Enabling Robust Epidemic Control via LLM-Elicited Causal Discovery

Jinwoo Go [1] **Xiaoning Qian** [1 2 3] **Byung-Jun Yoon** [1 2]

Abstract

Large language models (LLMs) can facilitate scientific prior elicitation, but it remains unclear how such priors might enhance high-stakes control under uncertainty. We study this in a zero-current-data epidemic-control setting on a deterministic compartmental epidemic model benchmark with five non-pharmaceutical intervention levers. Our LLM co-scientist pipeline elicits a consensus DAG, constructs an LLM-elicited sample pool, fits an LLM prior, and exposes the resulting artifacts to model-predictive control (MPC). Against a Uniform-prior MPC baseline and a literature prior, the LLM prior plays a bounded but useful role: its open-loop mode reduces cumulative infections by 25.8% relative to the content-free $\mathrm{Beta}(1, 1)$ (Uniform) baseline, but the intervention strengths it recommends inside the controller are less well calibrated than those of the literature-based prior. Keeping the full pool of LLM samples inside the controller works better than collapsing it into a single point prior. Disagreement among the samples then becomes an explicit measure of regime uncertainty. These results suggest that co-scientist systems are useful for prior structure and scenario coverage, while robust control should ultimately update the LLM prior with real observations and propagate posterior uncertainty through MPC.

1. Introduction

Epidemic control is a high-stakes decision task under structural uncertainty. A decision maker must choose non-pharmaceutical interventions (NPI) whose effects depend on unobserved parameters, behavioral response, and a dynamical system whose causal structure is itself uncertain. Classical model-predictive control (MPC) methods for epidemic mitigation usually assume the compartmental structure is given and recover parameters online (Carli et al., 2020; Köhler et al., 2021). The harder question is decided before control begins: which causal drivers of the effective contact rate should be controllable, and in what relative order they should matter.

Literature-derived priors compress historical epidemic evidence into sharp distributions over intervention magnitudes. This is useful when the new regime resembles past COVID-19, influenza, or respiratory-virus waves, but can be brittle when a novel variant, compliance collapse, or large social disruption pushes coefficients into regions that the literature prior assigns little mass. We therefore use large language models (LLMs) not as replacements for the literature, but as scenario-conditioned samplers. Given a curated causal graph and a regime description, the LLM proposes plausible component values for the infection-rate equation, including values from disrupted regimes that may not yet appear in data.

Table 1. Capability matrix across representative LLM-for-causal, LLM-SCM, AI co-scientist, and epidemic-control works (2022–2026). **DAG**: LLM proposes graph structure over domain variables. **SE**: LLM serves as structural equation at each node. **Surr**: tractable distilled surrogate exposed for audit. **MPC**: closed-loop control over intervention levers. **Gate**: explicit human-in-the-loop review artifact. **Epi**: evaluated on an epidemic task. ✓ full, ✓* partial, – absent.

Method (year)	DAG	SE	Surr	MPC	Gate	Epi
Darvariu et al. (2024)	✓	–	–	–	–	–
Kıcıman et al. (2024)	✓	–	–	–	–	–
Takayama et al. (2025)	✓	–	–	–	–	–
Du et al. (2025)	✓	–	–	–	–	–
Bynum and Cho (2025)	–	✓	–	–	–	–
Lu et al. (2024)	–	✓*	–	–	✓*	–
Gottweis et al. (2025)	–	✓*	–	–	✓	–
Du et al. (2024)	–	–	✓	–	–	✓
Polcz et al. (2025)	–	–	✓	✓	–	✓
Hewing et al. (2020)	–	–	✓	✓	–	✓
Mascaro et al. (2023)	✓*	–	–	–	–	✓
Ours	✓	✓	✓	✓	✓	✓

Accepted at ICML 2026 AI for Science Workshop. [1]Computing & Data Sciences, Brookhaven National Laboratory, Upton, NY, USA [2]Department of Electrical & Computer Engineering, Texas A&M University, College Station, TX, USA [3]Department of Computer Science & Engineering, Texas A&M University, College Station, TX, USA. Correspondence to: Byung-Jun Yoon <byoon@bnl.gov>.

Proceedings of the 43rd International Conference on Machine Learning, Seoul, South Korea. PMLR 306, 2026. Copyright 2026 by the author(s).

Recent work shows that LLMs can provide causal judgments (Kıcıman et al., 2024), edge priors for causal discovery (Darvariu et al., 2024; Takayama et al., 2025; Du et al., 2025), and structural equations at DAG nodes (Bynum and Cho, 2025). AI co-scientist systems automate larger parts of the research loop (Lu et al., 2024; Gottweis et al., 2025), while epidemic MPC and forecasting systems optimize or predict under mostly fixed model structure (She et al., 2022; Polcz et al., 2025; Du et al., 2024; Mascaro et al., 2023). What remains unclear is whether LLM prior content transfers into closed-loop decisions, which part transfers, and how the samples should be consumed. Our answer is to make the co-scientist contribution phase-separated and auditable, so the graph, component samples, surrogate weights, and MPC trajectories can be inspected separately.

Table 1 summarizes the closest work by capability. Prior LLM causal-discovery and structural causal model (LLM-SCM) methods provide graph or mechanism priors, AI co-scientist systems automate broader research workflows, and epidemic-control methods close the loop with MPC or forecasting. None combines LLM-elicited DAG construction, LLM-elicited structural-equation specification, surrogate distillation, closed-loop MPC, a human review gate, and epidemic evaluation in one co-scientist pipeline.

We study this question in the zero-current-data setting, where no observations from the current epidemic regime are available for posterior updating. We compare an LLM prior with a literature-based prior and a content-free Uniform baseline on a deterministic Susceptible-Infectious-Quarantined-Recovered benchmark (Hethcote et al., 2002; Odagaki, 2020) with five NPI control levers. The Uniform-prior MPC keeps the regularizer form but removes prior content, so performance differences isolate what the LLM or literature prior adds.

The results show a bounded but useful role for the LLM prior. In open loop, the LLM prior mode reduces cumulative infections by 25.8% relative to Uniform, while the literature prior mode reduces them by 91.3%. Inside MPC, replacing the Uniform Beta regularizer with the LLM prior improves cumulative infections from 382.18 to 313.66 per $1,000$, while the literature prior reaches 40.49. A scenario-softmin controller that consumes the same LLM sample pool without collapsing it reaches 140.07 ± 99.36, suggesting that the LLM samples may be more effective when used as a scenario ensemble rather than collapsed into a single point density.

Contributions. Our contributions are as follows.

1. **A six-phase, fully inspectable co-scientist prior-elicitation pipeline:** DAG drafting; human-curated consensus; scenario-conditioned component sampling; structural-equation fitting; surrogate distillation; and prior-regularized MPC.
2. **A source-matched comparison of four prior-usage modes** (Table 2): a Uniform $\mathrm{Beta}(1,1)$ form-only baseline, a point prior from the LLM, a point prior from the literature, and a scenario-ensemble controller that retains the LLM sample diversity without collapsing it into a density.
3. **An empirical decomposition of nominal accuracy versus scenario coverage** (Section 3). The literature prior gives the best nominal performance in the zero-current-data setting, but the LLM samples are more useful as a dispersed set of plausible regimes than as a collapsed point prior. Consuming them as a scenario ensemble lets the controller hedge against regime uncertainty and recover much of the degraded performance.

2. Methods

Pipeline overview. Our co-scientist pipeline has six auditable phases (Figure 1). DAG drafting $\rightarrow$ human-curated consensus $\rightarrow$ scenario-conditioned component sampling $\rightarrow$ structural-equation fit for $\beta(\cdot) \rightarrow$ linear surrogate distillation $\rightarrow$ prior-regularized MPC. Each phase produces an artifact that a reviewer can inspect. A linear surrogate (Section 2) fit on roll-outs of this co-scientist pipeline serves as the audit interface between the DAG and the controller.

