GSC Advanced Research and Reviews
eISSN: 2582-4597 CODEN (USA): GARRC2
Cross Ref DOI: 10.30574/gscarr
Journal homepage: https://gscarr.gsconlinepress.com/

(RESEARCH ARTICLE)

AN INTEGRATED CAUSAL-PREDICTIVE OPTIMIZATION FRAMEWORK FOR HEALTHCARE RESOURCE EFFICIENCY, FINANCIAL RESILIENCE, AND SUSTAINABLE SERVICE DELIVERY

Oluwasanya Ogunsakin [1,*], Temitope Iyamu Fadairo [2], Misheck Malangano [3] and Kevin Mukasa [4]

[1] Bournemouth University, UK.
[2] DePaul University, Chicago.
[3] University of West Georgia.
[4] Maharishi International University.
* Corresponding Author

GSC Advanced Research and Reviews, 2026, 29(02), 355–367

Publication history: Received on 10 October 2026; revised on 21 November 2026; accepted on 23 November 2026

Article DOI: https://doi.org/10.30574/gscarr.2026.29.2.0339

Copyright © 2026 Author(s) retain the copyright of this article. This article is published under the terms of the Creative Commons Attribution License 4.0

ABSTRACT

Hospitals must preserve surge capacity while meeting operating budgets, yet policy effects, workload forecasts, and allocation decisions require different kinds of evidence. This study develops an integrated framework and evaluates it using the Centers for Medicare & Medicaid Services Hospital Provider Cost Report, U.S. Department of Health and Human Services hospital capacity records, and Agency for Healthcare Research and Quality national utilization estimates. The analytic panel contained 35,062 fiscal-year reports from 4,760 short-term and critical access hospitals for 2017–2024. A state-level difference-in-differences analysis compared the 2021 Medicaid expansions in Oklahoma and Missouri with eight comparison states. Relative to the comparison states, the treated states recorded a 3.99 percentage-point increase in the Medicaid inpatient-day share and a 3.20 percentage-point decline in operating margin. Reference probabilities from 45 state-pair reassignments were .089 and .289, respectively, leaving the margin effect uncertain. In a 2024 holdout of 4,155 hospitals, ridge regression achieved a mean absolute error of 2,493 annualized patient-days, compared with 2,638 for persistence. The paired state-bootstrap interval for the difference was −245 to −43 patient-days. Under an illustrative 3% reserve budget and an 80% planning threshold, scenario-based allocation reduced the modeled capacity gap from 3.26 million to 1.79 million bed-days, a 45.2% reduction under equal reserve budgets. The estimated capacity gains are retrospective model outputs. Realized savings, care access and clinical outcomes require prospective evaluation.

Keywords: Causal Inference, Financial Resilience, Healthcare Resource Allocation, Hospital Capacity, Predictive Modeling, Stochastic Optimization

1 INTRODUCTION

Hospitals make staffing and capacity commitments before they know how many patients will arrive, how long those patients will stay, or which payers will reimburse the resulting care. Rising inpatient workload can strain capacity even when annual admissions are stable if average stays become longer. Conversely, an annual financial surplus can coexist with losses on patient operations when a hospital receives other income. Evidence from U.S. hospitals during the COVID-

19 pandemic illustrates why patient-service margins and broader financial results must be examined separately (Wang et al., 2022). Research on hospital strain also links short-term surges with poorer outcomes in specific pandemic settings, making the timing of capacity constraints relevant to service continuity (Bravata et al., 2021; French et al., 2021; Kadri et al., 2021).

Three questions arise when allocating resources. How much of a change in financial performance can plausibly be attributed to policy? How much inpatient workload is a hospital likely to face? Where should a health system place a fixed amount of reserve capacity? A historical policy contrast can inform financing assumptions but offers little direct guidance on the location of future demand. A forecast can estimate that demand, while its operational value depends on whether a feasible allocation improves the outcome that managers care about. An allocation model therefore needs a measure of workload, an explicit decision objective, and constraints that reflect the availability of staff and funding.

This study addresses these questions using publicly accessible U.S. federal sources. Hospital cost reports supply the financial and utilization measures for the hospital panel. HHS state-day observations and AHRQ national estimates place those results in the context of system strain and changes in service use. The framework treats sustainable service delivery as the ability to maintain access and operating capacity under a finite resource envelope. Financial resilience is assessed through patient-service margins, net income, and the current ratio. These are annual accounting indicators. An assessment of long-term solvency would additionally require debt-service schedules, cash flows, and access to working capital.

