulation of observation within that system reduces.

17.5 The Transferable Residue

The substitution does not transfer and the recognition of the floor does.

Claim 17.2. (An irreducible component of the uncertainty). Part of the uncertainty attaching to a claim about the condition of a governed system arises from the system having one history, and that part is unaffected by the volume, the frequency or the quality of observation. It is distinct from the component treated in Section 10, which grows with the interval, and from the component treated in Section 11, which is a matter of what the measurements reach. It is present at zero interval and under complete observability.

The claim is offered as a restriction on what improved evidence can deliver, and its practical force is comparative. A programme of measurement promising to resolve a question about a governed system's condition is subject to a floor whose height depends on how much of the variation in that condition is attributable to the singularity of its history, and a promise to resolve the question below that floor is a promise that no measurement discharges.

What remains is the trade between the two escapes, and it is the subject of Section 24. Astrophysics has both a time series of one object and an ensemble of objects at many epochs, and it uses the second to interpret the first. Governance has the time series and lacks the ensemble, and its attempts to construct one exchange the temporal difficulty for the contextual one. That exchange is the paper's principal negative result and is stated there.

Assignment 17.3. (Class of the element). Single realisation falls in class 1, precondition absent, since the homogeneity and independence the substitution requires are unavailable, with no secondary assignment. The recognition recorded in Claim 17.2 is a consequence of the failure and not a residue of the practice.

18 Standard Rulers and the Drift of Measurement Definitions

This section closes the second family with the element that bears most directly on the third mode of correspondence failure identified at Definition 4.4. Its objective is to establish what a comparison across epochs requires of the quantities compared, and to show that governance measurement lacks that property in a way that its records rarely mark. The section follows the template fixed in Section 7. Its result is a claim about the composition of longitudinal evidence and a requirement whose cost is documentation.

18.1 The Structure of the Practice in Astrophysics

Comparison across epochs requires a quantity whose meaning is stable across them. Cosmology obtains this from features of known physical extent whose size at a given epoch is fixed by theory in place of convention. The acoustic scale imprinted on the distribution of matter

supplies the standing example, and its detection in the correlation function of a galaxy sample furnishes a length against which distances at different epochs are calibrated [12].

The property that makes such a feature useful is worth isolating. Its extent is determined by physics that the observer does not administer, so that its use as a reference requires no agreement among observers and admits no revision by any of them. Where a reference of this kind is available, quantities measured at separated epochs are commensurable, and a change in a measured value is attributable to the system.

18.2 Presuppositions of the Calibration

Two conditions are required.

The reference is invariant across the epochs compared, or its variation is known and corrected.

The reference is administered by no party to the comparison, so that its stability requires no agreement and is not subject to revision in the interest of a result.

The first fails where the reference changes. The second fails where the reference is a convention maintained by parties, in which case its stability depends on their forbearance and its revisions carry the interests that prompted them.

18.3 The Governance Counterpart

Governance possesses no reference of this kind, and its counterpart is the definition. Whether a person is unemployed, whether a household is in poverty, whether an event is a recorded crime of a given class, whether a condition is a diagnosis, and whether an emission falls within an inventory boundary are all settled by definitions that are written, published and periodically revised.

Such definitions perform the office of the standard ruler: they fix what a measured value means so that values at different times may be compared. They differ from an acoustic scale in the respect the second presupposition identifies. A definition is administered, its stability is a choice, and its revision is an act with a date and an author.

18.4 Status of the Presuppositions in Governance

The first fails routinely and the failures are documented, since a statistical agency revising a definition ordinarily records the revision and frequently publishes a bridging estimate for the period of overlap. Unemployment definitions have been revised with changes in the treatment of discouraged workers and of marginal attachment; poverty measures have moved between absolute and relative constructions and between income and consumption bases; diagnostic criteria have been revised in ways that alter recorded prevalence with no change in the underlying condition; and classification rules for recorded offences have been revised in ways that alter recorded volumes. Each revision is a change in the ruler.

The second fails structurally, and its failure is the one this paper adds to the familiar observation about the first. The definitions are administered by the parties whose conduct the

resulting series are used to assess, which is Claim 8.1 in its longitudinal form. A revision of a definition is therefore an act available to a party with an interest in the level or the trend of a series, and the availability holds whether or not it is exercised.

Claim 18.1. (Attribution under a revised definition). Where a definition is revised between two measurements, a difference between the measured values is a compound of the movement of the system and the change in the ruler, and the two are separated only by a bridging estimate. Where a bridging estimate is absent, the difference is unattributable. Where a bridging estimate is present, it is constructed by the party administering the definition.

The claim completes the treatment of the mode of failure named in articulation at Definition 4.4. Failure in state concerns a system that has moved. Failure in dynamics concerns relations that have altered. Failure in articulation concerns the categories through which both are recorded, and it is the mode that defeats the remedy available for the other two, since observing more recently corrects a stale value and supplies no bridge across a change of definition.

A further consequence bears on the interval itself, and it was anticipated in Section 8. The longer the interval between the reference time of a body of evidence and the moment of decision, the greater the chance that a definition has been revised within it. Correspondence failure in articulation therefore grows with τ_G , and grows in a manner that a reader who attends only to the age of the evidence will not detect, since the value retains its name.

18.5 The Transferable Residue

The element transfers as a requirement and withholds the property that makes the astrophysical practice work.

Requirement 18.2. (Statement of definitional vintage). A series presented in support of a claim about change over time states the definitions in force in each period it spans, marks the points at which a definition was revised, and states whether a bridging estimate exists. A comparison across a revision is conducted on a bridged basis or is marked as spanning a change of definition, and an inference about the movement of the system across such a point states which portion of the difference the bridging estimate attributes to the revision.

The requirement is satisfiable from documentation that statistical agencies already produce, and its cost is the transmission of that documentation to the point of use. The pattern is by now familiar from Sections 13 and 16: the information required to qualify an inference exists in the statistical system and is discarded in transmission to the decision.

What does not transfer is the invariance itself. Governance has no quantity whose meaning is fixed by something no party administers, and the search for one is unpromising, since the categories of a governed system are constituted by the same social processes the system governs. A definition of unemployment answerable to no administrative choice would be a definition answerable to no purpose. The absence is therefore a feature of the subject matter and admits no remedy of the kind the astrophysical case supplies.

Assignment 18.3. (Class of the element). Standard rulers fall in class III, interested observer, on the ground that the reference is administered by parties to the comparison, with a sec-

ondary assignment in class iv for Requirement 18.2 and a further secondary assignment in class i for the unavailability of any party-independent reference.

19 Blind Analysis and the Interested Observer

This section opens the third family, which concerns the disciplines an observing community imposes on itself. The family differs from the two preceding it in a respect that determines its results: its elements are practices adopted by a community and sustained by its norms, so the question of transfer concerns what a governing arrangement would have to accept and not what a governed system would have to be. Its objective for the present section is to identify the practice by which physics separates an analyst from the result, to establish that the corresponding separation is available in governance, and to state the conditions under which it is available. The section follows the template fixed in Section 7. Its result falls in the class of elements that transfer, with a restriction on the range of cases.

19.1 The Structure of the Practice in Physics

A blind analysis withholds the result from the analyst until the procedure that produces it has been fixed. The concealment is arranged in several ways: the region of data containing the signal is hidden while selection criteria are settled; an unknown offset is added to the quantity being measured and removed once the analysis is complete; or the analysis is developed on simulated data and applied to the real data once. The procedure is then run, and the result is accepted whatever it is [27].

The rationale is a claim about a mechanism and not about the honesty of analysts. An analyst who can see the result while choosing among admissible analytical decisions will tend, without any intention to deceive, to stop adjusting when the result looks as expected and to continue when it does not. The record of measurements of physical constants over the twentieth century supplies the evidence that the mechanism operates. Blinding removes the mechanism by removing the information on which it runs.

19.2 Presuppositions of the Practice

Three conditions are required.

The analytical procedure can be fixed before the result is seen, which requires that the decisions constituting the analysis be enumerable in advance.

The result can be concealed from the analyst, which requires that it be unavailable through channels outside the analysis.

The community accepts the result the fixed procedure produces, which requires that no party hold the authority to set the procedure aside once the result is known.

The first fails where the analysis is exploratory and its decisions arise from the data. The second fails where the result is known independently. The third fails where the analysis is conducted for a party with the standing to reject it.

19.3 The Governance Counterpart

The counterpart is the analysis of evidence bearing on a policy whose outcome a party has an interest in. The mechanism described above operates here with greater force, since the analyst is frequently employed by a party to the question, the analytical decisions are numerous, and the direction in which the result is preferred is known.

Two established practices occupy positions adjacent to blinding. Pre-registration fixes the analytical procedure in advance and lodges it publicly, and it is now standard in several research fields. Evaluation conducted by a body separate from the implementing department separates the analyst from the interested party institutionally in place of informationally.

19.4 Status of the Presuppositions in Governance

The first is satisfiable for a defined and substantial class. Where a question is stated in advance, an indicator identified and a comparison specified, the analytical decisions are enumerable and the procedure can be lodged before the data are examined. The class includes the assessment of a programme against a stated target, which is the setting the review sheet described in §2.3 is built for.

