Arbitrary Switches and Pool Exclusion: A Reply to Schurz

Albaraa Mesawi

Preprint

Abstract

Schurz proposes excluding methods with arbitrary parameters from a meta-inductive comparison pool because they lack a regularity-based or causal route to predictive success. Arbitrarily scheduled exceptions, including infinitely recurring exceptions of zero density, can preserve an admitted learner's limiting-success event under every common outcome law. This preservation refutes the unrestricted rationale without warranting exceptional forecasts. A shared bound on departures instead permits omitting these variants as redundant performance comparators at an explicit score cost. That bound does not select a warranted next action. The distinction separates three grounds for pool decisions: sustained performance, current predictive warrant, and limits on tracking.

1 Why pool exclusion needs a reason

Object-induction projects observed patterns to future outcomes. Schurz's meta-induction combines accessible methods' forecasts with weights based on their past predictive success. The comparison pool contains the methods whose forecasts and success records the agent can track simultaneously (Schurz, 2026). The choice of a comparison pool affects whether meta-induction can justify object-induction. The first step of Schurz's account establishes the optimality of meta-induction relative to the accessible methods in that pool. The second appeals to the past success of object-inductive methods to justify following them. A guarantee relative to a pool does not by itself determine which methods belong in it, or whether their weighted predictions support an object-inductive forecast (Schurz, 2026).

Goodman methods make this dependence visible. Immediately before a scheduled reversal, a method that has copied an inductive learner has exactly its past forecasts and success record, yet recommends the opposite next forecast. Giving both methods equal weight can prevent the combined prediction from approximating the learner's prediction. Bakshi's indetermination objection asks how the second step selects object-induction when such rivals remain available, even in a finite pool (Bakshi, 2025). Schurz agrees that a reason is needed for disregarding these rivals (Schurz, 2026). Their removal can support the second step only if it is justified independently of the preference it is meant to establish.

Schurz proposes both a performance response and an epistemic exclusion. The epistemic exclusion calls a method arbitrary when it contains parameters unrelated to observed events or other recognizable cues. Significantly above-random success, he argues, requires either learning from a qualitative regularity or following an oracle causally connected with the predicted events. He infers that a method based on arbitrary parameters satisfies neither condition. Since his definition makes a method arbitrary whenever it contains such a parameter, the inference as stated reaches every composite that contains one. His final paragraph connects arbitrary future behavior, indefinitely

many parameter values and a vanishing chance of accidental success to exclusion from the pool (Schurz, 2026). The inference at issue moves from an unsupported parameter in a composite rule to the loss of the whole rule's sources of success. The counterexample below rejects that unrestricted implication while leaving action-based exclusions open.

The relevant antecedents already limit what a reply can add. Pitts discusses induction with occasional noninductive exceptions (Pitts, 2023), and Schurz expressly acknowledges that the claimed superiority of object-induction does not cover these mixed methods (Schurz, 2026). His earlier work treats multiple switches and concedes accidental perfect prediction by a rule that forecasts a fixed event sequence independently of observations (Schurz, 2023). That work also rejects a parameter-counting objection to Goodman methods: future switch points are not adjustable to the data, and Akaike's criterion gives a Goodman method with a fixed switch point the same expected success as object-induction (Schurz, 2023). That rejection concerns whether curve-fitting criteria penalize future switch points; the present note concerns limiting-success events and current warrant. Schurz's 2026 performance response appeals to fixed-share meta-induction, which tracks a finite set of qualitative base methods. Schurz reports long-run optimality against their concatenations when switch counts grow sublinearly. He seeks a further epistemic solution because of the greater short-run losses within human planning horizons (Schurz, 2026).

The central question is whether arbitrary switching itself removes a composite's regularity-based route to sustained success. Schurz's generalized class permits one or multiple future switches between qualitative methods (Schurz, 2026). A two-switch composite in this class can continue learning from the same observations while departing at one stage without evidence. I grant the regularity-oracle disjunction and test the unrestricted inference on this finite case, then extend the mathematics to recurring schedules. The success event below is that a method's limiting average score exists and exceeds one half. Schurz does not specify this event in his chance claim. The event supplies one precise long-run reading of that claim; Section 3 assesses his concern about current warrant separately.