2 RELATED WORK AND CONCEPTUAL FRAMEWORK

2.1 Coverage policy and hospital finances

Medicaid expansion can change which payer covers care previously reimbursed at lower rates or delivered without payment. Blavin and Ramos (2021) found that earlier expansions were associated with lower uncompensated care costs and stronger financial results, with differences across hospital types. Moghtaderi et al. (2020) examined how expansion changes the composition of hospital revenue, making payer-specific measures important when interpreting overall margins. Neither mechanism guarantees a particular margin change because the net result also depends on payment rates, service intensity, and expenses. The analysis therefore pairs an operating-margin outcome with the share of inpatient days reported under Title XIX as a secondary measure of payer composition.

A policy comparison depends on the counterfactual trend. Hernán et al. (2022) emphasize defining eligibility, exposure and follow-up before interpreting observational comparisons. Callaway and Sant'Anna (2021) and Sun and Abraham (2021) show why timing and comparison-group construction matter in difference-in-differences designs. The present analysis isolates a single adoption cohort and avoids comparisons with states already exposed to the policy. Its conclusions remain conditional on the comparison states approximating the treated states' post-period trajectory in the absence of expansion. Pre-period differences and the number of treated states are especially relevant to this judgment (MacKinnon & Webb, 2020; Rambachan & Roth, 2023; Roth et al., 2023).

2.2 Prediction of inpatient workload

A useful planning forecast should beat a transparent benchmark on data held out in time. Forecasting research emphasizes that the choice of error measure and baseline can change model rankings (Makridakis et al., 2022). Grinsztajn et al. (2022) provide a reason to test tree-based methods on tabular records, while the accuracy of any such model remains an empirical question. This study compares persistence, ridge regression, and histogram gradient boosting. It uses an interval derived from validation residuals to quantify uncertainty in workload and to supply alternative demand scenarios (Bracher et al., 2021). Reporting choices on missingness, evaluation and intended use draw on TRIPOD+AI and PROBAST+AI as transferable principles for transparent prediction research (Collins et al., 2024; Moons et al., 2025).

Akinyelure (2023) identifies incomplete representation and limited community engagement as risks when predictive analytics informs services for underserved groups. In a separate review, Akinyelure (2025) argues that community-generated information can supplement administrative and clinical records in health planning. Here, the forecasts estimate recorded inpatient use. A prospective evaluation could compare that signal with referrals, delayed admissions, and local accounts of barriers to care when considering where to deploy reserve capacity.

2.3 From forecast to allocation

Predictive-to-prescriptive research links forecasts to decisions by evaluating the loss incurred under an actual decision rule (Bertsimas & Kallus, 2020; Elmachtoub & Grigas, 2022). Healthcare applications have illustrated how forecast scenarios and resource-sharing constraints can guide responses to surges and shortages (Bertsimas et al., 2021;

Mehrotra et al., 2020). The practical challenge is that resources can be allocated only where they can be staffed and maintained. The study compares three allocations under the same state reserve envelope and common site-level limits. The objective measures the annual workload remaining above a planning threshold after reserve is assigned.

The framework connects the analyses in sequence. The policy comparison determines whether an assumed financial gain can credibly support a larger planning envelope. Historical utilization generates point forecasts and scenarios. Those scenarios then drive a constrained decision that is evaluated against held-out utilization. The HHS and AHRQ analyses characterize service pressure and help interpret the planning scenarios without equating state-day occupancy, national stays, and hospital fiscal-year records. This design makes the assumptions about fiscal headroom, forecast accuracy, and resource deployment visible at each stage. Figure 1 shows the evidence each stage supplies to the next decision.

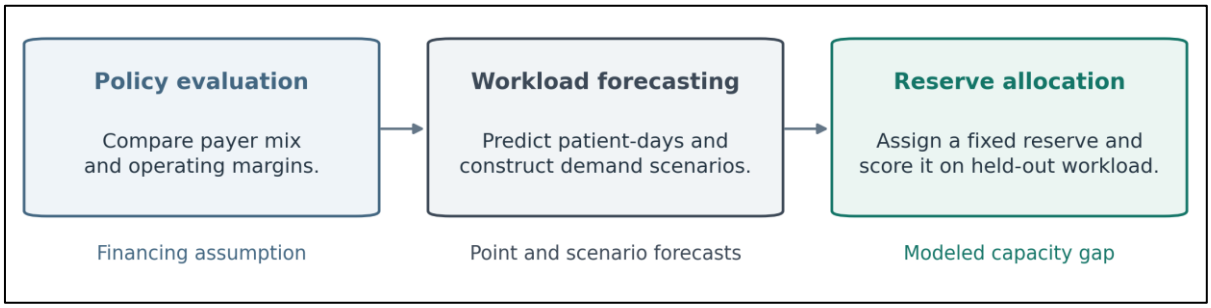

Figure 1: Causal, predictive, and optimization stages of the framework

Key: The policy comparison informs the financing assumption. Workload scenarios guide reserve allocation under that budget.