The first fails for the complementary class. Where the object is to characterise a situation whose relevant features are unknown at the outset, the analysis is exploratory by construction, and a procedure fixed in advance would fix the wrong procedure. The restriction is genuine and the section states it here in place of qualifying the result later.

The second is satisfiable with effort and is frequently violated in fact. A governing arrangement ordinarily holds partial knowledge of the result before the analysis concludes, through the operational reporting that runs alongside the statistical system. Concealment is therefore harder here than in a physics experiment, and it is achievable by separating the analytical function from the operational one.

The third is where the practice meets its principal obstacle. Acceptance of a result the fixed procedure produces requires that no party set the analysis aside once its direction is known, and a governing arrangement ordinarily retains the authority to commission a further analysis. The authority is legitimate, since a policy question is a political question and no analytical procedure settles it. What the authority permits is the selection of a result from a sequence of analyses, which reinstates the mechanism blinding was adopted to remove.

Claim 19.1. (Displacement of the selection). Where a party retains the authority to commission further analyses after seeing a result, the selection that blinding removes from within an analysis reappears in the choice among analyses. Blinding a single analysis is therefore effective only where the set of analyses commissioned is itself fixed in advance or recorded in full.

19.5 The Transferable Residue

The element transfers, with a restriction of range and an addition that the governance setting requires.

Requirement 19.2. (Fixed procedure and complete record). Where a question is stated in advance and its analysis is enumerable, the analytical procedure is lodged before the evidence is examined, and the result the procedure produces is reported whatever its direction. Where further analyses are commissioned after a result is known, the existence and the specification of each is recorded, so that the sequence of analyses is inspectable together with the result selected from it.

The addition in the second sentence is what Claim 19.1 requires, and it is the point at which the governance version departs from the physics one. Physics blinds an analysis. Governance must blind an analysis and record the population of analyses, since the second selection is available in governance and absent in the experimental setting.

The requirement costs no apparatus and costs something else, which should be stated plainly. It constrains a party at the moment the constraint bites, since the record of a commissioned and unused analysis is a record of a result the party preferred to set aside. A requirement of this kind is adopted by an arrangement that has decided to bind itself, and Section 29 records the paper's position on how much weight such requirements bear.

Assignment 19.3. (Class of the element). Blind analysis falls in class iv, transfers intact, for the class of questions statable in advance, with the addition recorded in Requirement 19.2. A secondary assignment in class i records the exploratory class, for which the first presupposition fails and the practice is unavailable.

20 Provenance, Re-Reduction, and the Time-Indexed Evidence Record

This section examines the practices by which an observing community preserves the conditions under which its observations were produced, and the operation those practices make possible. Its objective is to establish that governance discards at the point of transmission the information that three earlier sections found it would need, and to state the record structure that would retain it. The section follows the template fixed in Section 7. Its result is the requirement toward which Sections 13, 16 and 18 have each pointed, and it collects them.

20.1 The Structure of the Practice in Astrophysics

Observations are retained in a form independent of the processing applied to them, together with a record of the conditions of their production: the state of the instrument, the calibration in force, the version of the reduction pipeline, the observing conditions, and the selection criteria applied at each stage. Data are released in numbered versions, and a released version is identified by its number so that a result may be attributed to the version on which it was computed.

The operation this makes possible is re-reduction. An archive processed under a later calibration yields values differing from those first published, and the difference is attributable to the change in processing because the earlier processing was recorded. Observations decades

old are re-analysed and yield results their original analysts could not obtain, and the possibility rests entirely on retention: what was discarded at the time cannot be recovered by any later effort.

20.2 Presuppositions of the Practice

Two conditions are required.

The raw observation is retained separately from its processed form, so that a later processing may be applied to the same material.

The conditions of production are recorded with the observation in sufficient detail that a later analyst can determine how the published value was obtained.

Both admit failure. The first fails where only processed values are kept. The second fails where the conditions are undocumented, or documented in a location from which they do not travel with the value.

20.3 The Governance Counterpart

Statistical systems satisfy both conditions at origin and to a substantial degree. Microdata are retained under statutory arrangements, methodological documentation is published, and revisions are recorded. One field has built the corresponding archive explicitly: the collection of successive vintages assembled by Croushore and Stark [8] preserves what was available at each date, which is precisely the retention that permits a past decision to be assessed on the evidence its makers held.

The failure occurs downstream. A value assembled into a briefing, a dashboard, a review sheet or a paragraph of policy justification travels without its reference period, without its release date, without its vintage, without the definition in force when it was produced, and without the sampling interval of its source. The information exists at the origin and is stripped in transmission, and by the time the value reaches the decision it is a number with a name.

20.4 Status of the Presuppositions in Governance

Both are satisfied at origin and both fail at the point of use, and the location of the failure is what makes this element transfer.

The failure is not a failure of the statistical system. An agency that documents its methodology, publishes its revision history and records its reference periods has discharged the requirement. What fails is the transmission chain between that documentation and the moment a value is used to justify a decision, and no party in that chain is ordinarily obliged to carry the indices forward.

The consequence is that the qualifications developed in the preceding sections are unavailable at the point where they would operate. A party cannot compute a lookback profile from values that have lost their reference periods, cannot state a resolution limit for a series whose sampling interval has not travelled with it, and cannot detect a comparison spanning a definitional revision when the definitions are absent from the record it holds.

20.5 The Transferable Residue

The element transfers as a record structure, and the structure collects the requirements of the three sections that pointed toward it.

Requirement 20.1. (The indexed evidence item). An item of evidence travels to the point of decision with its value accompanied by the reference time of the condition it describes; the release date and, where the series is revised, the vintage; the uncertainty attaching to it, labelled with the time whose condition that uncertainty describes; the sampling interval of its source; and the definition in force at the reference time, together with a note of any revision since. An item lacking these indices is used as a description of an unspecified moment.

The requirement collects Requirement 16.3, Requirement 18.2 and the recording obligation identified in Section 13, and adds the labelling that Claim 10.1 showed to be missing. It is stated as a property of an item in transmission and not as a property of a statistical system, since the statistical systems ordinarily hold what it asks for.

Three features of the requirement are worth stating.

Its cost falls on the transmission chain and not on collection. Nothing new is measured, and the expense is the discipline of carrying forward what has already been recorded.

Its violation is detectable by inspection. A briefing can be examined for whether its figures carry reference times, and the examination requires no expertise in the subject matter.

Its satisfaction is asymmetric in time. An index recorded at origin is cheap and an index reconstructed later is expensive or impossible, so the requirement bears most heavily on present practice for the benefit of future assessment. The re-reduction operation depends on this asymmetry, and its governance counterpart, the reassessment of a past decision on the evidence its makers held, is available only where the vintages were kept.

Claim 20.2. (The assessment of past decisions). A decision taken on evidence available at its date can be assessed for its reasonableness only where the evidence of that vintage is recoverable. Where only revised series survive, a past decision is assessed against a description of the past that its makers did not possess, and the assessment attributes to them an error they had no means to avoid or a prescience they did not exercise. Retention of vintages is therefore a condition of the fair assessment of governing conduct, and its absence systematically distorts that assessment in a direction determined by the direction of the revisions.

The claim connects the requirement to a matter beyond epistemology. Where governing parties are held to account for past decisions, the evidential basis of that accounting is a record whose retention is administered by the parties themselves, and the temporal asymmetry noted above gives an arrangement wishing to escape assessment a means that costs nothing and looks like ordinary housekeeping.

Assignment 20.3. (Class of the element). Provenance and re-reduction fall in class iv, transfers intact, since the information is held at origin and the requirement concerns its transmission. A secondary assignment in class iii records that retention is administered by parties whose past conduct the retained record would be used to assess, per Claim 20.2.

21 Mock Pipelines and the Testing of the Measurement System

This section examines the practice of testing an analysis by applying it to synthetic data whose true content is known. Its objective is to distinguish that practice from the more familiar one of simulating a model of the world, and to establish that its governance counterpart would render several of the findings of Part IV operational. The section follows the template fixed in Section 7. Its result is a requirement that converts claims about what a source cannot show from assertions into demonstrations.

21.1 The Structure of the Practice in Astrophysics

A mock catalogue is a synthetic universe generated under a stated model, from which a synthetic observation is produced by passing the synthetic universe through a simulation of the instrument and of the reduction pipeline: the selection function, the noise, the calibration, the cuts applied at each stage. The analysis is then run on the synthetic observation and its output compared with the known content of the synthetic universe.

The distinction that matters is between simulating the world and simulating the observation of the world. A simulation of the first kind tests whether a model produces a pattern. A simulation of the second kind tests whether the analytical apparatus recovers, from what the instrument would actually deliver, a feature the analyst knows to be present [7]. The second tests the measurement system, and a feature the apparatus fails to recover is a feature the apparatus cannot establish.

21.2 Presuppositions of the Practice

Two conditions are required.

The measurement system is simulable, so that the transformation from a condition of the world to a published value can be represented well enough for the test to be informative.