Prior sources. We compare three prior sources throughout. The *literature prior* is a Beta distribution constructed from published epidemiological estimates of NPI effect sizes. The *LLM sample pool* contains 80 LLM-elicited, scenario-conditioned samples under the consensus DAG, with prompts designed to cover both nominal and disrupted regimes. The *LLM prior* is the per-lever Beta density fit to this sample pool. The *Uniform prior* is $\mathrm{Beta}(1,1)$ on each lever and serves as a content-free baseline that keeps the regularizer form while removing prior information. Control actions u_t are optimized by MPC at each time step and are not themselves prior samples.

SIQR dynamics. We adopt the standard Susceptible–Infectious–Quarantined–Recovered (SIQR) compartmental model (Hethcote et al., 2002; Odagaki, 2020) with a five-dimensional normalized control $u_t \in [0,1]^5$ that modulates the effective contact rate through the linear surrogate (Equation (2)). Full equations and hyperparameters are in Appendix C.

Phase 1: DAG drafting. We prompt the LLM to propose a DAG over a fixed set of SIQR-relevant variables covering covariates of the contact rate, mitigation response, and reporting lags. We draw $K{=}10$ candidate DAGs across

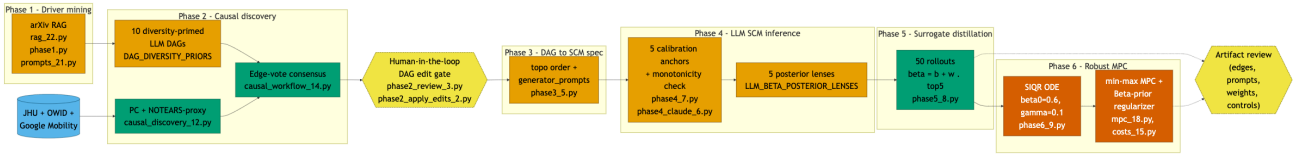

Figure 1. Six-phase co-scientist pipeline: Phase 1 driver mining; Phase 2 causal discovery with edge-vote consensus and an explicit human-in-the-loop DAG edit gate; Phase 3 DAG-to-SCM specification; Phase 4 LLM SCM inference with calibration samples and monotonicity checks; Phase 5 surrogate distillation; and Phase 6 robust MPC with the Beta-prior regularizer. Yellow nodes are co-scientist artifacts, green nodes are auditable objects distilled from those artifacts, and the hexagonal artifact-review node materializes the edges, prompts, weights, and controls exposed for audit.

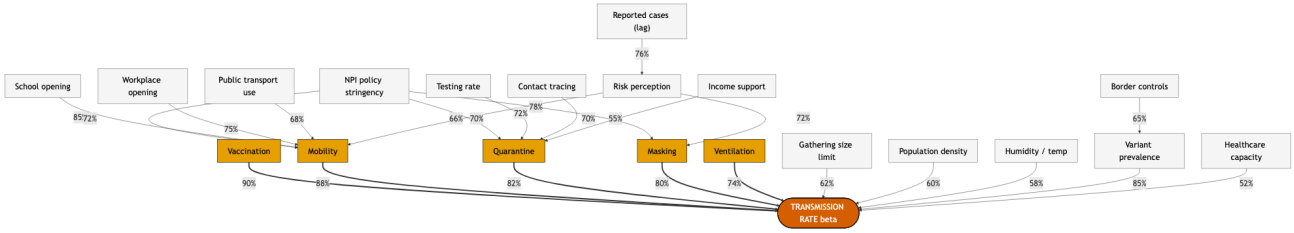

Figure 2. Consensus causal DAG over infection-rate drivers produced by Phase 2 (edge-vote across K LLM-drafted DAGs, followed by the human-in-the-loop edit gate). Five orange leaf nodes (*Vaccination, Mobility, Quarantine, Masking, Ventilation*) are the MPC control levers $u \in [0,1]^5$, and the remaining nodes are exogenous context variables. Edge labels show retention rates (%) across Phase 2 sampling rounds. Only edges above the consensus threshold survive into Phase 3. This is the object that makes Phase 2 (DAG elicitation) separately inspectable from Phases 3–5.

independent queries at `temperature`=0.8. Prompts are reproduced in Appendix G.

Phase 2: Human-curated consensus. Candidate DAGs are aggregated into an edge-frequency table. A human reviewer inspects the table and selects a single consensus DAG $G^\star$. Edges with support of at least 0.5 across the K candidates are accepted by default. Edges with lower support are accepted only if the reviewer records a written justification based on domain knowledge or prior literature. Figure 2 visualizes $G^\star$ with retained edge frequencies. The five orange leaves correspond to the MPC control levers $u \in [0,1]^5$.

Phase 3: Scenario-conditioned component sampling. Conditioned on $G^\star$, we prompt the LLM to emit a structural-equation specification for each component of the infection-rate equation. The specification includes a functional form, coefficient ranges, per-coefficient Beta summaries, and interaction signs. For each lever $k \in \{1,\dots,5\}$, we collect 80 samples and fit an independent activation prior $\mathrm{Beta}(\alpha_k, \beta_k)$. The prompt asks for plausible parameter values under both familiar epidemic regimes and disrupted regimes outside the historical literature. The same LLM sample pool is used to construct the LLM point prior and the Scenario-softmin ensemble. The fitted per-lever Beta parameters are listed in Appendix F.

Phase 4: Structural-equation fit. We fit a structural equation to the component pool, yielding

$$\beta(x) \;=\; f_{\boldsymbol{\theta}}\big(\phi(x);\, G^\star\big), \tag{1}$$

where $\phi(x)$ are the DAG-consistent features and $\boldsymbol{\theta}$ are coefficients estimated from the LLM component samples.

The fit is deterministic given the component samples, so stochasticity in $\beta(\cdot)$ is traceable to the upstream LLM sampling phase.

Phase 5: Linear surrogate. For MPC gradients and weight-level audit, we fit a linear surrogate

$$\hat{\beta}(x) \;=\; b \;+\; \mathbf{w}^\top \phi_{1:d}(x), \tag{2}$$

on N=50 pipeline rollouts using the top-d features screened against $G^\star$. We report per-weight bootstrap confidence intervals ($B = 1000$) so that each w_i can be compared directly with the corresponding DAG edge.

MPC objective. For rollout horizon H, we use the objective

$$\mathcal{J} \;=\; \sum_{t=0}^{H-1} \Big[\, \ell_{\mathrm{inf}}(u_t) \;+\; \alpha\|u_t\|^2 + \;\lambda\, \mathrm{NLL}_{\mathrm{prior}}(u_t) \;+\; \mu\|\Delta u_t\|^2 \Big], \tag{3}$$

with selected values λ=0.1 and μ=1 chosen from the grid in Appendix F.

The prior negative-log-likelihood $\mathrm{NLL}_{\text{prior}}(u_t) = \sum_{i=1}^{5} -\log \mathrm{Beta}(u_t^{(i)}; \alpha_i, \beta_i)$ is a per-lever soft penalty. The lever-specific (α_i, β_i) are the Phase 4 summaries of the Phase 3 LLM component samples and are reported in Appendix F. Setting λ=0 removes the prior-penalty pull. This recovers an unregularized controller.

Reproducibility protocol. The SIQR simulator and MPC optimizer are deterministic given the co-scientist artifacts. The only stochastic phases are DAG drafting and component sampling; we quantify their variability with DAG-edge Jaccard across three independent rounds in Section 3. Full conditions, seeds, and prompts are listed in Appendices C and G.