2.4 Research Objectives

The objectives of this study are to:

- Assemble and document a reproducible federal data pipeline for hospital utilization, expenditure, financial performance, and service demand.
- Estimate changes in payer composition and operating margin following the 2021 Medicaid expansions in Oklahoma and Missouri, and assess the sensitivity of the comparison.
- Compare competing forecasts of annualized hospital inpatient workload using a chronologically held-out fiscal year.
- Compare proportional, point-forecast, and scenario-based reserve allocations under equal resource budgets and quantify the modeled service-capacity implications.

3 MATERIALS AND METHODS

3.1 Federal data and analytic populations

The study downloaded the CMS Hospital Provider Cost Report public-use CSVs for the releases labeled 2016–2024, the HHS state time-series extract, and two HCUP Fast Stats national Excel workbooks (Agency for Healthcare Research and Quality [AHRQ], 2026a, 2026b; Centers for Medicare & Medicaid Services [CMS], 2026; U.S. Department of Health and Human Services [HHS], 2024). The CMS public-use files provide selected variables from annual cost reports in the Healthcare Provider Cost Reporting Information System. The public sources differ in population coverage, unit of observation and reporting calendar.

The CMS cohort was limited to the 50 states and the District of Columbia, with provider type 1 and a facility designation of short-term hospital (STH) or critical access hospital (CAH). Records were pooled before assigning the analytic year from the actual fiscal-year end date. This recovered reports that closed in the year after the file label. The cohort required a valid CMS Certification Number (CCN) and a reporting period of 350–380 days. For the one duplicate CCN-year, the record with the later fiscal-year end was retained. Leading zeros in CCNs were preserved. The result was 35,062 hospital-years and 4,760 distinct CCNs. The policy and prediction modules applied additional eligibility rules because their outcomes and longitudinal requirements differed.

HHS context used complete calendar years 2021–2023. State-days required a positive bed denominator, occupancy from 0% to 100%, and equal, positive reporting counts for bed supply and use. Equal counts are a comparability check and can occur when different facilities contribute to the two totals. HHS metrics summarize the reporting hospitals.

AHRQ estimates use the published overall categories of the National Inpatient Sample and Nationwide Emergency Department Sample. AHRQ's emergency department (ED) counts include visits resulting in admission at the same hospital, so ED visits and inpatient stays overlap. No individual hospital identifier was inferred from either contextual source.

Table 1: Federal datasets and their analytic roles

Source	Coverage used	Measures and role
CMS Hospital Provider Cost Report public-use CSVs	File labels 2016–2024. Hospital fiscal years ending 2017–2024.	CCN, facility type, reported patient-days, bed-days, Title XIX days, net patient revenue, operating expense, net income and current assets and liabilities. Primary panel.
HHS Reported Patient Impact and Hospital Capacity	State-day observations. Complete calendar years 2021–2023.	Inpatient beds, beds used, and reporting counts. Daily pressure and context for the threshold range.
AHRQ HCUP Fast Stats national Excel tables	National inpatient stays and ED visits. Calendar years 2019–2023.	Weighted stay and visit estimates, cost per inpatient stay and length of stay. National context.
CMS Medicaid state plan amendments	Oklahoma and Missouri. 2021 effective dates.	Formal eligibility timing. KFF implementation tracking cross-checks comparison state status and rollout.

Key: Exact source URLs and checksums are supplied with the replication materials.

3.2 Measures and data quality

Table 2: Measures and interpretation

Indicator	Calculation or source	Interpretation
Inpatient workload	Reported hospital patient-days × 365 ÷ report duration	Observed annualized utilization used as forecast target.
Bed occupancy	Recorded patient-days ÷ available bed-days	Annual utilization of reported capacity.
Medicaid-day share	100 × Title XIX inpatient days ÷ all reported inpatient days	Payer composition of recorded inpatient activity.
Operating margin	100 × (net patient revenue − operating expense) ÷ net patient revenue	Patient-service margin. Net income is read separately as reported.
Current ratio	Current assets ÷ current liabilities	Accounting liquidity at the report date.
Modeled capacity gap	Positive part of observed annualized workload minus threshold-adjusted base and reserve	Residual workload above a scenario capacity threshold.