A ground truth can be posited, so that the output of the analysis has something to be compared against.

The first fails where the transformation is undocumented. The second fails where no statement of a condition of the world can be formulated independently of the measurement system that would record it.

21.3 The Governance Counterpart

The governance counterpart is the simulation of a statistical or administrative measurement system, and it stands in the same relation to the practice of policy simulation that the astrophysical mock catalogue stands to a cosmological simulation.

Policy analysis simulates models of governed systems: microsimulation of a tax and benefit system, agent models of a population, projections under scenarios. These simulate the world. What is seldom simulated is the observation of the world: the sampling frame, the response process, the definitional boundaries, the processing, the release schedule and the revision policy through which a condition becomes a published figure.

The nearest existing practice is the power calculation, which asks whether a study of a stated size would detect an effect of a stated magnitude. Its scope is the sampling stage alone. The practice under examination here extends the same question to the whole apparatus, including the definitional boundary that determines what counts as an instance and the schedule that determines when the instance is registered.

21.4 Status of the Presuppositions in Governance

Both are satisfiable, and their satisfaction is what places this element in the transferring class.

The first is satisfiable because statistical systems document their transformations. A sampling frame, a response model, a set of definitional rules, a processing sequence and a release schedule can be represented in a simulation of the measurement system, and the representation need be no more faithful than the test requires.

The second is satisfiable in the specific form the test needs. The requirement is a posited condition of the world stated independently of the instrument, which is available wherever a policy question concerns a change of stated size in a stated quantity. A synthetic population in which a defined shift has been imposed supplies the ground truth, and no metaphysical claim about independent access to the world is involved.

21.5 The Transferable Residue

The element transfers, and its transfer converts three earlier findings into a testable procedure.

Requirement 21.1. (Recovery test of the measurement system). Before a source is relied upon to establish that a phenomenon of a stated character has occurred or has failed to occur, the measurement system producing that source is simulated, a synthetic condition containing the phenomenon is passed through the simulation, and the analysis to be used is applied to the synthetic output. Where the analysis fails to recover the phenomenon, the source establishes nothing about it, and a report drawing on that source states this.

The requirement operationalises what Part IV established by argument. Claim 16.1 held that the release schedule fixes the class of phenomena that can appear in evidence; a recovery test determines, for a particular phenomenon and a particular schedule, whether it does. Claim 16.2 held that fast variation is registered as slow structure; a recovery test exhibits the artefact by passing a known fast variation through the simulated schedule and displaying what emerges. Claim 15.1 held that selection varies with age; a recovery test under a period's recording régime shows what that régime would have registered.

Two features distinguish the requirement from the others in Part V. It is the only one whose satisfaction produces a positive demonstration in place of a record or a disclosure, since its output is an exhibit that a reader may examine. And it bears against the party conducting it in a specific way: an arrangement that runs the test and finds its source unable to recover a phenomenon has produced evidence that its own evidential basis is inadequate for a question it may prefer to answer.

The limit of the requirement follows from the first presupposition. A simulation of a measurement system is a model of that system, and a test conducted through an inaccurate simula-

tion establishes what an inaccurate simulation would deliver. The test is therefore informative in proportion to the documentation of the transformation, and the sections of a measurement system that are least documented are the sections about which the test says least. Recording practice by frontline officials, which Section 8 identified as part of the forward operator, is ordinarily the least documented part.

Assignment 21.2. (Class of the element). Mock pipelines fall in class iv, transfers intact, since the measurement system is documented and its simulation requires effort in place of new observation. A secondary assignment in class iii records that the fidelity of the simulation depends on documentation supplied by the party whose measurement system is being tested.

22 Ladder Calibration and the Retention of Persistent Tension

This section closes the third family with the practice by which a discipline handles two independent routes to one quantity that disagree. Its objective is to establish what the astrophysical handling requires, to identify the governance counterpart, and to derive from the comparison a result that supplies the paper with its one constructive instrument for estimating the movement of a system without a model. The section follows the template fixed in Section 7.

22.1 The Structure of the Practice in Astrophysics

Distances in cosmology are established by calibrated chains, each rung of which is anchored to the one below. Independent chains reach the same quantities by different routes, and the expansion rate of the universe is presently the subject of a persistent disagreement between a route running through the early universe and a route running through the local distance ladder [44].

The handling of that disagreement is what concerns the present argument. The disciplines have not averaged the two determinations, have not selected the one produced by the method judged more reliable, and have not suppressed the disagreement pending resolution. Both determinations remain in circulation with their uncertainties, and the disagreement is treated as information about the model connecting them. The persistence of the disagreement across improvements in each route is itself a reported finding, and the search for its source is a substantial research programme.

22.2 Presuppositions of the Practice

Three conditions are required.

The routes are genuinely independent, so that a disagreement between them bears on something other than a shared error.

The uncertainties are stated well enough that a disagreement can be distinguished from a fluctuation.

No party holds authority to require a single number, so that the disagreement may be retained without resolution for as long as it persists.

Each admits failure. The first fails where the routes share a calibration or a frame. The second fails where the uncertainties are unstated or understated. The third fails where a decision requires one figure.

22.3 The Governance Counterpart

Governance frequently possesses two routes to one quantity. Employment is measured by a household survey and by administrative payroll records. Crime is measured by police recording and by a victimisation survey. Health conditions are measured by clinical registration and by population survey. Income is measured by household reporting and by tax records. In each case the two routes have different frames, different definitional boundaries, different production intervals and different failure modes.

The routes disagree, and the disagreements are documented by the agencies producing them. What differs from the astrophysical case is the handling.

22.4 Status of the Presuppositions in Governance

The first is partly satisfied. The routes named above differ in frame and in method to a degree that makes their disagreement informative, though shared elements remain: both may rest on a common population estimate, and both may be affected by a common definitional revision.

The second is partly satisfied. Sampling uncertainty is stated for survey routes; the uncertainty attaching to an administrative count is ordinarily unstated, since the count is presented as enumeration.

The third fails, and its failure is the section's principal finding on the negative side. A decision requires a figure, and an arrangement holding two figures for one quantity resolves the disagreement in order to proceed. The resolution takes one of three forms: an average, a selection of the route judged authoritative, or the designation of one route as the official series and the other as a check. Each disposes of the disagreement, and each discards what the disagreement carried.

22.5 The Transferable Residue

The practice transfers as a requirement, and the comparison yields one result of a kind the rest of the paper does not supply.

Requirement 22.1. (Retention of a disagreement between routes). Where two routes to one quantity disagree beyond their stated uncertainties, both determinations are reported with the decision they inform, the disagreement is recorded as a quantity, and its persistence or resolution across successive periods is tracked. A single figure adopted for operational purposes is marked as an adoption and is accompanied by the range the two routes span.

Claim 22.2. (Information carried by a disagreement between routes of differing age). Where two routes to one quantity have different reference times, the disagreement between them is a compound of the differences in their frames, definitions and errors, and of the movement of the system across the interval separating their reference times. The frame and

definitional components are estimable from periods in which the two routes refer to the same moment. What remains after their removal carries information about the movement of the system over the interval, and it carries that information without a dynamical model of the system.

The claim requires care, and three restrictions belong with it.

It supplies a comparison and no absolute quantity. What is recovered is the change over the interval separating the two reference times, and where that interval is short relative to the lookback of the slower route the recovered quantity bears on a fraction of the interval that matters.

Its estimation requires periods of overlap, and the overlap must be one in which the two routes genuinely refer to the same moment. Where the faster route is available continuously and the slower periodically, such periods exist by construction, and the comparison of the fast route at the slow route's reference time with the fast route at the present supplies the interval movement directly. Where the slower route is the only source for a quantity, the claim is unavailable.

It rests on the frame and definitional differences being stable across the interval. Where a definition has been revised within the interval, the separation fails, which is Requirement 18.2 operating as a precondition of the present one.

Subject to those restrictions the claim supplies something the paper has otherwise withheld. The evidence horizon of Definition 10.2 was defined without a means of measuring it, and Claim 10.4 supplied an ordering in place of a magnitude on the ground that a magnitude would require a dynamical model. A pair of routes with differing reference times supplies a partial empirical substitute: it exhibits how much a quantity moves over an interval of a stated length, without asserting why it moved. Where such a pair exists, an arrangement can estimate what its lookback costs it in the specific quantity at issue.

Assignment 22.3. (Class of the element). Ladder calibration falls in class *iv*, transfers intact, for Requirement 22.1 and Claim 22.2, both of which require the retention and comparison of existing series. A secondary assignment in class *iii* records the failure of the third presupposition, since the party requiring a single figure is the party the figure is used to assess.

23 The Disanalogy of Reflexivity

This section opens Part VI, which assembles the results of the transfer tests. Its subject is the disanalogy identified in advance at §6.4 as the principal one: the astrophysical observer does not act on the observed object, and the observed object does not read its own description, while a governing party does the first and a governed system does the second. Its objective is to establish what follows for the correspondence requirement, and its result is that the interval τ_G is a coupling in governance and a delay in astrophysics, so that reconstruction of a present condition is differently posed here and not merely harder. The section closes with the conditions under which the disanalogy is weak, since a claim of this generality requires a statement of where it does not bite.