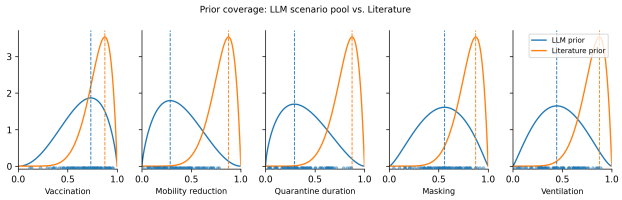

Figure 3. Per-lever prior coverage. Blue curve shows the LLM prior (fitted Beta density) and blue rug marks show the 80 samples drawn from Phase 3. Orange curve shows the literature prior. Dashed vertical lines mark the prior modes used by Fixed-PriorMode and prior-MPC rows in Table 2. The literature prior concentrates mass near high activation on all five levers. The LLM prior is broader by design and its modes sit lower on Mobility, Quarantine, and Ventilation. This breadth is the mechanism through which the LLM prior acquires scenario coverage at the cost of nominal accuracy.

3. Experimental Results

Setup. All closed-loop experiments use the SIQR dynamics (Hethcote et al., 2002; Odagaki, 2020) with step Δt=1 day, horizon H=60 days, and MPC lookahead 10 days. The controller acts on five normalized levers $u_t \in [0,1]^5$: Vaccination, Mobility reduction, Quarantine duration, Masking, and Ventilation. These levers are the top-5 causal parents of the infection rate in the consensus DAG $G^\star$.

We organize the comparison along two axes: *prior source* (Uniform, LLM, literature) and *prior usage* (open-loop mode, point density inside MPC, scenario ensemble inside MPC), giving seven conditions. The four closed-loop MPC rows use the shared operating point λ=0.1, μ=1 chosen from the grid in Appendix F. The simulator, MPC inner loop, and static Beta tables are deterministic, so every row except Scenario-softmin is a single-seed point estimate. Scenario-softmin is reported as mean $\pm$ std over seeds $\{0,1,2\}$ because each seed draws a different K=10 subset of the 80-sample LLM pool. Per-condition definitions and the scenario-softmin formula (Equation (4)) are in Appendix A.

Headline finding. In the zero-current-data setting, the literature prior wins on the nominal benchmark, but collapsing the LLM sample pool into a point prior discards its coverage value. Retaining the full pool via scenario-ensemble consumption recovers a large fraction of that value at the cost of higher variance.

Main results. Figure 3 shows the prior shapes before either enters the controller. Table 2 organizes the results by prior source and usage mode. Three findings stand out.

Open-loop ranking. Moving from Fixed-Uniform (456.72) to Fixed-PriorMode-LLM (338.93) cuts cumulative infections by 25.8% without any optimizer. Fixed-PriorMode-Literature pushes this further to 39.74, a 91.3% reduction. The ranking signal is real and transfers across prior sources. In the nominal benchmark regime the literature prior dominates because it is calibrated to historical epidemic waves.

Closed-loop magnitudes.

Uniform-prior MPC fixes the regularizer to Beta(1,1) and scores 382.18. Replacing the Uniform Beta with the LLM prior at the same (λ, μ) takes this to 313.66, an improvement of 68.52 per 1,000. The literature prior takes it to 40.49, an improvement of 341.69. The LLM magnitudes help inside MPC, but the gap to the literature prior is roughly half an order of magnitude. This establishes a clear nominal-regime gap between LLM prior magnitudes and literature-calibrated magnitudes.

Scenario-ensemble row.

Scenario-softmin MPC consumes the same 80-sample LLM pool under a seed-drawn K=10 subset rather than collapsing it into a point density. It lands at 140.07 ± 99.36 cumulative infections over seeds $\{0,1,2\}$, between the LLM and literature closed-loop rows. Because the simulator is deterministic, this variance reflects disagreement among sampled policies rather than rollout noise. Representative trajectories are visualized in Figure 7 of Appendix C.1.

3.1. When the LLM is load-bearing and when it does not

With Uniform-prior MPC in the table, the closed-loop performance decomposes into two pieces: the contribution of the regularizer form and the contribution of the prior Beta content. The form-only baseline is anchored at 382.18, and the row-to-row deltas in the Main result paragraph isolate the content piece. The roughly 5$\times$ gap between LLM and literature content deltas reflects the difference between a prior calibrated to the nominal benchmark regime and one designed to cover a wider scenario space. The LLM's lever ordering is already useful at Phase 2, while the remaining gap at Phase 5 reflects the cost of scenario breadth rather

Table 2. Closed-loop performance on the SIQR benchmark (H=60, Δt=1 day, 5-D control). Rows separate prior source and prior usage. Open-loop rows and Beta-prior closed-loop rows are deterministic point estimates at seed= 0. Scenario-softmin MPC reports mean ± std over seeds $\{0, 1, 2\}$ because each seed draws a different size-K=10 subset of the 80-sample LLM pool. Cum. Cost and Prior NLL are objective-function values. Bold marks the best closed-loop value among methods at matched (λ, μ)=(0.1, 1).

Method	Cum. Inf. (/ 1,000)	Cum. Cost	Peak $I+Q$	Mean $\|u\|_1/5$	Prior NLL
Open loop (no MPC)					
Fixed-Uniform	456.72	+5.74	0.151	0.400	−104.81
Fixed-PriorMode-LLM	338.93	+4.04	0.104	0.463	−162.78
Fixed-PriorMode-Literature	39.74	+0.61	0.012	0.670	−454.02
Closed loop (matched optimizer, surrogate, control space, λ=0.1, μ=1)					
Uniform-prior MPC	382.18	+6.40	0.140	0.432	+0.00
LLM-prior MPC	313.66	−11.55	0.093	0.473	−162.13
Literature-prior MPC	**40.49**	−42.89	**0.013**	0.669	−453.88
Scenario-softmin MPC (σ=0.15, K=10)	140.07±99.36	+15.53±3.30	0.041±0.027	0.632±0.144	+118.44±28.34

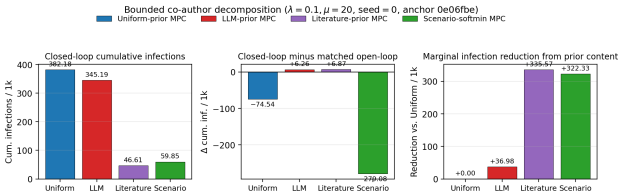

Figure 4. Decomposition of closed-loop performance into regularizer-form and prior-content contributions, at matched (λ, μ)=(0.1, 1). **Left:** Cumulative infections for the four closed-loop methods. Uniform-prior MPC (382.18) anchors the regularizer-form-only contribution. LLM-prior MPC (313.66) and Literature-prior MPC (40.49) measure the marginal effect of replacing Uniform Beta content with LLM and literature content respectively. **Center:** Closed-loop minus matched-open-loop cumulative infections per 1,000. LLM-prior MPC actually beats its static reference by 25.27 once feedback is allowed, while Literature-prior MPC is essentially tied with its static reference at +0.75. **Right:** Marginal infection reduction over Uniform-prior MPC attributable to prior content alone. LLM content reduces cumulative infections by 68.52 per 1,000, while literature content reduces them by 341.69, a roughly half-order-of-magnitude gap that the bounded contributor thesis makes explicit.

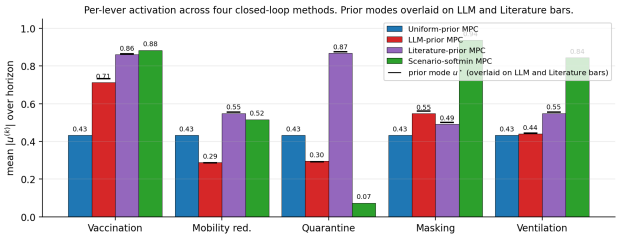

Figure 5. Per-lever mean activation $\overline{|u^{(k)}|}$ over the 60-day horizon for the four closed-loop methods, with the LLM and Literature prior modes overlaid as black dashes on the corresponding bars. Uniform-prior MPC and LLM-prior MPC track lower magnitudes because their content-bearing modes are themselves lower than the surrogate's free optimum. Literature-prior MPC tracks the literature mode lever-by-lever. Scenario-softmin departs from any single mode because the ensemble removes the point-density target.

than a failure of the elicitation procedure.