The argument first shows why even recurring arbitrary departures can preserve this success event. It then separates that result from the warrant needed to reject an exceptional forecast. Finally, it asks whether preserving a variant's success route requires tracking that variant separately. A common error allowance can justify representing a family by an independently retained learner for performance comparisons. These are distinct reasons for a pool decision: an unwarranted exception need not destroy sustained success, and a successful variant need not add a separately tracked comparator.

The unrestricted inference also bears on Schurz's concession. A mixed method that departs from the learner on a schedule selected without predictive evidence or a recognizable cue contains an arbitrary parameter, so the stated rationale would deny it both routes to above-random success and exclude it from the pool. Section 2 shows, under its stated assumptions, that such a method keeps the learner's regularity-based route to sustained success when its departures have zero density. For these mixed methods the inference fails. The elaboration that Schurz leaves for future work (Schurz, 2026) therefore cannot exclude them by appeal to a lost route to sustained success. Excluding them requires another justification, such as a premise about the exceptional action.

2 Finite and recurring arbitrary exceptions

Consider a hypothetical pair of forecasters who predict zero or one from the same observed outcomes. The second has copied the first so far. At the end of stage two, the second adopts this

rule: copy the first, reverse its fourth forecast, then copy it again. Stage four is chosen without predictive evidence. This single departure changes the total number of correct forecasts by exactly one after that stage. Its share of an ever longer record tends to zero. Immediately before stage four, the forecasters still have equal records but give opposite next forecasts. The example separates the disputed action from its effect on sustained performance. Departures that never end provide a stronger test of that separation.

Here a qualitative base method uses logical combinations of qualitative predicates without designated time points, relative to the chosen primitive vocabulary (Schurz, 2026). Let $f : \{0,1\}^{<\mathbb{N}} \to \{0,1\}$ be a total deterministic qualitative learner. At stage j, it predicts $f(x_{<j})$ from the actual observed history. Fix the present stage m and a set $S \subseteq \{m+2, m+3, \ldots\}$ of exceptional stages, selected without predictive evidence or a recognizable cue. Finite S suffices for the correction. Infinite schedules extend the finite-horizon concatenation setting in the cited passages. Define

$$g_S(x_{<j}) = \begin{cases} 1 - f(x_{<j}), & j \in S, \\ f(x_{<j}), & j \notin S. \end{cases}$$

Both methods process the same actual outcomes. The complement of a qualitative binary rule contains no designated time point. A concatenation applies base methods over successive stretches of stages; its switch schedule determines which method applies (Schurz, 2026). The clock enters the concatenation. Formal membership of the complement in this base set supplies no warrant for using it on a particular occasion.

For an outcome sequence ω, let $C_f(N,\omega)$ count correct forecasts through stage N, let $a_f = C_f/N$, and define C_{g_S} and a_{g_S} in the same way. Each correct forecast receives one unit. Put

$$I_j(\omega) = \mathbf{1}\{f(x_{<j}) = x_j\}, \qquad D_S(N) = |S \cap \{1, \ldots, N\}|.$$

At an exceptional stage the two forecasts differ, so exactly one of them is correct. At every other stage the forecasts agree. Therefore

$$\begin{aligned} C_{g_S}(N,\omega) - C_f(N,\omega) &= \sum_{\substack{j \le N \\ j \in S}} (1 - 2I_j(\omega)), \\ |a_{g_S}(N,\omega) - a_f(N,\omega)| &\le \frac{D_S(N)}{N}. \end{aligned} \tag{1}$$

For a single exception $S = \{t\}$, the cumulative scores differ by exactly one in absolute value after t. The method uses f, then its complement for one forecast, then f again. Its segment boundaries $t-1$ and t are both future when it is introduced at stage m.

More generally, suppose S has zero asymptotic density: $D_S(N)/N \to 0$. Equation (1) proves, separately on every path, that the limiting average scores exist together and agree whenever they exist. Their lower and upper limits also agree on paths with no limiting average. In particular,

$$B_f = \{\omega : \lim_{N\to\infty} a_f(N,\omega) \text{ exists and exceeds } \tfrac{1}{2}\} = B_{g_S}.$$

Because the two events coincide as sets of paths, every probability law on the binary outcome space assigns them the same probability. The events are measurable in the cylinder sigma-algebra generated by finite binary histories: each finite score depends on a finite history, and the limit condition is a countable combination of such conditions. The conclusion holds for any fixed threshold. The argument introduces no probability law tailored to the forecaster: a law already justified for f

assigns the same probability to this success event for g_S. The equality holds for every zero-density S, so the number of schedules the agent could have chosen does not affect this probability. The composite retains the learner's regularity-based route because it continues using the learner on all but a vanishing fraction of stages. The unsupported timing of its departures does not remove what it learns on the remaining stages.