23.1 The Interval in Its Coupling Form

An astrophysical source emits, the signal propagates, and the observer receives. Nothing the observer does enters the sequence. The interval is therefore a pure delay: what is received is a state the source occupied for reasons entirely its own.

A governing party observes in order to act, and acts on the system it observes. The condition recorded at the reference time is accordingly a condition the system occupies partly in consequence of the previous round of governance, and the party reading a lagged indicator is reading, among other things, the trace of its own earlier intervention arriving late.

Claim 23.1. (The interval as a coupling). In governance the interval between the reference time of a body of evidence and the moment of decision is an interval during which the system responds to the intervention that the previous body of evidence informed. The evidence read at t_d describes a condition at t_e that was itself produced in part by an action taken on evidence with reference time earlier still. The interval therefore couples successive rounds of governance to one another, and a party that corrects for the delay without accounting for the coupling attributes to the system a movement that its own prior action produced.

The claim has a consequence for Requirement 11.4. The residual monitoring recommended there compares an anticipated value with a received one, and a persistent discrepancy was read as evidence that the account has departed from the system. Under Claim 23.1 a discrepancy may equally record the system's response to the intervention the account informed, which is a different finding with a different implication. The requirement stands, since detecting that something has changed remains informative, and its interpretation acquires an alternative that must be entertained.

23.2 Inversion of the Lag Under Anticipation

The second element of the disanalogy produces an effect with no counterpart at all in the observational case, and it is the most disruptive consequence of reflexivity for the paper's framework.

Parties within a governed system read descriptions of that system, including descriptions produced for the purpose of deciding upon them, and they adjust in anticipation of the intervention those descriptions are expected to prompt. The adjustment occurs before the intervention.

Claim 23.2. (Anticipatory inversion). Where the parties within a governed system anticipate an intervention, part of the system's response to that intervention occurs before it is adopted, and is recorded by the measurement system at reference times preceding the decision. The observational record therefore contains effects of the intervention at times earlier than the intervention, and the sign of the interval between cause and recorded effect is not fixed.

The claim bears against a natural reading of the correspondence requirement. Sections 4 through 18 have treated the problem as one of evidence describing a condition the system has left. Claim 23.2 identifies a case in which the evidence describes a condition the system entered because of a decision that has not yet been taken, so that the evidence is stale with

respect to one process and premature with respect to another. Correction by more recent observation addresses the first and compounds the second, since the most recent observations are the ones most contaminated by anticipation.

The phenomenon is documented in domains where anticipation is strong. Announcements of tax changes shift the timing of transactions before the change takes effect; announced enforcement campaigns alter recorded conduct before enforcement begins; the prospect of a threshold-based allocation alters the quantities near the threshold before the allocation is made.

23.3 The Constitution of the Object by Its Description

A third element completes the disanalogy. The categories through which a governed system is described are available to the parties within it, and their availability changes what the parties do and how they understand what they are doing. A description of a population as at risk, of a practice as an offence, or of a condition as a diagnosis enters the conduct of those so described. The claim is Giddens' [15], and it is imported here without amendment.

Claim 23.3. (Reconstruction under constitution). The reconstruction of a governed system's present condition from lagged observation is conducted in categories that participate in constituting the condition being reconstructed. The object of the reconstruction is therefore partly a product of the descriptive apparatus applied to it, and the reconstruction is not the recovery of an independently standing condition. Reconstruction in governance is accordingly differently posed from reconstruction in astrophysics, and its difficulty is of a different kind.

Claim 23.3 supplies the general form of the failure that Section 8 and Section 18 found in particular places. The forward operator is interested because its administrators act on its outputs; the standard ruler is administered because its categories are constitutive of the phenomena they measure. Both are instances of a system whose description is internal to it.

23.4 Conditions Under Which the Disanalogy Is Weak

A claim of this generality requires a boundary, and the boundary is stated in terms of what reflexivity requires.

Reflexivity requires that parties within the system have access to the description, that they have an interest in the intervention it informs, and that they possess the capacity to adjust. Where any of the three is absent the disanalogy weakens, and the astrophysical situation is approached.

Three classes of governance question satisfy the conditions weakly. Physical and ecological systems governed for their own condition, of which air quality, water systems and habitat are instances, contain no parties who read the description, though the human parties acting upon them do. Quantities recorded automatically by instruments outside the awareness of those they concern are less exposed to anticipation than quantities recorded through self-report. And decisions whose subject matter is unknown to the parties affected, whether through the ordinary obscurity of administrative process or through deliberate withholding, exhibit weak anticipation and strong ethical difficulty of a different kind.

The boundary has a consequence worth recording. The elements of Part III whose failure was assigned to reflexivity fail less severely in the weakly reflexive domains, so forward modelling is a better-founded practice for an ecological system than for a labour market. The paper's negative results are therefore graded across subject matters, and a governing arrangement may determine where on that gradient a particular question falls by asking the three conditions of this subsection.

24 The Ensemble and the Time Series

This section states the paper's principal negative result. Its objective is to establish that the manoeuvre by which astrophysics converts its lookback condition from a defect into an instrument requires something governance lacks, that the substitute available to governance exchanges one difficulty for another, and that the exchange is ordinarily neither chosen nor declared. The section proceeds by setting out the astrophysical manoeuvre, identifying its requirement, establishing the trade, and drawing three consequences, of which the third bears on an operation examined in Part III.

24.1 The Conversion of Lookback into an Instrument

The condition described in Section 6 is a limitation, and astrophysics has made it the foundation of a research programme. Objects at different distances are observed at different epochs, and where those objects are instances of one class governed by one physics, the collection of them at many distances constitutes a record of how that class evolves. Galaxy evolution is studied in this way. The observer does not wait to see a galaxy change; the observer looks further away and sees an earlier one.

The manoeuvre substitutes an ensemble for a time series. A time series of one object over the relevant timescale is unavailable, and an ensemble of objects at many epochs stands in its place. What was a defect, the impossibility of observing the present state of a distant object, becomes the means by which the past is accessible at all.

24.2 The Requirement of the Substitution

The substitution requires that the instances be instances of one thing. Where objects at different distances are governed by one physics, an object observed at an earlier epoch informs about the earlier condition of an object observed now, and the inference proceeds through the shared law. Where the objects are governed by different laws, or where their differences are unmodelled, the collection is a collection of unlike things and the ensemble carries no information about evolution.

Section 17 recorded the two conditions in the form the cosmological case requires, statistical homogeneity and approximate independence, and found both unavailable in governance. The present requirement is the more fundamental one on which those two rest.

24.3 The Trade Between the Two Mismatches

Governance holds one system and one history. Its available substitute for an ensemble is comparison across cases: jurisdictions, regions, organisations, populations. The substitute is used constantly and is the basis of most evidence offered for policy, since a finding from elsewhere is what an evidence base ordinarily contains.

Proposition 24.1. (The trade between temporal and contextual mismatch). Two routes are available for informing a decision about a system's present condition and the effects of an intervention upon it. The first uses evidence from the same system at earlier times, and incurs the temporal mismatch this paper has examined: the evidence describes a condition the system may have left, under dynamics that may have altered, in categories that may have drifted. The second uses evidence from other systems at comparable times, and incurs the contextual mismatch examined in the external validity literature: the evidence describes a system whose support factors may differ from the target's. Astrophysics escapes both, because its instances at different epochs are governed by one law, so that a temporally distant instance is a contextually equivalent one. Governance possesses no such law, and the two mismatches are therefore exclusive alternatives between which a choice is made whenever evidence is assembled.

The proposition is stated as a trade and not as an impossibility, and the distinction matters. Neither route is closed. What is closed is the position astrophysics occupies, in which reaching into the past and reaching across instances are the same operation. In governance they are different operations with different defects, and a body of evidence assembled from both incurs both.

24.4 Consequences of the Trade

Three consequences follow, and the third is the sharpest.

Claim 24.2. (The undeclared choice). An evidence base assembled for a decision ordinarily contains items of both kinds without marking which mismatch each carries. A synthesis reporting a pooled estimate does not report how much of its weight rests on evidence distant in time from the target and how much on evidence distant in context. The choice between the two mismatches is therefore made in the composition of the base, by whoever assembled it, and is neither stated nor deliberated.

Claim 24.3. (Ambiguity in the demand for further evidence). A demand that an evidence base be strengthened has no determinate direction. Extending it backward in time deepens the temporal mismatch, since older evidence is further from the present condition and, per Claim 15.1, more selected. Extending it across further cases deepens the contextual mismatch, since additional cases are drawn from settings progressively less like the target. The two extensions improve precision and degrade correspondence in different currencies, and no exchange rate between them is available. A recommendation to gather more evidence is therefore incomplete until it states which extension it recommends and which mismatch it accepts.