When the prior helps and when it does not. The per-lever activation averages clarify the mechanism. Literature-prior MPC and Scenario-softmin place high activation on Vaccination and Ventilation, matching the two high-weight levers in the surrogate. LLM-prior MPC and Uniform-prior MPC use lower average activation, consistent with weaker magnitude information in the point prior. Thus lever ordering explains the open-loop ranking gain, while lever magnitudes determine how much the closed-loop optimizer can exploit the prior once the objective form is fixed. Per-lever means and prior modes are shown in Figure 5.

Surrogate uncertainty. The audit depends on the linear surrogate being accurate enough to interpret. The surrogate achieves train R^2=0.88 and held-out R^2=0.85 with RMSE 0.02 (Figure 6). All five weights have the expected negative sign, and Masking and Vaccination are the two largest in magnitude. Bootstrap intervals are shown in Figure 6 and tabulated in Table 4.

Stability of the consensus DAG. Phases 1 and 3 are the LLM-driven, non-deterministic parts of the pipeline. We rerun Phases 1–5 under three independent elicitation rounds and measure DAG-edge Jaccard between pairs of rounds. Pairwise Jaccard on the merged edge sets is J_{12}=0.800, J_{13}=0.762, J_{23}=0.857 (mean 0.806). Restricting to the direct parents of β raises this to J_{12}=0.938, J_{13}=0.882, J_{23}=0.941 (mean 0.920).

The top-5 decision-relevant factors are less stable (J-mean 0.587), but the two surrogate-heaviest levers are robust: Masking appears as a direct β-parent in 30/30 candidates and Vaccination in 29/30. Thus the exact edge set varies, but the dominant decision-relevant structure is stable. Per-round edge-support tables are in Appendix E.

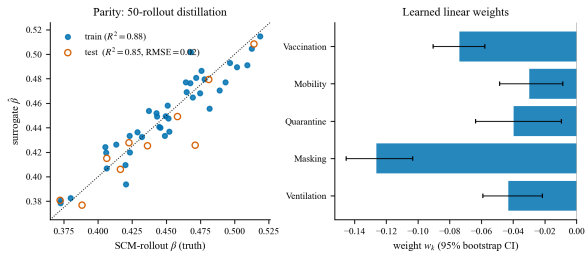

Figure 6. Surrogate distillation (Phase 5). **Left:** parity plot between the SCM-rollout β ground truth and the surrogate $\hat{\beta}$ prediction on 50 held-out rollouts; train R^2=0.88, held-out R^2=0.85 with RMSE 0.02. **Right:** learned linear weights w_k with 95% bootstrap CIs (B=1000). All weights are negative as expected (each lever reduces transmission) and the ranking $|w_{\text{Masking}}| > |w_{\text{Vaccination}}| > |w_{\text{Ventilation}}| > |w_{\text{Quarantine}}| > |w_{\text{Mobility}}|$ is the object audited in Section 3.

4. Discussion

Why the literature prior wins on the nominal benchmark. The nominal SIQR benchmark lies close to the regime for which the literature prior is calibrated. Its Beta modes are high on the same levers that the surrogate weights identify as most effective, especially Vaccination and Ventilation. As a result, Literature-prior MPC is regularized toward actions that are already close to the surrogate's low-transmission optimum. The LLM prior recovers the qualitative ordering of important levers, but its modes are lower on these high-weight levers, so the closed-loop controller under-intervenes relative to the literature prior. We therefore interpret the gap as a calibration gap in magnitudes, not as a failure of causal ranking.

Where the LLM prior adds value. The LLM prior should not be read as a cheaper substitute for a well-calibrated epidemiological prior. Its role is different: it supplies structured support over regimes that may be absent from the literature. Novel variants, abrupt compliance changes, reporting disruptions, or conflict-driven mobility changes can move intervention effects outside the high-mass region of a sharp literature prior. In a zero-current-data setting, such a prior can be confidently wrong before observations arrive to correct it. The LLM sample pool is useful precisely because it keeps multiple plausible regimes reachable.

How to consume LLM priors. The results suggest that collapsing LLM samples into a single point-density prior is the least attractive use of the elicited pool. A point Beta summary is easy to optimize against, but it hides disagreement among samples and converts scenario breadth into a single magnitude target. Scenario-softmin preserves this disagreement and lets the controller reduce the ensemble under the current state rather than before optimization begins. When no literature prior is available, the safer default is therefore to expose the sample pool to the controller, while reporting the induced variance as uncertainty rather than treating it as noise. Phase 2 (consensus DAG curation) is a human-in-the-loop gate, and we treat it as a load-bearing design choice rather than a bottleneck to be automated away.

Future directions. Our present study deliberately stops at the zero-current-data setting. The natural next step is to pair the pipeline with collaborators who can provide real outbreak, mobility, hospitalization, or intervention-compliance observations. In that setting, the LLM prior is only the initial distribution. As observations arrive, Bayes' rule updates it into a posterior over transmission and intervention-effect parameters,

$$p(\theta \mid y_{1:t}) \propto p(y_{1:t} \mid \theta)\, p_{\text{LLM}}(\theta).$$

The control problem should then be posed over posterior uncertainty rather than over a fixed prior, so that MPC decisions become robust to what has and has not yet been learned from data.

This raises a second question: which samples actually affect the controller? Posterior samples may change the MPC policy through the infection forecast, the prior penalty, the scenario ensemble, or the risk measure used in the objective. Tracing these pathways would show when sample diversity improves robust decision making and when it is washed out by the optimizer. Warm-starting from LLM modes, disrupted-regime probes, and state-dependent (λ, μ) schedules are useful ablations, but the broader goal is a prior-to-posterior-to-control pipeline that connects LLM elicitation, empirical observations, and robust epidemic decisions.

5. Limitations

No current-regime observations. This study deliberately evaluates the zero-current-data setting. We do not condition on real outbreak, mobility, hospitalization, compliance, or intervention-response observations from the target regime. The reported priors are therefore elicited or literature-derived initial distributions, not posteriors. In deployment-facing work, these priors should be updated with observations through Bayes' rule before being used for robust decision making. Collaborations with groups that hold such data are necessary to test whether the LLM-elicited support remains useful after posterior updating.

No posterior-control loop. Our MPC controller consumes either point Beta summaries or scenario samples from the elicited prior. It does not yet propagate posterior samples through the SIQR dynamics or optimize against posterior predictive uncertainty. As a result, we can measure whether LLM samples improve the fixed-prior controller,

but we cannot yet claim that they produce calibrated uncertainty under real observations. A full posterior-to-control pipeline would need to identify which samples change the MPC policy, which are ignored by the optimizer, and how uncertainty should affect the intervention schedule.

Sample size and LLM stochasticity. The linear surrogate is trained on $N{=}50$ pipeline rollouts, with cross-validation and bootstrap confidence intervals. The closed-loop numbers in Table 2 are deterministic given the consensus DAG and Phase 5 surrogate, but the upstream Phase 1 and Phase 3 LLM calls are stochastic. We report DAG-edge Jaccard across three elicitation rounds in Section 3, but we do not claim prompt-level determinism or full coverage of LLM sampling variability.

Reproducibility Statement

The SIQR dynamics, MPC objective, solver settings, hyperparameters, and random-seed protocol are described in Appendices C and D. The prompts used at Phases 1 and 3 are reproduced in Appendix G. The reported closed-loop experiments use deterministic solvers except for the explicitly reported LLM sampling and Scenario-softmin subset draws.

Ethics Statement

This work studies LLM-assisted epidemic control entirely in simulation. The SIQR benchmark and the five-lever NPI vector are research abstractions, and none of the reported numbers correspond to a deployed policy or a real population. We do not claim counterfactual validity against historical epidemic waves and explicitly flag this as the primary follow-up in Section 5. Any real-world use of an LLM prior or controller regularized by an LLM prior in epidemic policy would require clinical and public-health oversight that this system does not provide.