CMS utilization uses the adult and pediatric hospital worksheet aggregates for recorded inpatient days and available bed-days. The term "workload" denotes realized service use. Patients who were diverted, left without treatment, or could not obtain care are outside that measure. Reports with zero patient-days or occupancy outside 0–100% were excluded from utilization and forecasting summaries. Occupancy at the panel level was calculated as the sum of patient-days divided by the sum of available bed-days. The financial subset required positive patient revenue and operating expense and an operating margin in [−100%, 100%]. Balance-sheet ratios were defined only for nonnegative current assets and positive current liabilities. The policy analysis additionally required a Medicaid-day share in [0%, 100%] in every analyzed year. Missing values remained missing. CMS dollar totals are nominal, while AHRQ's inflation-adjusted cost series uses its published GDP price-index adjustment with a 2010 base.

3.3 Policy evaluation and statistical inference

Oklahoma and Missouri both have federal state-plan effective dates of July 1, 2021 (CMS, 2020, 2021). Missouri began processing applications in October with retroactive coverage (KFF, 2026). The comparison states were Alabama, Florida, Kansas, Mississippi, South Carolina, Tennessee, Texas, and Wyoming. Georgia was excluded because its limited adult coverage pathway began during follow-up. Wisconsin was excluded because its pre-existing adult coverage to the poverty level changes the comparison. The primary pre-period was FYs ending in 2018–2019. The primary post-period was FYs ending in 2023–2024, with 2021–2022 excluded to avoid partial implementation and periods with inconsistent

exposure. FY2020 was included only in diagnostics and a sensitivity specification because the onset of COVID-19 substantially disrupted hospital utilization and finances.

The balanced cohort required a usable observation in 2017–2020 and 2023–2024 and stable state assignment. For each outcome, hospital values were averaged within state-year. The estimator then subtracted the equal-state-weighted mean change among comparison states from the corresponding mean among treated states. The estimand weights Oklahoma and Missouri equally and represents the eligible, continuously reporting hospitals in those states. A causal interpretation would require parallel untreated trends, limited interference between states, and a stable reporting composition that is not materially determined by the policy. The analysis tested the influence of baseline periods and individual comparison states. It also inspected the annual treated–comparison gap relative to the 2018–2019 reference.

With two treated states, inference based on the number of hospital records would overstate precision. The analysis enumerated all 45 ways to label two of the ten states as treated, recomputed the same contrast, and calculated the fraction with absolute magnitude at least as large as observed. These are two-sided reference probabilities. Their interpretation as exact randomization probabilities would require exchangeable assignments, which the policy adoption process does not establish (MacKinnon & Webb, 2020). Alternative-period and leave-one-state results describe sensitivity. A matched or synthetic difference-in-differences design would warrant a larger donor pool, credible weights and additional pre-period observations (Arkhangelsky et al., 2021).

3.4 Forecast design and temporal validation

The forecast target was inpatient days multiplied by 365 and divided by report duration. Capacity inputs were normalized in the same way. Pairs used the same CCN and state two fiscal years apart, with known urban or rural designation and valid workload measures. The benchmark copied workload from two years earlier. The fitted models used log lagged patient-days, log lagged available bed-days, lagged occupancy, rural status, and CAH status. Ridge regression standardized the predictors, used a penalty of 10, and modeled log patient-days. The boosted model used a maximum of 160 iterations and 15 leaves, a minimum of 40 observations per leaf, and an L2 penalty of 1. Both predictions were exponentiated without a retransformation correction. Forecast bias was evaluated on the original scale.

Training targeted 2019–2022 and included 16,820 paired records. FY2023 supplied 4,238 validation records and determined the model with the lowest mean absolute error (MAE). The model was frozen before comparing predictions with 4,155 FY2024 records. The two-year fiscal lag reduces reliance on recent filings. Because the snapshots include subsequent revisions and use an ex post matched test roster, this is an outcome-year holdout under retrospective data availability. A live deployment would require archived as-of snapshots, a prospectively frozen roster, and a procedure for new or closed hospitals.