Claim 24.4. (Simultaneous incurrence in meta-analytic synthesis). Meta-analytic synthesis pools along both axes at once. Its constituent studies were conducted at different times and

in different settings, and the pooling operation treats variation along both axes as draws from a common distribution. The resulting estimate therefore incurs the temporal mismatch of Claim 12.1 and the contextual mismatch of the external validity literature simultaneously, and reports a single figure in which the two are indistinguishable. The heterogeneity term absorbs both, and its magnitude carries no information about which of them produced it.

Claim 24.4 completes the argument begun at Claim 12.1 and identifies why the operation examined there is the paper's central applied target. A synthesis is the form in which evidence is ordinarily delivered to a governing party, it is the point at which the two mismatches are combined, and its output is a number whose composition along both axes is unrecoverable from the number itself. Requirement 12.2 addresses one axis by asking that estimates be ordered by reference date. The corresponding treatment of the other axis is the subject of the external validity literature, and the present claim is that the two treatments must be conducted together, since a variance component absorbing both cannot be attributed to either.

24.5 The Position the Result Leaves

The result is negative and its consequence for the paper's conclusion is direct.

Where a decision concerns a system's present condition, the evidence available describes that system at an earlier time or describes other systems at a comparable time, and both descriptions stand at a distance from the object of the decision. The distances are of different kinds and neither is eliminable. What remains available is the declaration of which distance a body of evidence carries, the examination of the base for the composition Claim 24.2 leaves unstated, and the allocation of the burden of justification developed in Section 28.

The result also fixes the limit of the paper's comparison. Astrophysics has been used throughout as a discipline that confronted the historical character of observation and built an epistemology around it, and Parts III to V examined what of that epistemology transfers. The present section identifies what does not and cannot: the manoeuvre that made the condition productive rests on a shared law among instances, and the absence of such a law in governance is a feature of the subject matter admitting no remedy of the kind the comparison supplies. The transfer is therefore bounded from above by Proposition 24.1, and what lies below that bound is assembled in Section 25.

25 The Classification of Transfer Failures

This section assembles the assignments recorded at the close of each element section and states what their distribution establishes. Its objective is to present the classification as a whole, since the argument of Parts III to V travels through the distribution and not through the order of exposition, and to draw the result that the distribution supports. The section presents the table, reads it by class, and states the finding together with the restrictions on it.

Table 2: Classification of the fifteen elements by the presupposition that fails in governance. Classes are I precondition absent, II reflexivity, III interested observer, IV transfers intact. Primary assignments are given first.

Element	Presupposition at issue in governance	Class
<i>Reconstruction of a past state</i>		
Inverse inference	Operator fixed independently of the answer	III, IV
Forward modelling	System closed to the modeller	II, IV
Uncertainty propagation	Law and bounded model family	I, IV
Data assimilation	Observability of the state	I, III, IV
Retrodiction	One system across epochs (holds)	IV, I
<i>The observing situation</i>		
Past light cone	Intervals known and commensurable	I, IV
Horizon	Boundary determinable (holds)	IV, III
Selection at depth	Selection function characterisable	I, IV
Survey cadence	Limit computable from schedule (holds)	IV, III
Single realisation	Homogeneity and independence	I
Standard rulers	Reference administered by no party	III, IV, I
<i>Discipline of the observing community</i>		
Blind analysis	Procedure fixable, result accepted	IV, I
Provenance and re-reduction	Conditions recorded and transmitted	IV, III
Mock pipelines	Measurement system simulable (holds)	IV, III
Ladder calibration	No authority requiring one figure	IV, III

25.1 The Distribution of Primary Assignments

Table 2 records the assignments. Read by primary class, five elements fall in class I, one in class II, two in class III, and seven in class IV.

The distribution across families is the feature that carries the argument. The first family, the reconstruction of a past state, contains four elements whose primary assignment records a failure and one that transfers. The second family divides, with three failures and three transfers. The third family transfers entire.

The pattern admits a statement that does not depend on the counts.

Claim 25.1. (The line the classification draws). An element of astrophysical practice transfers to governance where its presupposition concerns the conduct of the observing community, and fails where its presupposition concerns a property of the observed system or of the relation between that system and the instruments trained on it. Every element in the third family transfers. Every element whose presupposition requires a dynamical law, an observability condition, a characterisable selection function, a common distribution across instances, or a reference administered by no party fails, and the failures are properties of governance as a subject matter, distinct from defects of practice within it.

25.2 Class III Read Across the Table

Class III appears twice as a primary assignment and seven times as a secondary one, and the pattern is worth reading in its own right.

Where the interested observer is the primary failure, the element requires a measurement

apparatus settled independently of the party assessed by its outputs, which the inverse inference and the standard ruler both do. Where it is secondary, the element is otherwise available and its operation is conditioned by an interest: the boundary of the unobserved region is set by the party declaring it, the schedule fixing a resolution limit is set by the party the series assesses, the fidelity of a simulated measurement system depends on documentation the same party supplies, the retention of vintages is administered by the party whose past conduct they would be used to assess, and the demand for a single figure comes from the party the figure assesses.

Claim 25.2. (Interest as a modifier). The interested position of the observer is rarely the sole obstacle to an element's transfer and is present in almost every case. It operates as a condition on the elements that transfer, determining how far each is available in practice, and the requirements yielded by those elements are accordingly requirements a party imposes on itself.

Claim 25.2 identifies the respect in which the class iv results are weaker than their classification suggests, and Section 29 states the consequence. An element that transfers intact transfers into a setting where its adoption is a decision of the party it constrains.

25.3 Class II Read Across the Table

Class ii appears once as a primary assignment, and the appearance understates the reach of the disanalogy that Section 23 examined.

The reason is a feature of the classification. An element is assigned to class ii where its pre-supposition explicitly requires that the observed system evolve independently of the observation, and forward modelling is the only element whose statement contains that requirement directly. Reflexivity nevertheless conditions the interpretation of results throughout: it supplies an alternative reading of the residual monitored under Requirement 11.4, it confounds the consistency examination of Requirement 12.2, and Claim 23.3 identifies it as the general form of the failures assigned to class iii in the inverse inference and the standard ruler.

The counts therefore understate class ii and the table should be read with Section 23 alongside it.

25.4 The Character of the Transferable Requirements

The elements assigned primarily to class iv yield the following requirements: Requirement 12.2 on temporal consistency, Requirement 14.1 on declaration of the unobserved region, Requirement 16.3 on the resolution limit, Requirement 19.2 on fixed procedure and complete record, Requirement 20.1 on the indexed evidence item, Requirement 21.1 on the recovery test, and Requirements 22.1 and 22.2 on the retention and reading of disagreement. Two further requirements, 9.2 on predicted evidence and 11.4 on residual monitoring, arise as secondary residues of elements whose primary assignment is a failure.

Claim 25.3. (The character of the transferable). Every requirement yielded by the transfer test is a discipline of declaration, of record, of procedural constraint or of comparison. None

requires the construction of an estimation system, the acquisition of a dynamical model, or the enlargement of the measurement apparatus. What a mature observational science supplies to governance is therefore a set of practices for handling observations honestly, and not a technique for recovering the present condition of a system from lagged observation of it.

Claim 25.3 states the result the paper has been assembling and hands it to Part VII. Its two components are stated separately in Sections 27 and 28: what follows for the domain within which estimation remains legitimate, and what follows for a party proposing to act where it does not.

26 Grammatical Tense and the Displacement of Responsibility

This section closes Part VI with the element of astrophysical practice identified at §6.2 as the one bearing most directly on governance, and it is not a method. Its objective is to establish that the grammatical form in which policy evidence is reported asserts more than its production supports, to identify what the surplus costs, and to connect the requirement of this paper to the third of the limitations set out in Section 3. Its result is that the correspondence failure and the displacement of responsibility onto evidence are connected through the tense in which evidence speaks, so that the fourth limitation aggravates the third.

26.1 The Concession Restated

Astrophysics reports its objects in the past tense of a stated epoch. A quantity is given at a redshift, a property is attributed to a population at a stated cosmic time, and a claim about a present condition is marked as an extrapolation carried by a named model.

Three things follow from that discipline, and they were stated at §6.2 in a form the present section now uses. An inference from observation to present condition becomes visible as an inference. Being visible, it carries an assumption that someone has stated. Carrying a stated assumption, it has a defender and a place at which an objection may be lodged.

26.2 The Tense of Policy Evidence

Policy evidence speaks in the present. An unemployment rate is reported as what unemployment is; an evaluation reports what a programme does; a dashboard displays a condition. The reference period appears in documentation and the assertion appears in the present indicative.

Claim 26.1. (Surplus of the grammatical form over the warrant). A present-tense report of a quantity produced through a finite interval asserts a claim about a condition at the time of reading, while its production supports a claim about a condition at the reference time. The difference between the two is an inference, and the grammatical form conceals that an inference has been made. The surplus is therefore carried without an assumption being stated, and consequently without a defender and without a place at which it may be contested.

The claim is modest in what it asserts and consequential in what follows. It does not hold that any party is deceived, since a reader familiar with statistical production knows that figures have reference periods. It holds that the form of the report leaves the inference unstated, so that the question of how far the system may have moved is never raised at the point where it would bear.