The mathematical extension permits exceptions that never end. Fix an arbitrary integer $t \geq m + 2$ and take

$$S_t = \{t + k^2 : k = 0, 1, 2, \ldots\}.$$

For $N \geq t$, its exception count is $1 + \lfloor \sqrt{N - t} \rfloor$, so its density tends to zero. The composite departs infinitely often and never permanently resumes f. The clock tells the agent when the composite departs and supplies no predictive support for those departures. Excluding only finitely modified methods does not resolve this case. Counting infinitely many departures is also insufficient to show loss of the learner's long-run success. If the generalized class is restricted to finitely many switches while permitting two, the single reversal already gives the correction: it returns to f and preserves its limiting-success event. A finite switch count alone does not ensure preservation when the final segment permanently replaces f.

Schurz's maximum-rule makes the result concrete. It predicts one when the observed frequency of ones exceeds one half and zero otherwise (Schurz, 2026). Set its empty-history forecast to zero. On an all-one sequence it makes one initial error. The square-schedule composite makes additional errors exactly on S_t, yet both limiting accuracies equal one. A permanent complement from t has limiting accuracy zero. Departing on every other stage has limiting accuracy one half. The density of departures distinguishes these cases; an arbitrary parameter occurs in each.

For a numerical instance of this hypothetical all-one sequence, take $m = 2$ and $t = 4$. The square schedule reverses forecasts at stages $4, 5, 8, 13, 20, \ldots$. By stage 100 it has made ten exceptions: the base has 99 correct forecasts and the variant has 89. By stage 10,000 it has made 100 exceptions: the scores are 9,999 and 9,899. The accuracy gap falls from 10/100 to 100/10,000, even though further exceptions keep occurring. A single reversal at stage four instead gives gaps of 1/100 and 1/10,000. Each comparison uses the same outcome record for both methods.

The pathwise preservation does not depend on unit correctness. For a common measurable per-stage score s in a fixed interval of length $c < \infty$, write a_h^s for method h's average score. Agreement outside S gives

$$|a_{g_S}^s(N, \omega) - a_f^s(N, \omega)| \leq cD_S(N)/N.$$

Zero density again preserves lower and upper limits, the existence and value of the limit, and the event that it exceeds any fixed threshold. This includes Schurz's scores defined as one minus a loss normalized to $[0, 1]$ (Schurz, 2023). The accuracy calculations and the threshold 1/2 above concern unit correctness.

The constructions assume deterministic binary forecasts and common actual histories. They do not concern interventions whose forecasts change the future outcomes. Zero density is sufficient for preservation; it need not be necessary for every learner and path.

3 What an epistemic exclusion can establish

The equal-record problem concerns what to predict next, the question Bakshi frames as whether object-induction is justified now (Bakshi, 2025). Equality of limiting-success events cannot answer that question: it permits a wrong or unwarranted forecast at the very stage under dispute.

Schurz's strongest surviving rationale therefore assesses the exceptional action. The reason he gives for seeking an epistemic solution, the greater short-run losses of fixed-share within human planning horizons (Schurz, 2026), likewise concerns finitely many stages. A limiting-success event cannot register such losses: changing a method's forecasts at finitely many stages leaves the event unchanged. At a finite horizon N, equation (1) bounds the gap between a_f and a_{g_S} by $D_S(N)/N$ on every path, which is at most $1/N$ for a single exception. The assessment of the exceptional action requires no lottery over possible methods. Suppose an independently justified outcome law gives

$$p = P\{f(x_{<t}) = x_t \mid x_{<t}\} > \tfrac{1}{2}$$

at an observed history of positive probability. The complement then has conditional success probability $1-p$. Under unit correctness, choosing it instead of f incurs an expected loss of $2p-1$ at that decision. The equality $B_f = B_{g_S}$ remains true. A justified local comparison can therefore reject the exceptional forecast despite equal limiting-success probabilities. The construction supplies no such law or conditional probability.