The pipeline is also designed against a specific failure mode: an LLM that drafts a controller objective without leaving an inspectable trail. We caution against removing this gate in deployment-facing variants of the pipeline, since the consensus DAG and LLM prior are exactly the artifacts a downstream reviewer needs to sign off on. The dual-use surface of automated policy recommendation, including the risk that a confident but poorly calibrated LLM prior is mistaken for a literature prior, is one of the reasons we report the LLM-versus-literature magnitude gap as a central finding rather than as a caveat.

References

Lucius E. J. Bynum and Kyunghyun Cho. Language models as causal effect generators. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2025. URL https://aclanthology.org/2025.emnlp-main.107/. Introduces sequence-driven structural causal models (SD-SCMs).

Raffaele Carli, Graziana Cavone, Nicola Epicoco, Paolo Scarabaggio, and Mariagrazia Dotoli. Model predictive control to mitigate the COVID-19 outbreak in a multi-region scenario. *Annual Reviews in Control*, 50:373–393, 2020. doi: 10.1016/j.arcontrol.2020.09.005. URL https://doi.org/10.1016/j.arcontrol.2020.09.005.

Victor-Alexandru Darvariu, Stephen Hailes, and Mirco Musolesi. Large language models are effective priors for causal graph discovery, 2024. URL https://arxiv.org/abs/2405.13551.

Hongru Du, Jianan Zhao, Yang Zhao, Shaochong Xu, Xihong Lin, Yiran Chen, Lauren M. Gardner, and Hao Frank Yang. Advancing real-time pandemic forecasting using large language models: A COVID-19 case study, 2024. URL https://arxiv.org/abs/2404.06962.

Huaming Du, Yujia Zheng, Baoyu Jing, Yu Zhao, Gang Kou, Guisong Liu, Tao Gu, and Hongtu Song. Causal discovery through synergizing large language model and data-driven reasoning. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '25)*, 2025. doi: 10.1145/3711896.3736874. URL https://doi.org/10.1145/3711896.3736874.

Juraj Gottweis, Wei-Hung Weng, Alexander Daryin, Tao Tu, Anil Palepu, Petar Sirkovic, Artiom Myaskovsky, Felix Weissenberger, Keran Rong, Ryutaro Tanno, Khaled Saab, Dan Popovici, Jacob Blum, Fan Zhang, Katherine Chou, Avinatan Hassidim, Burak Gokturk, Amin Vahdat, Pushmeet Kohli, Yossi Matias, Andrew Carroll, Kavita Kulkarni, Nenad Tomasev, Yuan Guan, Vikram Dhillon, Eeshit Dhaval Vaishnav, Byron Lee, Tiago R. D. Costa, José R. Penadés, Gary Peltz, Yunhan Xu, Annalisa Pawlosky, Alan Karthikesalingam, and Vivek Natarajan. Towards an AI co-scientist, 2025. URL https://arxiv.org/abs/2502.18864.

Herbert Hethcote, Zhien Ma, and Shengbing Liao. Effects of quarantine in six endemic models for infectious diseases. *Mathematical Biosciences*, 180 (1–2):141–160, 2002. doi: 10.1016/S0025-5564(02)00111-6. URL https://doi.org/10.1016/S0025-5564(02)00111-6.

Lukas Hewing, Kim P. Wabersich, Marcel Menner, and Melanie N. Zeilinger. Learning-based model predictive control: Toward safe learning

in control. *Annual Review of Control, Robotics, and Autonomous Systems*, 3:269–296, 2020. doi: 10.1146/annurev-control-090419-075625. URL https://doi.org/10.1146/annurev-control-090419-075625. Canonical survey of learning-based MPC; cited for the LB-MPC category in the capability matrix.

Emre Kıcıman, Robert Ness, Amit Sharma, and Chenhao Tan. Causal reasoning and large language models: Opening a new frontier for causality. *Transactions on Machine Learning Research*, 2024. URL https://arxiv.org/abs/2305.00050. arXiv:2305.00050.

Johannes Köhler, Lukas Schwenkel, Anne Koch, Julian Berberich, Patricia Pauli, and Frank Allgöwer. Robust and optimal predictive control of the COVID-19 outbreak. *Annual Reviews in Control*, 51:525–539, 2021. doi: 10.1016/j.arcontrol.2020.11.002. URL https://doi.org/10.1016/j.arcontrol.2020.11.002.

Chris Lu, Cong Lu, Robert Tjarko Lange, Jakob Foerster, Jeff Clune, and David Ha. The AI scientist: Towards fully automated open-ended scientific discovery, 2024. URL https://arxiv.org/abs/2408.06292.

Steven Mascaro, Yue Wu, Owen Woodberry, Erik P. Nyberg, Ross Pearson, Jessica A. Ramsay, Ariel O. Mace, David A. Foley, Thomas L. Snelling, and Ann E. Nicholson. Modeling COVID-19 disease processes by remote elicitation of causal Bayesian networks from medical experts. *BMC Medical Research Methodology*, 23(1):76, 2023. doi: 10.1186/s12874-023-01856-1. URL https://doi.org/10.1186/s12874-023-01856-1.

Takashi Odagaki. Exact properties of SIQR model for COVID-19. *Physica A: Statistical Mechanics and its Applications*, 564:125564, 2020. doi: 10.1016/j.physa.2020.125564. URL https://doi.org/10.1016/j.physa.2020.125564.

Péter Polcz, István Z. Reguly, Kálmán Tornai, János Juhász, Sándor Pongor, Attila Csikász-Nagy, and Gábor Szederkényi. Smart epidemic control: A hybrid model blending ODEs and agent-based simulations for optimal, real-world intervention planning. *PLOS Computational Biology*, 21(5):e1013028, 2025. doi: 10.1371/journal.pcbi.1013028. URL https://doi.org/10.1371/journal.pcbi.1013028. Introduces LUT-based $\beta \to$NPI mapping inside an MPC loop.

Baike She, Shreyas Sundaram, and Philip E. Paré. A learning-based model predictive control framework for real-time SIR epidemic mitigation. In *2022 American Control Conference (ACC)*, pages 2565–2570, 2022. doi: 10.23919/ACC53348.2022.9867851. URL https://doi.org/10.23919/ACC53348.2022.9867851.

Masayuki Takayama, Tadahisa Okuda, Thong Pham, Tatsuyoshi Ikenoue, Shingo Fukuma, Shohei Shimizu, and Akiyoshi Sannai. Integrating large language models in causal discovery: A statistical causal approach. *Transactions on Machine Learning Research*, 2025. URL https://arxiv.org/abs/2402.01454. arXiv:2402.01454.

A. Per-condition Definitions and Scenario-softmin Formula

The seven conditions in Table 2 are organized along two axes (prior source and prior usage). The three open-loop rows are **Fixed-Uniform** (with $u_t \equiv (0.4, 0.4, 0.4, 0.4, 0.4)$ for all t, ignoring both DAG and prior), **Fixed-PriorMode-LLM** (with u_t held at the LLM Phase 4 prior mode $(0.733, 0.286, 0.292, 0.560, 0.444)$, using LLM ranking without an optimizer), and **Fixed-PriorMode-Literature** (same construction with the literature-based mode vector reported in Table 3).