The test metrics were MAE and root mean squared error (RMSE) in patient-days, weighted absolute percentage error (WAPE), and signed percentage bias. WAPE divides the sum of absolute errors by the sum of observed workload. Bias divides the sum of forecast minus observed workload by the same denominator. The interval for the paired MAE difference was based on 5,000 state-cluster bootstrap draws, resampling states and their hospitals together. It measures cross-state sampling variation conditional on the fixed test year and fitted models. Random draws used seed 31. Validation residuals, defined as observed minus forecast days divided by lagged bed-days, were summarized at the 10th, 25th, 50th, 75th, and 90th percentiles. The two outer quantiles defined a nominal 80% interval whose empirical coverage was evaluated on the test year.

3.5 Constrained reserve allocation

The decision unit was a hospital within a state. For each state s, let I_s denote its hospitals, C_i the lagged annualised available bed-days at hospital i, x_i its incremental reserve allocation and θ the planning threshold. The threshold-adjusted base capacity was θC_i. For each residual quantile q_k, the model calculated the workload scenario as follows.

$$D_{ik} = max\{0, \hat{D}_i + q_k C_i\}, \quad i \in I_s, \quad k = 1, \ldots, K.$$

Here, $\hat{D}_i$ is the selected workload forecast and K is the number of scenarios. Each quantile applied a common percentile shock to hospitals in the same state, scaled by their lagged bed-days. The scenarios had equal objective weights and served as stress cases. The weights did not represent estimated joint probabilities. The point-forecast policy used a single scenario with D_{i1} equal to the selected prediction.

3.5.1 Optimisation model

The model used u_{ik} for the remaining shortfall at hospital i in scenario k and z_i for the deviation from its proportional reserve allocation. The reserve share is b, the site-specific lower bound is ℓ_i, and the deviation penalty is $\lambda = 10^{-5}$. For each state, the linear programme was as follows.

$$min_{x,u,z} \ \frac{1}{K} \sum_{i \in I_s} \sum_{k=1}^{K} u_{ik} \ + \ \lambda \sum_{i \in I_s} z_i$$

Subject to

$$u_{ik} \geq D_{ik} - \theta C_i - x_i, \quad i \in I_s, \quad k = 1, \ldots, K.$$

$$u_{ik} \geq 0, \quad i \in I_s, \quad k = 1, \ldots, K.$$

$$\sum_{i \in I_s} x_i = b \sum_{i \in I_s} C_i.$$

$$\ell_i \leq x_i \leq 0.15 C_i, \quad i \in I_s.$$

$$z_i \geq x_i - bC_i, \quad i \in I_s.$$

$$z_i \geq bC_i - x_i, \quad i \in I_s.$$

The first objective term penalised total workload above the base capacity and assigned reserve, averaged across scenarios. The final two constraints ensure that z_i is at least the absolute deviation from bC_i. Because the objective penalises z_i, it equals that absolute deviation at the optimum. The small penalty favours allocations closer to the proportional benchmark when shortfall objectives are tied or nearly tied.

The statewide reserve budget was b times the state's total lagged available bed-days. Each hospital could receive up to 15% of its own lagged bed-days. The minimum was 0.5% for rural hospitals and zero for urban hospitals. The proportional benchmark allocated bC_i to each hospital. The proportional, point-forecast and scenario-based policies used the same statewide reserve budget, rural floor and hospital caps.

The analysis evaluated planning thresholds of 75%, 80% and 85%, and reserve shares of 2%, 3% and 4%. The HHS state-day patterns informed the choice of 80% as an illustrative midpoint for planning. The parameter was used for scenario comparison and has no clinical safety interpretation. The linear programmes were solved using SciPy's HiGHS interface, and the solutions were checked for feasibility.

3.5.2 Backtest scoring

Each allocation was scored against observed FY2024 patient-days. Let Y_i denote observed annual workload. The remaining gap for hospital i was calculated as follows.

$$g_i = max\{0, Y_i - \theta C_i - x_i\}.$$

The outcomes were the total gap, the gap at rural hospitals, the number of hospitals with a positive gap and the amount of assigned reserve unused under the annual scoring rule. Observed workload was kept at its original hospital, and the backtest assumed that each unit of assigned reserve was available at its allocated site. Referral changes were outside the scoring model.

Paired state-level scores were resampled to quantify variation in the relative gap reduction. The comparison was conditional on the same observed workload, fixed physical reserve budgets and modelling assumptions. The backtest did not estimate site-specific deployment costs.