A procedural norm can require grounds for departing from any method already warranted for use. Its method-neutral form could be independent of object-induction. Applying it to f still requires an account of f's warrant and an evidential difference between following f and departing. That account must also address how the pool was determined when the pool affects the warrant. A stronger reconstruction could instead assess eligibility before selecting a base, using a general criterion against unsupported scheduled changes. Such a criterion need not assume that the selected base is already warranted. It requires its own defense, which the score calculation neither supplies nor rules out. The present argument leaves these reconstructions open. It refutes the unrestricted inference from an arbitrary parameter to loss of sustained success; it does not show that every remaining procedural exclusion depends on induction.

Schurz's final paragraph also permits a narrower target: a method permanently replaces the successful learner with a successor unrelated to its earlier success (Schurz, 2026). Neither the one-stage exception nor the square schedule satisfies that description. The square schedule rules out equating permanent abandonment with infinitely recurring departures. A restriction to a permanently unsupported successor remains available, provided its epistemic justification is established. It does not apply to every member of the generalized multiple-switch class.

4 When one comparator can represent a family

A rule determined solely by the probability of the specified limiting-success event must treat f and g_S alike under any common law. Equality of the events establishes this restriction even when the law is justified. Yet denying that particular ground for exclusion does not require an agent to track every variant. The separate question is how much a performance comparison loses when it represents variants by a retained learner. Answering it avoids making pool reduction depend on the false claim that those variants have lost their learning route.

Here a comparator is a method whose score provides a reference for assessing another forecaster. Suppose every variant can exceed the retained learner's total score by at most one. A forecaster that falls no more than five correct predictions behind that learner can fall no more than six behind any such variant. Omitting the variants from this comparison adds a known allowance. The allowance says nothing about whether deleting the variants from a forecasting algorithm would change its predictions or weights.

For a concrete performance rule, retain an independently selected base learner f. Consider a family $\mathcal{G}_f$ of variants. Suppose a known common bound $d(N)$ limits, on every outcome path, the number of forecasts at which any member differs from f. Let A be any forecaster evaluated on those same paths, with average correctness a_A. Then

$$\sup_{g \in \mathcal{G}_f} a_g(N, \omega) - a_A(N, \omega) \leq a_f(N, \omega) - a_A(N, \omega) + \frac{d(N)}{N}. \tag{2}$$

Each g can gain at most one unit over f at each differing forecast; taking the supremum proves the bound. If A's average-score deficit relative to f is at most $r(N)$, its deficit relative to the entire family is at most $r(N) + d(N)/N$.

Equation (2) yields a restricted rule for pool reduction: a retained learner may represent a family of variants as performance comparators when their common disagreement bound fits the stated score tolerance. For a long-run comparison, $d(N) = o(N)$ suffices. For a finite horizon, the additional allowance is $d(N)/N$ and must be acceptable at that horizon. All single-override variants share $d(N) = 1$. All square-schedule variants above share the bound $d(N) = 1 + \lfloor \sqrt{N} \rfloor$. Neither family needs separate tracking to preserve this conditional comparison guarantee.

Weighted averaging permits a related, conditional comparison. At stage j, abbreviate the two forecasts as f_j and g_j. Write a pool's weighted forecast as $q_j = R_j + u_j f_j + v_j g_j$, where the weights are nonnegative and sum to one, and R_j collects the other weighted forecasts. Coalesce the pair's weight onto f by defining $\bar{q}_j = R_j + (u_j + v_j) f_j$, using the weights from the same original run. Then

$$q_j - \bar{q}_j = v_j(g_j - f_j) = v_j(1 - 2f_j)\mathbf{1}\{j \in S\}.$$

The aggregates agree outside S. For any common score on their forecasts in an interval of length c, their average scores differ by at most $cD_S(N)/N$. This holds even when the original weights depend on past performance. It does not describe deleting g_S and recalculating the pool's adaptive weights. That operation can change the other weights, so the coupled comparison establishes no invariance for the recalculated aggregate. The variant can affect the disputed forecast even when the coupled averages have the same limit.

The common bound is an additional premise. Preserving each member's limiting success does not supply it. A pointwise guarantee treats each fixed member separately; a uniform guarantee uses one bound for all members at the same horizon. For every positive integer L, let a variant complement f throughout stages $t, \ldots, t + L - 1$ and resume afterward. Each variant has a finite exceptional set and preserves the limiting-success event.