26.3 The Connection to the Displacement of Responsibility

Section 3 recorded as the third established limitation that an appeal to evidence relocates the ground of a decision from the party deciding to a procedure, so that the decision becomes attributable to an authority occupying no office. The present section identifies what the relocation requires.

Claim 26.2. (Dependence of the displacement on the tense). The formula by which a decision is attributed to the evidence operates because the evidence appears to describe the situation the decision addresses. Where evidence is reported in the present tense, the party deciding may present itself as having followed a description of the present, and the description bears the responsibility. Where evidence is reported with its reference time and the inference to the present is marked, the same formula asserts that the party acted on a description of an earlier condition together with a judgement about the interval. The judgement is the party's own, is unsupported by the evidence, and is therefore attributable to the party. The fourth limitation accordingly aggravates the third, and the correction of the tense partly disarms it.

One consequence of the claim belongs with the account given in Section 3. A position from which a warrant issues must be able to speak about the act it warrants, and an authority whose statements are marked as descriptions of an earlier condition cannot occupy that position for a decision taken now. The present tense is therefore a condition of the installation, and the marking recommended below removes it. What the marking withdraws is not the authority of the evidence over its own subject matter, which is untouched, but its availability as the locus from which a present act receives its guarantee.

The claim also identifies the point at which the present paper meets the analysis of the political conditions of evidence use noted in §5.6. That analysis concerns the interests shaping which evidence is produced and how it is interpreted. The present claim concerns a feature of the evidence itself which determines how much resistance those interests encounter: a report whose inferential surplus is unmarked offers a party a description it may adopt without acknowledging that it has judged anything, and a report whose surplus is marked requires the party to state a judgement of its own.

26.4 The Requirement and Its Weakness

Requirement 26.3. (Tense of a report of condition). A report of the condition of a governed system is written in a tense its production interval supports. A statement about the condition at the time of reading is marked as an extrapolation from the reference condition, and the ground of the extrapolation is named. Where no ground is available, the report states the

reference condition and states that the present condition is undetermined by the evidence adduced.

The requirement costs nothing in apparatus and it is the weakest of the paper's requirements in a specific respect that should be stated plainly. It changes no fact about what is known. A party that marks its inferences knows exactly what a party that conceals them knows, and a decision taken under a marked inference may be the same decision.

Its value lies elsewhere and is worth stating precisely. It relocates the inference from the implicit background of a report into the body of it, where the three consequences set out at the opening of this section follow: an assumption is stated, a party stands behind it, and an objection has somewhere to go. The requirement produces no knowledge and creates the conditions under which the absence of knowledge can be raised.

One objection deserves an answer. It may be held that governing arrangements require assertive language, that a report hedged at every figure is unusable, and that the marking recommended here would produce documents no minister reads. The answer is that the discipline is available in a form that costs little: a reference time attached to a figure and a single marked sentence where an extrapolation is relied upon. Astrophysics maintains the discipline in papers of high technical density without rendering them unreadable, and the cost of the alternative is that the assumption on which a decision rests goes undefended because it goes unstated.

27 The Partition of the State Space and the Domain of Estimation

This section opens Part VII by stating the consequence of Proposition 11.1 for the standing of estimation in governance. Its objective is to establish where estimation from lagged observation is legitimate, where the question of estimation does not arise, and what a governing arrangement owes at the boundary. The section is deliberately affirmative in its first half, since a framework that declined estimation where estimation works would be a framework no practitioner should accept, and the paper's negative results are worth only as much as its positive ones are credible.

27.1 The Region in Which Estimation Is Legitimate

Where the quantities bearing on a decision lie within the component recoverable from the instituted measurements, the reconstruction of a present condition from lagged observation is a well-posed problem. Its difficulty is a matter of degree, its errors are estimable, and the machinery developed in the fields that practise it applies.

Epidemic surveillance is the standing instance and its standing is literal. The estimation of current infection incidence from case reports arriving with a delay distribution is the reconstruction of a present condition from lagged observation, conducted for the purpose of governing, and it works. Macroeconomic nowcasting occupies the same position. The existence of these practices is what establishes that the paper's negative results are bounded, and a framework denying them would be refuted by them.

Claim 27.1. (The domain of legitimate estimation). Where the quantities bearing on a decision are observable through the instituted measurements, where a dynamical account of their evolution is available and testable, and where the interval falls within the evidence horizon of Definition 10.2, estimation of the present condition is legitimate and its products carry the standing their construction supports. The paper's negative results concern the complement of this region and establish nothing against practice within it.

Two features of the claim are worth stating. Its three conditions are conjunctive, so a domain satisfying two of them lies outside it: a quantity observable and modelled but read across an interval beyond its horizon supports no claim about the present. And the conditions are assessable in advance of any estimate, which is what makes the claim operative.

27.2 The Region in Which the Question Does Not Arise

Where the quantities bearing on a decision lie in the unobservable component, no estimate of their present values exists to be constructed, and the deficiency is structural in the sense established at Proposition 11.1 and Equation 13.

Claim 27.2. (Priority of the partition over the choice of method). The determination of which region a quantity occupies is prior to the selection of an estimation method and is conducted by different means. It requires an examination of the relation between the instituted measurements and the quantity, and it is settled before any question of technique arises. A methodological dispute conducted over a quantity in the second region is conducted over a problem that is undefined, and the disputants' preference among methods is idle.

The claim supplies the answer to a question the introduction left open. Section 1 declined a methodological proposal, and the refusal might be read as a general scepticism about technique. It is not. The paper's position is that technique is appropriate within a region whose extent is determinable, and that the determination is the step ordinarily omitted. A governing arrangement that conducts it will find some of its questions inside the region and will proceed with the methods those fields supply.

27.3 The Obligation at the Boundary

The boundary between the regions is partly a matter of what the world permits and partly a matter of what was instituted, per Claim 11.2, and the composition of the boundary is what a governing arrangement owes an account of.

Requirement 14.1 stated the obligation in its general form: an account of a system's condition states which of the quantities bearing on the decision the measurements reach and which they do not. The present section adds what the declaration exposes.

Claim 27.3. (Attribution of the boundary). Where the unobserved region is declared, its contents divide into quantities beyond every available instrument and quantities beyond the instituted measurements alone. The second division is attributable to the parties that instituted the measurements, and a claim of unknowability concerning a quantity in that division is a claim about an arrangement, distinct from a claim about the world. The declaration there-

fore converts an appeal to unknowability from a stopping point into an assertion that may be examined.

The claim gives the partition its political form. An arrangement may enlarge the unobservable region by declining to measure, and the enlargement is invisible while the boundary is undeclared. Declaring the boundary makes the enlargement visible without requiring that it be reversed, and the question of whether it should be reversed becomes a question that can be put.

28 Restraint Under Temporal Correspondence Failure

This section states the paper's normative content. Its objective is to convert the findings of Parts III to VI into an allocation of the burden of justification, to enumerate the conditions under which that allocation operates, to distinguish it from a presumption against action, and to identify the one positive recommendation the argument supports. The section states the criterion, its triggers, its symmetry, what discharges it, and its differentiation from the nearest established formulation.

28.1 The Criterion

Criterion 28.1. (Allocation of the burden under correspondence failure). Where a course available to a party would produce effects that are substantially irreversible, and where temporal correspondence fails for the evidence adduced in support of that course, the burden of justifying its adoption lies with the party adopting it. Where correspondence holds, or where the effects of the course are revisable at proportionate cost, the evidence bears its ordinary weight and the choice falls to ordinary political and administrative settlement.

Three terms carry the criterion and are fixed here.

A *course* is anything a party may do, including declining to act, and the subsection on symmetry below states why the inclusion is required.

Substantial irreversibility concerns the effects and is relative to available means and to a horizon that the party states. It is a factual matter admitting dispute, and the criterion requires that the horizon be stated whenever it is applied.

Correspondence failure is established by the conditions of the following subsection and by nothing else. The criterion does not fire on uncertainty as such.

28.2 The Conditions Establishing Failure

The transfer tests supply six determinate conditions, each of which was derived in its own section and each of which is assessable without a dynamical model of the governed system.

The interval exceeds the evidence horizon of Definition 10.2 and Equation 10, so the distribution over conditions consistent with the evidence no longer separates the alternatives at issue.

A quantity on which the choice turns lies in the unobservable component of Proposition 11.1 and Equation 13, so no estimate of its present value exists.

The residual sequence of Requirement 11.4 exhibits consistent sign or growing magnitude, so the account held has departed from the system.

The estimates constituting the evidence base exhibit trend by reference date under Requirement 12.2, and the party pooling has not discharged the burden of showing that pooling is warranted.

The recovery test of Requirement 21.1 fails, so the source relied upon would not register the phenomenon at issue.

Two routes to a quantity disagree beyond their stated uncertainties under Requirement 22.1, and the disagreement is unexplained by the frame and definitional differences between them.

The list is a set of sufficient conditions and makes no claim to completeness. Each is determinable from records a governing arrangement holds or can construct, and none requires the apparatus the paper has declined throughout.