3.5.3 Budget assumptions and pilot costing

The policy analysis assessed whether the allocation budget could incorporate expansion-related financing. The historical evidence did not establish a reliable increase in operating margin, so the executed allocation assumed zero policy-related budget uplift. The policies were compared under fixed physical reserve budgets.

A funded pilot would add a separate constraint requiring total facility-specific incremental deployment costs and coordination costs to stay within the approved contingency budget. Eligible costs would include staffing, activation, supplies and associated transfer logistics. Published hospital operating expenses are institution-level accounting totals

and do not identify these marginal costs. The pilot would price the reserve explicitly and compare incremental costs with independently measured savings or capacity benefits to assess the break-even point.

4 RESULTS

4.1 Observed utilization and financial performance

Table 3: CMS utilization and financial indicators by fiscal year

FY end	Eligible reports	Occupancy %	Operating expense $bn	Weighted margin %	Median margin %	Median current ratio
2017	4,432	56.15	861.1	-0.21	-2.11	2.14
2018	4,390	56.76	900.5	0.80	-2.15	2.12
2019	4,345	57.38	947.8	0.29	-1.75	2.12
2020	4,380	54.80	994.2	-5.29	-6.72	1.60
2021	4,418	58.68	1,085.8	-1.30	-2.36	1.81
2022	4,418	60.15	1,171.1	-3.73	-4.77	2.13
2023	4,377	60.54	1,248.3	-2.32	-4.53	2.23
2024	4,302	61.91	1,322.5	-1.40	-2.26	2.18

Key: Operating expense is a nominal total for the financial subset. Occupancy is weighted by available bed-days and the weighted margin is weighted by net patient revenue. In 2024, 4,268 reports met utilization filters and 4,201 met finance filters.

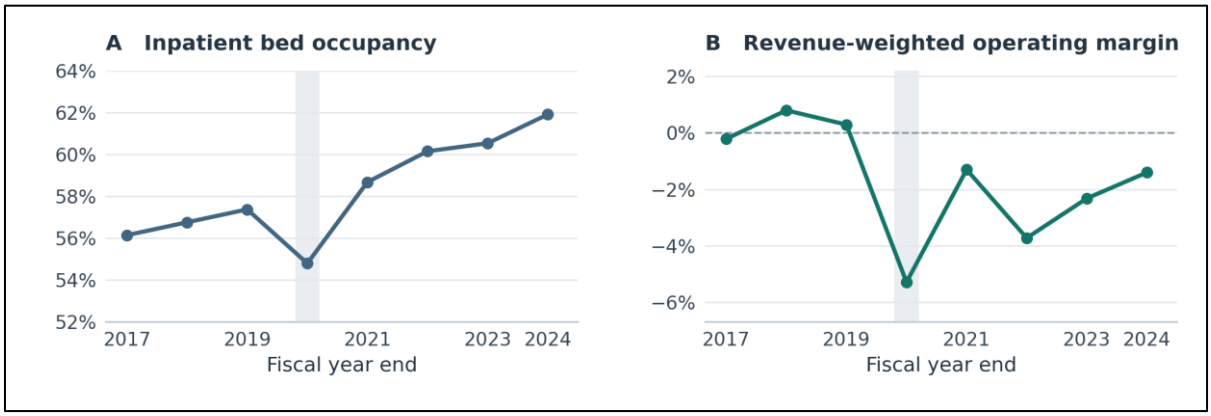

Figure 2: Observed inpatient bed occupancy and operating margin in the CMS hospital panel

Key: FY2017–FY2024. Shading marks FY2020. Occupancy is weighted by available bed-days, and operating margin is weighted by net patient revenue. The two indicators use the eligible subsets defined in Section 4.2.

Aggregate occupancy among eligible reports increased from 57.38% in 2019 to 61.91% in 2024, after falling to 54.80% in 2020. Nominal operating expenses in the eligible finance sample rose from $947.8 billion to $1,322.5 billion between 2019 and 2024. The annual cohorts vary, and the totals reflect changes in prices, activity and the institutions represented. The revenue-weighted operating margin moved from 0.29% in 2019 to −5.29% in 2020 and −1.40% in 2024. The median current ratio declined from 2.12 to 1.60 in 2020 and reached 2.18 in 2024. Median reported net income was positive in 2024 at $4.50 million while the median patient operating margin was −2.26%. These are separate cohort-level summaries of different accounting measures. Figure 2 shows the annual trajectories of occupancy and operating margin.