For a hypothetical numerical example, set $m = 2$ and $t = 4$ again. The member with $L = 7$ departs only at stages 4 through 10. Its disagreement fraction is 7/100 at horizon 100 and 7/1,000 at horizon 1,000. But the same family includes $L = 97$ and $L = 997$, with disagreement fractions 97/100 and 997/1,000 at those respective horizons. Every rule and block length is fixed in advance. The member supplying the largest gap can change as the comparison horizon grows.

At every $N \geq t$, some member departs on every stage from t through N. The supremum disagreement fraction equals $(N - t + 1)/N$, which tends to one. On the outcome sequence defined recursively by $x_j = 1 - f(x_{<j})$, the corresponding finite-block variant gains one unit on every changed stage. The uniform score difference therefore also tends to one. Separate zero-density results for each member cannot replace a common vanishing bound for the family. Figure 1 illustrates this distinction.

Schurz's earlier fixed-share response already tracks concatenations through their base methods (Schurz, 2021). The constructions here use f and its complement as two base methods. At a

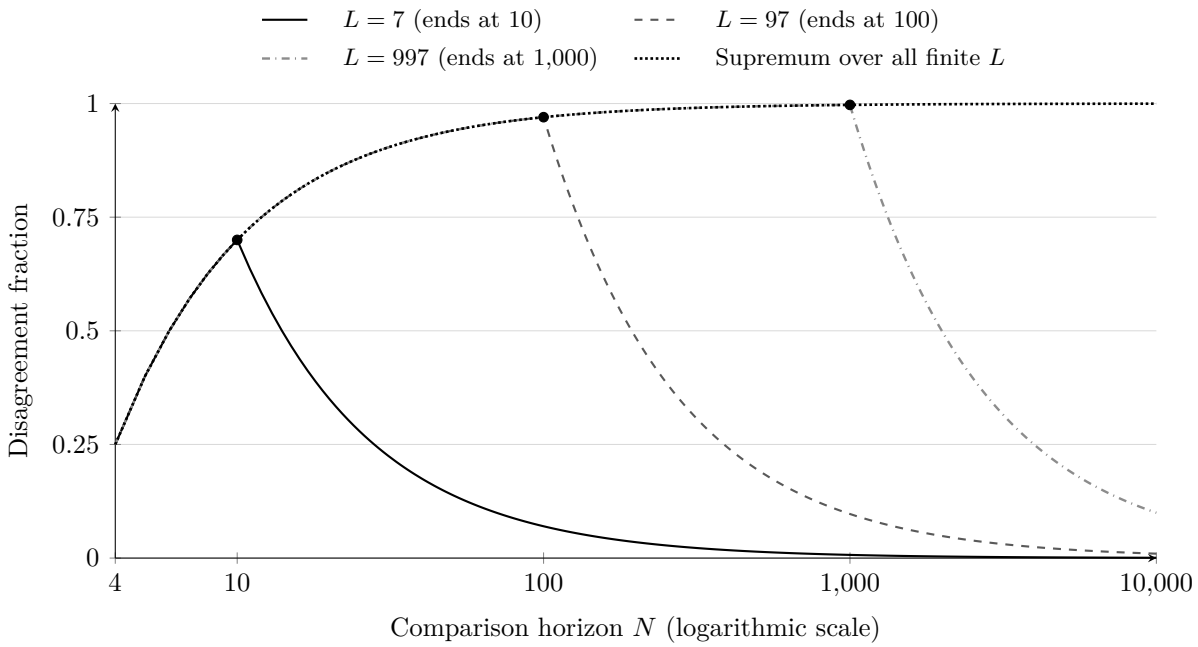


Figure 1: Disagreement fractions for variants that reverse f from stage 4 through stage $3 + L$ and copy it afterward. The three fixed members have $q_L(N) = \min(L, N-3)/N$ and each tends to zero. The dotted curve is the supremum over all positive integer L: $Q(N) = (N-3)/N$, which tends to one. At each horizon a member with $L \geq N - 3$ attains it. Points mark block endpoints. Values use integer horizons $N \geq 4$; joined lines guide the eye. These deterministic counts are independent of outcomes and are not empirical success rates.