The four closed-loop rows all share the same optimizer and the same operating point λ=0.1, μ=1. **Uniform-prior MPC** replaces the Beta prior with Beta(1,1), so the regularizer becomes a constant and the controller sees no prior content (form-only baseline). **LLM-prior MPC** exposes the LLM Phase 4 Beta prior as a soft NLL term. **Literature-prior MPC** replaces that Beta with the literature-based one at the same (λ, μ). **Scenario-softmin MPC** consumes the LLM Phase 4 sample pool of 80 samples and, instead of reducing them to a Beta density, evaluates the stage cost on each of a seed-drawn size-K=10 subset:

$$-\frac{1}{\tau}\log\sum_{j=1}^{K}\exp\big(-\tau\, J(u_t;\, u_{\text{LLM}}^{(j)})\big), \tag{4}$$

with τ tied to λ and per-draw Gaussian perturbation σ=0.15. Across seeds $\{0, 1, 2\}$ the row averages cum. inf. 140.07±99.36, cum. cost +15.53±3.30, peak $I{+}Q$ 0.041±0.027, mean $\|u\|_1/5$ 0.632±0.144, prior NLL +118.44±28.34.

B. Future Threads (extended)

Risk-adaptive prior weights. A fixed (λ, μ) pair is a blunt instrument. Early in an outbreak, when the Infectious compartment is still small, the prior's conservative modes are close to the conditional optimum. Once I_t crosses a threshold, they are not. Replacing scalar weights with a state-dependent schedule $\lambda(I_t), \mu(I_t)$ that decays as risk rises would keep the prior where it helps (structure at low

risk) and release it where it hurts (magnitudes at high risk). The experimental cost is a second grid sweep, but the conceptual payoff is a principled reconciliation of the two columns of Table 2.

Multi-agent co-scientist audits. The decomposition in Table 2 is written for a single LLM contributor. In a multi-agent co-scientist setting, different agents may contribute to different pipeline phases (one drafts the DAG, another fits the SCM components, a third proposes controllers). An auditable version of our four-row comparison, with one ablation row per agent's contribution, would extend the human-in-the-loop account from "where did this LLM help" to "which agent, at which phase, was load-bearing." We view this as the most direct path from our single-paper audit to a reusable template for multi-agent co-scientist systems.

C. SIQR Dynamics and MPC Details

Continuous-time SIQR ODEs. The benchmark dynamics are a standard four-compartment SIQR model (Hethcote et al., 2002; Odagaki, 2020) with a five-dimensional NPI vector $u_t \in [0,1]^5$. The vector u_t is mapped by the Phase 5 surrogate to a scalar effective transmission rate $\beta_t(u_t)$, which enters the SIQR dynamics:

$$\dot{S} = -\,\beta_t(u_t)\, S\, I/N, \tag{5}$$

$$\dot{I} = \;\; \beta_t(u_t)\, S\, I/N \;-\; (\gamma + q)\, I, \tag{6}$$

$$\dot{Q} = q\, I \;-\; \lambda_Q\, Q, \tag{7}$$

$$\dot{R} = \gamma\, I \;+\; \lambda_Q\, Q. \tag{8}$$

with $S{+}I{+}Q{+}R = N$ and $N{=}1$ (population-normalized). The transmission rate β_t is produced by the Phase 4 fitted structural equation (Section 2) and is *not* a free parameter of the MPC. We discretize with explicit Euler at step $\Delta t{=}1$ day over an evaluation horizon of $H{=}60$ steps, initialized at $(S_0, I_0, Q_0, R_0) = (0.995,\, 0.005,\, 0,\, 0)$.

MPC problem. At each outer step t we re-solve a receding-horizon optimal control problem over a lookahead of $L{=}10$ steps:

$$\min_{u_{t:t+L-1}\in[0,1]^{5L}} \quad \sum_{k=0}^{L-1}\Big[(I_{t+k} + Q_{t+k}) + \alpha\, \|u_{t+k}\|_2^2 + \lambda\, \ell_\pi(u_{t+k}) + \mu\, \|\Delta u_{t+k}\|_2^2\Big]. \tag{9}$$

subject to the Euler-discretized SIQR dynamics above. Only the first control $u_t^\star$ is applied; the horizon then advances by one step. $\ell_\pi(\cdot,\cdot)$ is the Phase 4 prior negative log-density (see Section 2).

Inner solver. The inner problem is solved with `scipy.optimize.minimize` using `L-BFGS-B` with box constraints $u_k \in [0,1]$, warm-started at $u_k^{(0)}{=}\mathrm{clip}(u_{t-1},\, 0.05,\, 0.95)$, `maxiter`$=50$, and `ftol`$=10^{-5}$. The β-forecast over the lookahead is produced by evaluating the Phase 4 structural equation at the predicted states; the forecasting RNG is forked deterministically from the outer loop so that the MPC inner problem is reproducible.

Hyperparameters. Table 3 lists all dynamics and MPC hyperparameters used to produce Table 2 and the (λ, μ) grid in Appendix F. The values of α and of (λ, μ) are the only method-distinguishing quantities; all other values are shared across all seven conditions in Table 2.

Reproducibility. All closed-loop runs use fixed seeds and deterministic SciPy solvers, and the β-forecast RNG is forked from the outer seed. A single seed reproduces every deterministic row in Table 2 and every cell in Table 7. The Scenario-softmin row of Table 2 additionally depends on the random $K{=}10$ subset drawn from the 80-sample LLM pool, and we report it as mean$\pm$std over three seeds.

C.1. Additional results

This section provides additional qualitative results for the closed-loop trajectories of the four controllers in Table 2. The main table reports aggregate performance metrics. In contrast, the trajectories below show how each method behaves over time. They show both the epidemic state and the control effort assigned to the five intervention levers. This helps identify whether each controller suppresses active cases by using strong controls across many levers or by using a more selective control strategy.

Figure 7 shows the active-case trajectory $I(t) + Q(t)$ and the control activations $u^{(k)}(t)$ for vaccination, mobility reduction, quarantine, masking, and ventilation. The Uniform-prior MPC allows active cases to increase substantially before they decline. This suggests a delayed and moderate control response. The LLM-prior MPC gives a lower peak than the Uniform-prior MPC, but it still allows a sizable epidemic wave. In contrast, the Literature-prior MPC and Scenario-softmin MPC keep active cases much lower over the full horizon by using stronger and more persistent controls. The Literature-prior MPC relies heavily on vaccination and quarantine. The Scenario-softmin MPC places more emphasis on masking, vaccination, ventilation, and mobility reduction, while using relatively little quarantine. These trajectories show that the different priors and scenario constructions affect not only the final cost values, but also the qualitative intervention strategies.

Table 3. SIQR dynamics and MPC hyperparameters used in every closed-loop run. α, λ, μ are the only method-distinguishing values; all other rows are shared.

Symbol	Value	Description
SIQR dynamics		
γ	0.10 / day	recovery rate $I \to R$
q	0.15 / day	quarantine rate $I \to Q$
λ_Q	0.05 / day	release rate $Q \to R$
N	1.0	normalized population
(S_0, I_0, Q_0, R_0)	(0.995, 0.005, 0, 0)	initial condition
Δt	1 day	Euler step size
H	60 steps	evaluation horizon
MPC inner loop		
L	10 steps	lookahead (receding horizon)
Solver	L-BFGS-B	box-constrained quasi-Newton
`maxiter`	50	inner iterations per outer step
`ftol`	10^{-5}	inner convergence tolerance
$u^{(0)}$	$\mathrm{clip}(u_{t-1}, 0.05, 0.95)$	warm-start
Cost weights (method-distinguishing)		
FIXED-UNIFORM	$u_t \equiv (0.40, \ldots, 0.40)$, α=0, λ=0, μ=0	open loop, no optimizer
FIXED-PRIORMODE-LLM	$u_t \equiv (0.733, 0.286, 0.292, 0.560, 0.444)$	open loop at LLM prior modes
FIXED-PRIORMODE-LITERATURE	$u_t \equiv (0.865, 0.555, 0.875, 0.500, 0.555)$	open loop at literature prior modes
LLM-PRIOR MPC	α=0.05, λ=0.1, μ=1, $p_k = \mathrm{Beta}_k^{\mathrm{LLM}}$	MPC, point LLM prior
LITERATURE-PRIOR MPC	α=0.05, λ=0.1, μ=1, $p_k = \mathrm{Beta}_k^{\mathrm{Lit}}$	MPC, point literature prior
SCENARIO-SOFTMIN MPC	α=0.05, λ=0.1, μ=1, K=10 subsampled from 80, σ=0.15	MPC, scenario ensemble

D. Surrogate Details

Table 4. Linear surrogate bootstrap 95% CIs (top-5 features, ranked by $|\hat{w}|$). Weights are fit by OLS on the n_{train}=40 Phase 5 training rollouts, and CIs are obtained from B=500 nonparametric bootstrap resamples of the training split. All five CIs exclude zero, consistent with the sign and ranking reported in the main text (cf. Figure 6).