Table 4: AHRQ national inpatient and emergency department indicators

Calendar year	Inpatient stays millions	ED visits millions	Mean stay days	Real cost per stay 2010 USD
2019	35.42	143.43	4.70	$11,505
2020	32.36	123.28	4.94	$12,611
2021	33.33	126.97	5.16	$12,908
2022	32.89	136.97	5.18	$12,694
2023	33.72	142.19	5.07	$12,971

Key: The values are published national estimates. Real costs use AHRQ's adjustment. ED visits include visits ending in same-hospital admission.

The AHRQ national estimates show that 2023 inpatient stays remained 4.8% below 2019, while average length of stay increased from 4.70 to 5.07 days. The published inflation-adjusted cost per stay increased approximately 12.7%. ED visits recovered to 142.2 million in 2023 following the 2020 decline to 123.3 million. In the filtered HHS records, pooled inpatient bed use was 74.25% in 2021, 76.01% in 2022, and 75.18% in 2023. The 90th percentiles of the state-day occupancy ratios were 81.26%, 83.01%, and 82.42%, respectively. The HHS ratios and CMS ratios use different reporting populations and bed definitions. Together, these sources document increased intensity per inpatient stay and variation in capacity pressure over time.

4.2 Medicaid expansion and hospital financial resilience

The policy panel comprised 706 hospitals, including 89 in Missouri and 69 in Oklahoma. Average Medicaid share increased by 2.88 percentage points across the two treated states while declining by 1.10 points in comparison states, a 3.99-point difference after rounding. The contrast was larger in Oklahoma than in Missouri. Operating margins decreased in the treated states and also fell in the comparison states, yielding a relative decline of 3.20 points. The reference probabilities were .089 for the Medicaid share and .289 for margin. Neither outcome crossed a .05 reference threshold, and the nonrandom adoption process further limits the inferential interpretation.

Table 5: Medicaid expansion difference-in-differences estimates

Outcome	Treated change pp	Comparison change pp	Difference pp	Reference p
Operating margin	-3.90	-0.70	-3.20	.289
Medicaid-day share	+2.88	-1.10	+3.99	.089

Key: The comparison uses 2018–2019 versus 2023–2024. The state-pair reassignment reference distribution contains 45 assignments. pp = percentage points.

The margin gap relative to the 2018–2019 baseline was −1.36 points in 2017 and −1.57 points in 2020. Changing the baseline to 2017–2019 or 2018–2020 gave margin estimates of −2.75 and −2.68 points. Removing one comparison state at a time gave estimates from −3.94 to −2.61 points. These checks show that the observed direction persisted in the specifications evaluated. The size of the pre-period differences, the two-state treatment group, and possible differences in state-specific pandemic recovery weaken attribution. The Medicaid-day share is consistent with a payer-mix change, while the fiscal evidence provides insufficient support for assuming that expansion independently created a reserve budget.

4.3 Out-of-time forecast performance

Table 6: Temporal validation and holdout forecast performance

Forecast	2023 MAE	2024 MAE	2024 RMSE	2024 WAPE %	2024 bias %
Persistence	2,587	2,638	6,106	9.49	-2.08
Ridge	2,470	2,493	5,547	8.96	+0.53
Boosting	2,726	2,784	10,467	10.01	-0.54

Key: MAE and RMSE are 365-day-equivalent patient-days. The 2024 holdout contains 4,155 hospitals across 50 states and the District of Columbia.

Ridge regression had the lowest validation MAE and was carried forward for allocation. It improved the holdout MAE by 145 patient-days per hospital, or 5.5%, relative to persistence. The paired 95% state-bootstrap interval for ridge minus persistence was −245 to −43 patient-days. Ridge also reduced test WAPE from 9.49% to 8.96% and had a signed

bias of +0.53%. Its MAE was 3,477 patient-days in urban hospitals and 1,710 in rural hospitals. WAPE was 9.15% and 8.69%, respectively, which puts the difference in absolute error in the context of different workload volumes. The validation-derived nominal 80% interval contained 81.6% of test observations. These metrics describe the 2024 holdout and do not establish performance under a future shock.

4.4 Modeled allocation on observed workload

At an 80% planning threshold and a 3% reserve share, each policy assigned 5.61 million annualized reserve bed-days across the eligible hospitals. The proportional policy left a 3.26 million bed-day gap under the scoring rule. The point-forecast program reduced that gap to 2.11 million, and the scenario program reduced it to 1.79 million. The scenario result represents a 45.2% modeled reduction relative to the proportional benchmark. The conditional state-bootstrap interval was 35.6% to 57.3%. The rural gap declined from 1.58 million to 0.83 million bed-days, while the number of hospitals with a positive annual gap decreased from 314 to 238. These counts are scored comparisons of assignments against recorded workload.