28.3 Symmetry Across Acting and Declining

The criterion is stated over courses, and the inclusion of deferral within that term is required by an argument that must be made explicitly, since a criterion keyed to correspondence failure would otherwise operate as a presumption against action.

Claim 28.2. (Irreversibility on the side of delay). Deferral produces effects, and its effects are sometimes irreversible. A transition postponed past the point at which it could have prevented a threshold crossing forecloses as completely as an intervention does; an opportunity for a settlement that closes, a population that ages out of a window of effect, and a habitat lost while a decision was pending are foreclosures produced by declining to act. A criterion stated over intervention alone would place no burden on the deferring party in these cases, and would therefore favour whichever course the existing arrangement already follows.

The consequence for the present criterion is direct and is the paper's principal protection against a misreading. Correspondence failure raises the burden on any course whose effects are substantially irreversible, and the party proposing to continue as it is bears the burden as fully as the party proposing to change, where continuation forecloses. The criterion allocates a burden and states no presumption in favour of the status quo, and a party invoking it to resist a change must accept it when the change is resisted.

28.4 The Discharge of the Burden

The burden is argumentative and is discharged by argument, since by hypothesis the evidence does not reach the condition at issue. A party discharges it by stating four things: the foreclosure the course would produce and the horizon over which it is irreversible; which of the six conditions listed above obtains and what is therefore undetermined by the evidence; the ground on which the party nevertheless judges the course warranted, identified as theory, as analogy, as the experience of the parties, or as interest, per Claim 14.2; and what would be observed if the judgement were wrong, together with when that observation would arrive.

The fourth element is Requirement 9.2 operating as a component of discharge, and it is what keeps the burden from being satisfied by assertion. A party that cannot say what would show its judgement to be mistaken has not discharged the burden; a party that says so has produced a commitment against which it may later be assessed, subject to the retention that Claim 20.2 requires.

28.5 Differentiation from Precaution

The criterion resembles the precautionary principle and differs from it on two axes, and the differences run in opposite directions.

Its trigger is narrower. Precaution in its familiar forms fires on uncertainty together with the prospect of serious harm, and uncertainty attends every course including refusal, which is the ground of the standing objection that such a principle speaks against the whole choice set. The present criterion fires on a determinate failure of correspondence established by one of six assessable conditions, and a course supported by evidence within its horizon, on quantities within the observable region, with a stable residual and consistent routes, does not trigger it however uncertain its outcome may be.

Its consequence is weaker. Precaution in its strong forms operates as a bar, requiring a demonstration of safety before a course may proceed. The present criterion allocates a burden that argument discharges and forecloses no course, and the question remains decidable even where the burden falls on every alternative, since a burden falling on all alternatives is discharged by whichever party states its ground most adequately.

The pairing of a narrower trigger with a weaker consequence is what permits the criterion to reach cases precaution releases, of which a foreclosure whose outcome is certain is the standing instance, while releasing cases precaution catches.

28.6 The Positive Recommendation

The argument supports one recommendation that is more than a discipline of reporting, and it follows from Claim 16.4.

Production intervals and sampling intervals are set by administrative convention and by the cost of collection, and the association between them means that the variables carrying a system toward a stability boundary are observed by the sources that arrive latest and resolve most coarsely. The association is contingent. An arrangement may collect a slow structural variable more frequently and release it more quickly, at a cost, and the cost buys a reduction in the region where Criterion 28.1 fires.

Claim 28.3. (Measurement effort as a governance choice). The allocation of measurement effort across variables determines the extent of the region in which correspondence fails, and the allocation is a decision an arrangement takes and may revise. An arrangement that concentrates its measurement effort on quantities that are cheap to observe and fast to arrive has enlarged the region in which its own courses carry the burden of Criterion 28.1, and the enlargement is attributable to it.

The claim is the one place at which the paper recommends the expenditure of resources, and the recommendation is specific: measurement effort directed toward the slow variables carrying structural condition buys correspondence where correspondence is worth most. It is a recommendation about what to measure and how often, and it requires no estimation apparatus and no dynamical model.

29 Scope, Residue, and Directions

This section states what the paper establishes, what it withholds, and where its arguments are weakest. Its objective is to fix the boundary of the claims so that a reader may assess them against the right standard, and to record the questions the argument raises and leaves open. It closes with the paper's position.

29.1 The Boundary of the Claims

The paper establishes that the instruments by which evidence-based policy making assesses evidence contain no category for correspondence between the condition an item of evidence represents and the condition obtaining at the moment of decision, and it establishes this by inspection of the instruments in place of inference from their consequences.

It establishes that the requirement is independent of the problem of induction, in the sense that the concession of everything that problem places in doubt leaves the requirement in force, and that it is constructible from premises an empirical method licenses.

It establishes, by comparison with a discipline whose observations are historical by construction, that the elements of that discipline's practice transfer to governance where their presuppositions concern the conduct of an observing community and fail where they concern properties of the observed system or of its relation to the instruments, and it exhibits the distribution in Table 2.

It establishes that the manoeuvre by which that discipline converts its lookback condition into an instrument requires a shared law among instances, that governance possesses no such law, and that its available substitute exchanges temporal mismatch for contextual mismatch without an exchange rate between them.

It establishes nothing about astrophysics, proposes no method, supplies no measure of staleness, and ranks no class of evidence above another.

29.2 The Weakest Points

Four weaknesses are stated here by the paper, in preference to leaving them to be found in it.

The requirements yielded by the transfer test are self-imposed. Claim 25.2 established that the interested position of the observer conditions almost every element that transfers, and the consequence is that each requirement is adopted by the party it constrains. A party that would prefer its inferential surplus unmarked, its unobserved region undeclared, its analyses unrecorded and its vintages discarded is under no compulsion from this paper. The re-

quirements describe what an arrangement that has decided to bind itself would do, and the question of what would produce that decision lies outside the paper's argument.

The criterion of Section 28 fires on a property whose instantiation is contestable exactly where it matters. Substantial irreversibility is a factual matter admitting dispute, and a party wishing to avoid the burden may contest the determination in place of discharging it. The same holds of several of the six conditions, of which the adequacy of a recovery test and the sufficiency of an explanation for a disagreement between routes are the most contestable.

The paper supplies no procedure and administers its requirements through no office, and the cost of that refusal is that an arrangement seeking to comply has nothing to comply with. The refusal is deliberate, on the ground that machinery specified from outside is administered by the party whose conduct is at issue and is satisfied by the production of documents. The cost is real and is not answered by the ground.

The association on which two of the paper's sharper claims rest is empirical and contingent. Claim 13.2 and Claim 16.4 both depend on slow-moving structural variables being observed by sources with long production intervals and coarse sampling. The association holds in the statistical systems the paper has considered and is a fact about how those systems have been resourced. A reader who finds a counterexample has narrowed the claims and has not refuted the framework.

29.3 Questions the Argument Raises and Leaves Open

Four questions are recorded, in descending order of tractability.

Claim 22.2 supplies a means of estimating the movement of a quantity over an interval from the disagreement between two routes of differing reference time, subject to the availability of overlap periods and to definitional stability. Whether the instrument generalises beyond the paired series named in Section 22, and what precision it can attain, is a question that existing data would answer.

Claim 12.3 identified the underdetermination of a temporal trend in pooled estimates between the research process and the movement of the system, and Claim 15.1 supplied an age-dependent mechanism operating on the first reading. Separating the two requires evidence about the selection régime of each period, and the construction of such evidence for a domain in which the régimes are documented would settle the question for that domain.

Claim 23.2 established that anticipation places effects of an intervention at reference times preceding it, so that the sign of the interval between cause and recorded effect is not fixed. The consequences for the correspondence requirement have been stated here and not developed. A treatment of correspondence under anticipation, in which the interval is signed and the contamination of recent observations is represented, would be a substantial extension.

The evidence horizon of Definition 10.2 has been defined, shown measurable in one domain, and supplied with an ordering in place of a magnitude. Whether horizons can be estimated for domains lacking a record of subsequent adequacy determinations, and whether the ordering can be strengthened without importing the dynamical model whose absence the paper has established, is the deepest of the four and the least tractable.

29.4 The Position

Evidence describes a condition a system occupied. Governance acts on the condition a system occupies. The interval between them is a property of how evidence is produced and survives every improvement in the instruments producing it, and the requirement that the two conditions correspond is therefore permanent and is not registered by the apparatus through which evidence is assessed.

A discipline that met the same condition in a purer form gave up the present tense about its objects and rebuilt its epistemology on the assumption that observation is historical. What that rebuilding supplies to governance is a set of practices for handling observations honestly: carrying the indices, declaring the boundary, stating what a source could not have shown, retaining what was held at each date, and reporting a disagreement in place of resolving it. What it does not supply is a technique for recovering the present, and the reason it does not is that governance lacks the shared law that made the recovery possible there.

Where the recovery is unavailable and the effects of a course are irreversible, what remains is the allocation of the burden of justification to the party proposing the course, and the requirement that the party state the ground on which it acts and what would show it to be mistaken. That is a modest conclusion, and it is what the comparison supports.