Feature	$\hat{w}$	95% CI	w/o 0?
Masking compliance	-0.126	$[-0.145, -0.103]$	✓
Vaccination coverage	-0.074	$[-0.090, -0.058]$	✓
Ventilation rate	-0.043	$[-0.059, -0.022]$	✓
Quarantine duration	-0.040	$[-0.064, -0.009]$	✓
Mobility reduction	-0.030	$[-0.049, -0.009]$	✓

We keep the surrogate strictly linear. Higher-capacity alternatives (polynomial features, small MLP) would absorb residual nonlinearity but would also obscure the per-lever weight read in Table 4, which is the object the audit reaches into. The reported R^2=0.85 on held-out rollouts is sufficient for the ranking to be stable, and the linear form is what makes the bootstrap CIs interpretable as confidence in individual lever contributions rather than in a single nonlinear summary.

E. DAG Stability

We rerun Phase 1 (candidate DAG drafting) under three independent elicitation rounds, each drawing K=10 candidates at `temperature`=0.8 with distinct seeds and the same 20 SIQR-relevant factor vocabulary. All other pipeline phases are held fixed. Each round is passed through the same support-$\geq$ 0.5 merge rule described in Phase 2 of Section 2.

Table 5 reports pairwise Jaccard similarity on three representations of the merged per-round DAG: the full edge set, the direct parents of the target node β, and the LLM's self-nominated top-5 decision-relevant factors. Edges into β are substantially more stable than either the full edge set or the top-5 self-nominations, which is consistent with our finding that the LLM's *ranking* of levers is transferable but its full causal structure is not.

Table 5. Pairwise DAG Jaccard similarity across three independent elicitation rounds (K=10 candidates each, merge rule support$\geq$ 0.5). Merged edge counts: $|E_1|$=17, $|E_2|$=19, $|E_3|$=20.

	J_{12}	J_{13}	J_{23}	mean
Full edge set	0.800	0.762	0.857	0.806
Direct β-parents	0.938	0.882	0.941	0.920
LLM top-5 factors	0.667	0.429	0.667	0.587

Per-round edge support. Table 6 lists every edge with support $\geq$ 0.5 in at least one round. An entry is the fraction of the round's K=10 candidate DAGs that contained that edge; a blank cell means the edge fell below 0.5 in that round. All but four rows are direct β-parents, as one would expect from a target-ranking prompt.

Table 6. Edge support fraction across $K{=}10$ candidates per round. Rows sorted by average support. "β" denotes the target node *Transmission rate beta*. Blank cells indicate support <0.5 in that round.

Source	Target	$r{=}1$	$r{=}2$	$r{=}3$
Contact tracing efficiency	β	1.0	1.0	1.0
Household density	β	1.0	1.0	1.0
Masking compliance	β	1.0	1.0	1.0
Quarantine effectiveness	β	1.0	1.0	1.0
Superspreader events	β	1.0	1.0	1.0
Variant strength	β	1.0	1.0	1.0
Ventilation rate	β	1.0	1.0	1.0
Mobility reduction	β	1.0	0.9	1.0
Vaccination coverage	β	1.0	1.0	0.9
Vaccine efficacy	β	0.9	0.9	1.0
Seasonality	β	0.8	1.0	0.8
Social distancing measures	β	0.9	0.5	0.8
Quarantine duration	β	0.9	0.6	0.6
Age structure	β	0.8	0.7	0.6
Population mixing	β	0.7	0.7	0.6
Public health interventions	β		0.5	0.5
Quarantine compliance	β			0.5
Quarantine compliance	Quar. effectiveness	0.7	0.5	
Quarantine capacity	Quar. effectiveness		0.5	0.5
Mobility patterns	Mobility reduction		0.5	0.5
Economic factors	Mobility patterns	0.5		0.5

What this table tells us. (i) Nine direct β-parents are unanimously supported in at least one round and carry support ≥ 0.9 in all three; these are the edges that survive regardless of the elicitation seed. (ii) The two surrogate-heaviest levers from Equation (2), Masking and Vaccination, sit in this unanimous-or-near-unanimous band, so the lever-ranking claim in Table 2 is not an artifact of a single elicitation run. (iii) Non-β mediated edges appear at exactly the 0.5 threshold and drift across rounds, which is why we required a written reviewer justification for any sub-threshold edge in Phase 2: at this elicitation scale the LLM is reliable on *which factors touch* β but not on *how they interact upstream*. (iv) The LLM-nominated top-5 has the lowest Jaccard because it collapses ≥ 9 near-unanimous parents into a length-5 list; the merged β-parent set (which we actually use) is a strictly stronger statistic.

F. Prior- and Smoothness-Weight Grid

Table 7 reports LLM-prior MPC on a $4{\times}4$ grid over (λ, μ). The prior parameters used throughout are the Phase 4 Beta summaries $(\alpha_i, \beta_i) = (3.2, 1.8), (1.8, 3.0), (1.7, 2.7), (2.4, 2.1), (2.2, 2.5)$ for Vaccination, Mobility reduction, Quarantine duration, Masking, and Ventilation, respectively. The $(\lambda, \mu){=}(0, 0)$ cell is omitted because it removes both the prior pull and the smoothness regularizer and leaves an underdetermined controller.

Within the active cells, cumulative infections move monotonically along each axis. Increasing μ at fixed λ raises infections (the smoothness penalty makes the controller less responsive). Increasing λ at fixed μ also raises infections (the prior pull biases the controller toward LLM-elicited modes that are less aggressive than the surrogate's free optimum). The Table 2 operating point at $(\lambda, \mu){=}(0.1, 1)$ is not the cell-wise minimum. We report at this cell because it is the smallest grid cell at which both the prior pull ($\lambda{>}0$) and the smoothness penalty ($\mu{>}0$) are nontrivially active, which lets us attribute closed-loop differences to elicited content rather than to a regularizer corner. The qualitative ranking among Uniform-prior MPC, LLM-prior MPC, and Literature-prior MPC is preserved across the broader region $\lambda \in [0.01, 0.1]$ and $\mu \in [0, 1]$, so the choice is a robustness-aware reporting decision rather than a narrow tuning artifact.

Table 7. Cumulative infections per 1,000 for LLM-prior MPC over the (λ, μ) grid (single seed, deterministic). The $(\lambda, \mu){=}(0, 0)$ cell is omitted as N/A. The Table 2 operating point $(\lambda, \mu){=}(0.1, 1)$ is shown in bold.