horizon $N \geq 2$, a single override has at most two switches; every square schedule has at most $2(1+\lfloor\sqrt{N}\rfloor)$. Capping each allowance at $N-1$ gives a common sublinear switch allowance b_N. With parameters chosen for N and b_N, the standard fixed-share bound for a horizon known in advance gives a uniform upper bound on average regret that tends to zero (Cesa-Bianchi and Lugosi, 2006). Average expected regret is the randomized forecaster's expected loss per stage in excess of that of a comparator following f or its complement with at most b_N switches. The forecaster's expected loss is the weighted mean of the base methods' losses. Sampling a binary forecast from those weights gives the corresponding conditional expected loss. It does not guarantee the realized correctness score of a deterministic binary aggregate.

The same known-horizon result also handles the finite-block family uniformly. Every block has at most two switches, regardless of its length. Setting $b_N = \min\{2, N-1\}$ gives uniform average expected regret of order $\sqrt{\log N/N}$ while tracking only f and its complement (Cesa-Bianchi and Lugosi, 2006). The family therefore remains trackable despite its supremum disagreement fraction relative to f tending to one.

Equation (2) supplies a simpler conditional bound with narrower scope: any established guarantee against f transfers to a sufficiently close family at a stated cost. It neither improves fixed-share nor provides a forecasting algorithm. A common $d(N) = o(N)$ is sufficient for this representative-f guarantee. It is not necessary for uniform tracking or comparator compression in general. The finite-block family separates these claims: representation by f alone loses a vanishing uniform guarantee, whereas fixed-share tracks the family through two base methods.

5 The correction to the exclusion rationale

The elementary score bound does not resolve Bakshi's next-action objection or discover a new class of mixed methods. The bound identifies a gap in the unrestricted 2026 rationale: containing an arbitrary parameter does not establish that a method lacks a regularity-based route to sustained success. That rationale therefore does not remove the constructed finite and zero-density composites. Under every common law, their limiting-success event has the learner's probability, regardless of how many values their schedules could have taken. Schurz's concession concerns a broader class of mixed methods. The recurring extension shows that merely exempting finite modifications would still leave the same inference invalid for some infinite schedules. An elaboration that excludes these composites needs a ground other than a lost route to sustained success. A rule based on current warrant or on permanent replacement by an unsupported successor remains available, with its own justification.

The distinction matters even when different rules recommend omitting the same method. Rejection of its next forecast requires a warranted evidential or procedural premise. For the representative-f guarantee, omission as separate performance comparators uses a common error allowance appropriate to the comparison horizon. Fixed-share supplies another route to tracking without separate records for each variant. These performance guarantees do not resolve the warrant for the next action. Stating which reason supports a pool decision prevents an approximation in performance evaluation from doing unearned work in the justification of object-induction.

Acknowledgments

AI-assisted writing. I designed and directed this paper, made the editorial decisions, and edited and verified the manuscript. OpenAI GPT-6 Astra and GPT-6.1 Sol assisted with drafting and manuscript review. I take full responsibility for the paper.

References

Bakshi, Kabir S. 2025. Identification and Indetermination in the Meta-Inductive Approach to Induction. *Philosophy of Science* 92(5):1427–1437. doi:10.1017/psa.2025.10130.

Cesa-Bianchi, Nicolò, and Gábor Lugosi. 2006. *Prediction, Learning, and Games*. Cambridge University Press. doi:10.1017/CBO9780511546921.

Pitts, J. Brian. 2023. Does Meta-induction Justify Induction: Or Maybe Something Else? *Journal for General Philosophy of Science* 54:393–419. doi:10.1007/s10838-022-09620-7.

Schurz, Gerhard. 2021. Metainduction over Unboundedly Many Prediction Methods: A Reply to Arnold and Sterkenburg. *Philosophy of Science* 88:320–340. doi:10.1086/711587.

Schurz, Gerhard. 2023. In Search for Optimal Methods: New Insights About Meta-Induction. *Journal for General Philosophy of Science* 54:491–522. doi:10.1007/s10838-023-09649-2.

Schurz, Gerhard. 2026. Object-Induction and Pool-Determination in the Meta-Inductive Approach. A Reply to Bakshi. *Philosophy of Science*, First View, 1–10. doi:10.1017/psa.2026.10247.