References

- [1] Bareinboim, Elias, and Judea Pearl. “Causal Inference and the Data-Fusion Problem.” *Proceedings of the National Academy of Sciences* 113, no. 27 (2016): 7345–7352.
- [2] Campbell, Donald T. “Assessing the Impact of Planned Social Change.” *Evaluation and Program Planning* 2, no. 1 (1979): 67–90.
- [3] Cabinet Office, Government of Japan. “Naikakufu ni okeru EBPM e no torikumi” [Efforts toward EBPM in the Cabinet Office]. <https://www.cao.go.jp/others/kichou/ebpm/ebpm.html> (accessed 19 August 2026).
- [4] Cartwright, Nancy, and Jeremy Hardie. *Evidence-Based Policy: A Practical Guide to Doing It Better*. Oxford: Oxford University Press, 2012.
- [5] Schünemann, Holger J., Julian P. T. Higgins, Gunn E. Vist, Paul Glasziou, Elie A. Akl, Nicole Skoetz, and Gordon H. Guyatt. “Completing Summary of Findings Tables and Grading the Certainty of the Evidence.” In *Cochrane Handbook for Systematic Reviews of Interventions*, chapter 14. Cochrane, current version.
- [6] Cranmer, Kyle, Johann Brehmer, and Gilles Louppe. “The Frontier of Simulation-Based Inference.” *Proceedings of the National Academy of Sciences* 117, no. 48 (2020): 30055–30062.
- [7] Cranmer, Kyle, Johann Brehmer, and Gilles Louppe. “The Frontier of Simulation-Based Inference.” *Proceedings of the National Academy of Sciences* 117, no. 48 (2020): 30055–30062.

- [8] Croushore, Dean, and Tom Stark. "A Real-Time Data Set for Macroeconomists." *Journal of Econometrics* 105 (2001): 111–130.
- [9] Deaton, Angus, and Nancy Cartwright. "Understanding and Misunderstanding Randomized Controlled Trials." *Social Science and Medicine* 210 (2018): 2–21.
- [10] Deutsch, Karl W. *The Nerves of Government: Models of Political Communication and Control*. New York: Free Press of Glencoe, 1963.
- [11] Egami, Naoki, and Erin Hartman. "Elements of External Validity: Framework, Design, and Analysis." *American Political Science Review* (2022).
- [12] Eisenstein, Daniel J., et al. "Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies." *The Astrophysical Journal* 633, no. 2 (2005): 560–574.
- [13] Forrester, Jay W. "Counterintuitive Behavior of Social Systems." *Technology Review* 73, no. 3 (January 1971): 52–68.
- [14] Friedman, Milton. "The Lag in Effect of Monetary Policy." *Journal of Political Economy* 69, no. 5 (1961): 447–466.
- [15] Giddens, Anthony. *The Constitution of Society: Outline of the Theory of Structuration*. Cambridge: Polity Press, 1984.
- [16] Goodhart, Charles A. E. "Problems of Monetary Management: The U.K. Experience." First circulated in *Papers in Monetary Economics*, vol. 1 (Reserve Bank of Australia, 1975); reprinted in *Monetary Theory and Practice: The U.K. Experience*, 91–121. London: Macmillan, 1984.
- [17] Zhang, Yuqing, Pablo Alonso-Coello, Gordon H. Guyatt, et al. "GRADE Guidelines: 19. Assessing the Certainty of Evidence in the Importance of Outcomes or Values and Preferences: Risk of Bias and Indirectness." *Journal of Clinical Epidemiology* (2018).
- [18] Cabinet Secretariat, Government of Japan, Administrative Reform Promotion Headquarters. "EBPM no suishin" [Promotion of EBPM]. Materials of the EBPM Promotion Committee, including the EBPM policy for fiscal year 2026 adopted 17 February 2026. <https://www.gyokaku.go.jp/ebpm/index.html> (accessed 19 August 2026).
- [19] Haasnoot, Marjolijn, Andrew Warren, and Jan H. Kwakkel. "Dynamic Adaptive Policy Pathways (DAPP)." In *Decision Making under Deep Uncertainty: From Theory to Practice*, edited by Vincent A. W. J. Marchau, Warren E. Walker, Pieter J. T. M. Bloemen, and Steven W. Popper. Cham: Springer, 2019.
- [20] Habermas, Jürgen. *Legitimation Crisis*. Translated by Thomas McCarthy. Boston: Beacon Press, 1975. Originally published as *Legitimationsprobleme im Spätkapitalismus* (Frankfurt: Suhrkamp, 1973).

- [21] Höhle, Michael, and Matthias an der Heiden. "Bayesian Nowcasting during the STEC O104:H4 Outbreak in Germany, 2011." *Biometrics* 70, no. 4 (2014): 993–1002.
- [22] Hume, David. *A Treatise of Human Nature*. London, 1739–40.
- [23] Ioannidis, John P. A. "Contradicted and Initially Stronger Effects in Highly Cited Clinical Research." *JAMA* 294, no. 2 (2005): 218–228.
- [24] Kalman, Rudolf E. "A New Approach to Linear Filtering and Prediction Problems." *Transactions of the ASME, Journal of Basic Engineering* 82, no. 1 (1960): 35–45.
- [25] Kalnay, Eugenia. *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge: Cambridge University Press, 2003.
- [26] Klein, Joshua R., and Aaron Roodman. "Blind Analysis in Nuclear and Particle Physics." *Annual Review of Nuclear and Particle Science* 55 (2005): 141–163.
- [27] Klein, Joshua R., and Aaron Roodman. "Blind Analysis in Nuclear and Particle Physics." *Annual Review of Nuclear and Particle Science* 55 (2005): 141–163.
- [28] Lacan, Jacques. *Écrits: The First Complete Edition in English*. Translated by Bruce Fink. New York: W. W. Norton, 2006.
- [29] Lawless, Jerald F. "Adjustments for Reporting Delays and the Prediction of Occurred but Not Reported Events." *Canadian Journal of Statistics* 22 (1994): 15–31.
- [30] Lempert, Robert J. "Robust Decision Making (RDM)." In *Decision Making under Deep Uncertainty: From Theory to Practice*, edited by Vincent A. W. J. Marchau, Warren E. Walker, Pieter J. T. M. Bloemen, and Steven W. Popper, 23–51. Cham: Springer, 2019.
- [31] Lucas, Robert E., Jr. "Econometric Policy Evaluation: A Critique." *Carnegie-Rochester Conference Series on Public Policy* 1 (1976): 19–46.
- [32] Malmquist, Karl Gunnar. "On Some Relations in Stellar Statistics." *Meddelanden från Lunds Astronomiska Observatorium*, Series I, no. 100 (1922); also *Arkiv för Matematik, Astronomi och Fysik* 16 (1922). The effect is elaborated in Malmquist's later papers and is surveyed in P. Teerikorpi, "Observational Selection Bias Affecting the Determination of the Extragalactic Distance Scale," *Annual Review of Astronomy and Astrophysics* 35 (1997): 101–136.
- [33] Marchau, Vincent A. W. J., Warren E. Walker, Pieter J. T. M. Bloemen, and Steven W. Popper, eds. *Decision Making under Deep Uncertainty: From Theory to Practice*. Cham: Springer, 2019.
- [34] Martínez García, Laura, Andrea Juliana Sanabria, et al. "The Validity of Recommendations from Clinical Guidelines: A Survival Analysis." *Canadian Medical Association Journal* 186 (2014): 1211–1219.

- [35] Munger, Kevin. "Temporal Validity as Meta-Science." *Research and Politics* 10, no. 3 (2023).
- [36] Parkhurst, Justin. *The Politics of Evidence: From Evidence-Based Policy to the Good Governance of Evidence*. Abingdon: Routledge, 2017.
- [37] Pawson, Ray, and Nick Tilley. *Realistic Evaluation*. London: Sage, 1997.
- [38] Reiss, Julian. "Against External Validity." *Synthese* 196, no. 8 (2019): 3103–3121.
- [39] Rosenzweig, Mark R., and Christopher Udry. "External Validity in a Stochastic World." NBER Working Paper 22449, 2016.
- [40] Shannon, Claude E. "Communication in the Presence of Noise." *Proceedings of the IRE* 37, no. 1 (1949): 10–21.
- [41] Shekelle, Paul G., Eduardo Ortiz, Shannon Rhodes, Sally C. Morton, Martin P. Eccles, Jeremy M. Grimshaw, and Steven H. Woolf. "Validity of the Agency for Healthcare Research and Quality Clinical Practice Guidelines: How Quickly Do Guidelines Become Outdated?" *JAMA* 286, no. 12 (2001): 1461–1467.
- [42] Sterman, John D. *Business Dynamics: Systems Thinking and Modeling for a Complex World*. Boston: Irwin/McGraw-Hill, 2000.
- [43] Verde, Licia, Tommaso Treu, and Adam G. Riess. "Tensions between the Early and the Late Universe." *Nature Astronomy* 3 (2019): 891–895.
- [44] Verde, Licia, Tommaso Treu, and Adam G. Riess. "Tensions between the Early and the Late Universe." *Nature Astronomy* 3 (2019): 891–895.