$\mu \backslash \lambda$	0	0.01	0.1	1
0	–	193.04	309.17	335.67
0.1	107.76	199.13	309.76	335.72
1	175.06	226.97	**313.66**	336.12
5	253.69	279.13	324.98	337.73

G. Prompts

We use a single instruction-tuned LLM at two elicitation phases. Phase 1 (DAG drafting) is run with $K{=}10$

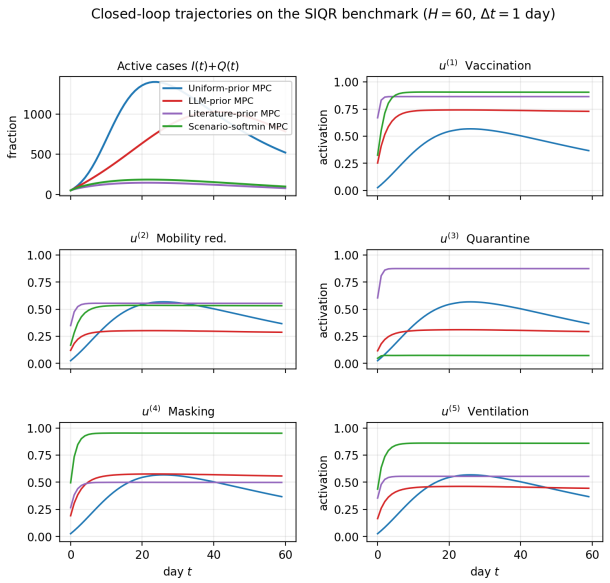

Figure 7. Closed-loop trajectories for the four controllers in Table 2. The top-left panel shows active cases $I(t)+Q(t)$. The remaining panels show the control activations $u^{(k)}(t)$ for vaccination, mobility reduction, quarantine, masking, and ventilation. The Uniform-prior, LLM-prior, and Literature-prior MPC curves are single deterministic trajectories. The Scenario-softmin MPC trajectory is shown for seed 0, using a random subset of $K=10$ samples from the 80-sample LLM pool.

candidate queries per elicitation round, each conditioned on a different diversity-prior instruction (listed below) at `temperature`=0.8. Three independent rounds are used for the Jaccard stability analysis in Section 3. Phase 3 (component sampling) is run at `temperature`=0.7 with 80 draws, which seed both the LLM-prior Beta fit and the 80-sample LLM pool from which scenario-softmin MPC samples. We do not fix the LLM's internal sampling seed and treat run-to-run variation as part of the upstream stochasticity reported in Section 5. Both prompts are implemented as Python f-string templates in `src/causal_workflow.py`, and the placeholders (`{factor_block}`, `{candidate_id}`, etc.) are filled in at runtime from the curated factor table and the selected candidate.

Diversity priors used at Phase 1. Each Phase 1 query is conditioned on one of the following ten instructions, drawn in order so the $K=10$ candidates span structurally different hypothesis classes.

```
Policy-first DAG: make public
    interventions upstream of behavioral
    changes and beta.
Behavior-first DAG: emphasize mobility,
    masking, gatherings, and social
    distancing as mediators.
Environment-first DAG: emphasize
    seasonality, ventilation, temperature,
    and air filtration effects.
Quarantine/testing-first DAG: emphasize
    testing, tracing, and quarantine
    pathways into beta.
Vaccination-first DAG: emphasize
    vaccination as the main upstream driver
     with downstream policy/behavior links.
Variant-first DAG: emphasize variant
    strength as an upstream driver
    affecting multiple downstream factors.
Household/community DAG: emphasize
    household density and superspreader
    pathways.
Sparse DAG: keep edges minimal and only
    include the strongest causal
    assumptions.
Mediated DAG: prefer multi-step mediation
    over many direct edges into beta.
Intervention-heavy DAG: emphasize
    controllable policy levers that
    ultimately change beta.
```

Phase 1: DAG drafting prompt. Reproduced verbatim from `src/causal_workflow.py`. The target node `"Transmission rate beta"` is appended as a sink, and the factor list is the curated 20-factor table from the upstream factor-extraction phase.

```
You are helping with a toy SIQR causal
    modeling study for methodological
    research.

Task:
- Propose exactly one literature-informed
    draft DAG candidate over the given 20
    factors.
- Treat this as a toy hypothesis graph,
    not as validated real-world policy
    advice.
- Add one sink node named "Transmission
    rate beta".
- Use the exact factor names provided
    below. Do not invent new factor names.
- Keep the graph acyclic.
- Include 12 to 22 directed edges.
- Include some mediated relationships
    between factors, not only direct edges
    into beta.
- Pick exactly 5 factors as the current
    most decision-relevant factors for
    later SD-SCM and control analysis.

Diversity instruction for this candidate:
- {diversity_prior}

Factors:
{factor_block}

Return JSON only in this schema:
{
  "candidate_id": "{candidate_id}",
  "title": "short descriptive title",
```

```
  "target_node": "Transmission rate beta",
  "nodes": ["Transmission rate beta", <
      factor names>],
  "edges": [
    {"source": "factor A", "target": "
        factor B or Transmission rate beta",
         "reason": "short reason"}
  ],
  "beta_parents": ["factor names that
      directly point into Transmission rate
       beta"],
  "top5_factors": ["exact factor name", "
      exact factor name", "exact factor
      name", "exact factor name", "exact
      factor name"],
  "rationale": "2-4 sentence summary"
}

Important rules:
- The only allowed node names are: {
    factor_list}, Transmission rate beta
- Use exact spelling from the factor list.
- Do not include markdown fences.
- Do not include any explanation outside
    the JSON.
```

Phase 1 system prompt for upstream factor extraction. The 20-factor table that feeds the Phase 1 DAG drafter is itself elicited from the literature with a separate single-shot prompt, reproduced verbatim from `src/prompt.py`.

```
You are an expert epidemiologist
    specializing in SIQR (Susceptible-
    Infected-Quarantined-Recovered) models.

From the provided recent papers (arXiv
    abstracts and content), extract exactly
     20 distinct factors that directly
    influence the transmission rate beta(t)
     in an SIQR model.

For each factor, output in this exact
    format:
1. Factor name | Hypothesized causal
    pathway | Reference (paper title/year)

Rules:
- Factor name must be short and clear (e.g.
     "Mobility reduction", "Ventilation
    rate")
- Hypothesized pathway must explain HOW it
     affects beta (e.g. "reduces contact
    rate between S and I", "increases
    effective transmission probability")
- Prioritize quarantine-related, mobility,
     ventilation, masking compliance,
    variant strength, superspreader events,
     seasonality, household density, policy
     interventions
- Do not repeat similar factors
- Only use information from the provided
    papers
- Output exactly 20 lines, nothing else.
```

Phase 3: component sampling (SD-SCM spec) prompt. Reproduced verbatim from `src/causal_workflow.py`. Phase 3 consumes the consensus DAG selected in Phase 2 and emits, for each node, a structural-equation specification (parents, role, state space, and a natural-language generator prompt) that downstream phase render into Beta priors over the five intervention levers.

```
You are helping with a toy SIQR causal
    modeling study.

Task:
- Convert the selected draft DAG into an
    SD-SCM style specification.
- Treat the DAG as fixed.
- Do not discover new edges.
- For each node, describe how an LLM could
     generate its value from its parents.
- Focus on methodological use in a toy
    model only.

Selected DAG:
- candidate_id: {candidate_id}
- title: {title}
- target node: {target_node}
- top5 factors: {top5_factors}

Node descriptions:
{factor_descriptions}

Edges:
{edges_text}

Return JSON only in this schema:
{
  "candidate_id": "{candidate_id}",
  "graph_title": "{title}",
  "target_node": "{target_node}",
  "top5_factors": [...],
  "node_specs": [
    {
      "name": "node name",
      "parents": ["parent 1", "parent 2"],
      "role": "root|mediator|target|
          intervention",
      "state_type": "binary|ordinal|
          continuous|categorical",
      "state_space": "short textual
          description",
      "generator_prompt": "natural-language
          structural equation prompt for
          this node",
      "intervention_values": ["example value
           1", "example value 2"],
      "notes": "short note"
    }
  ],
  "simulation_protocol": {
    "topological_order": ["ordered nodes"],
    "root_node_strategy": "how to
        initialize roots",
    "beta_generation_note": "how the target
         node is generated from its parents"
```

```
      ,
    "example_interventions": [
      {"node": "factor", "value": "value", "
          expected_effect_on_beta": "short
          note"}
    ]
  }
}

Rules:
- Include every node from the DAG exactly
    once in node_specs.
- Use the DAG parents exactly as given.
- Keep the output concise but complete
    enough to drive an SD-SCM
    implementation.
- Do not include markdown fences.
```