Table 7: Retrospective allocation scores at the primary scenario

Allocation policy	Total gap bed-days	Rural gap bed-days	Unused reserve bed-days	Sites with gap
Proportional	3,263,046	1,576,783	4,412,416	314
Point forecast LP	2,108,367	1,074,275	3,257,736	278
Scenario LP	1,786,566	832,176	2,935,936	238

Key: The planning threshold was 80%. Each policy allocated 5,606,361 annualized reserve bed-days. Outcomes reflect a modeled capacity proxy evaluated against observed workload.

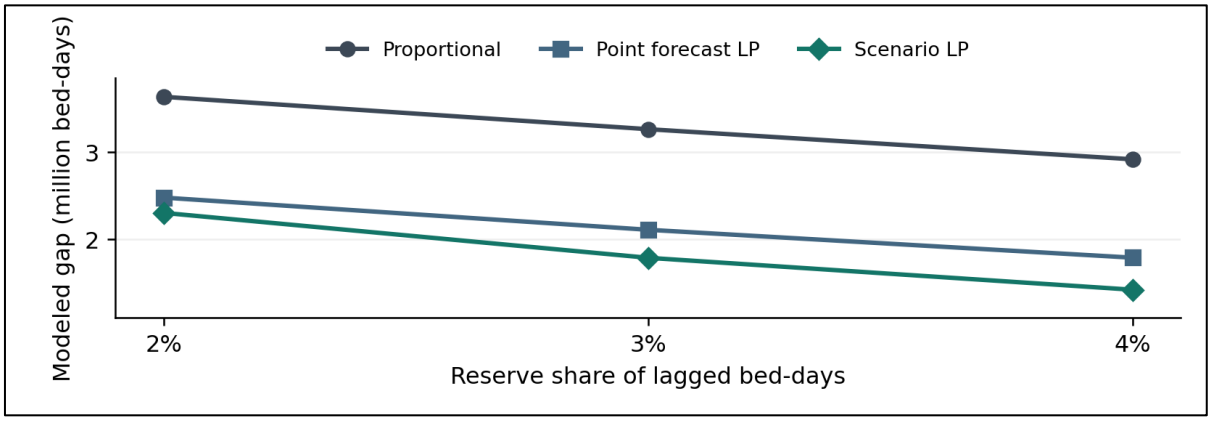

Figure 3. Modeled capacity gaps across reserve budgets at an 80% planning threshold

Key: The curves represent retrospective scenarios scored on FY2024 utilization. Lower values indicate less workload above the modeled capacity threshold.

Across the nine combinations of planning threshold and reserve share, scenario allocation reduced the modeled gap relative to proportional allocation by 26.3% to 54.5%. The remaining gap was sensitive to the threshold. At a 3% reserve, it was 3.42 million bed-days under a 75% threshold and 0.87 million under an 85% threshold. A lower threshold preserves more nominal headroom in the base and increases the modeled shortfall at the same reserve budget. Under the primary scenario the scenario program left 2.94 million allocated bed-days unused by the annual scoring rule, compared with 4.41 million under proportional allocation. Reserve can still have value for daily surges even when the annual proxy scores it as unused. Figure 3 displays the reserve-budget comparison at the 80% threshold.

5 DISCUSSION

5.1 The value of separating policy and operational evidence

The policy analysis produced a larger relative change in Medicaid-day share than the statistical reference distribution would typically generate, but the degree of change differed materially between the two adopting states. The negative operating-margin contrast occurred alongside substantial pre-period variation. Earlier work reports improvements in some financial outcomes following previous expansion waves (Blavin & Ramos, 2021; Moghtaderi et al., 2020). The present comparison concerns two later adopters, a period of unusual financial and service disruption, and a stable set

of hospitals. The fiscal decision that follows from this evidence is limited to the planning assumption used here. Reserve capacity in the model is funded within the specified envelope without drawing on an estimated policy dividend.

The distinction between operating margin and net income also matters for health-system managers. A hospital may retain positive net income while patient services generate an operating deficit. An assessment of resilience should therefore track both measures alongside short-term liquidity and separately identify one-time financial support. The annual current ratio offers a view of balance-sheet coverage at the reporting date. It can miss intrayear cash pressures and the timing of receivables and debt obligations. A reserve budget should be based on available funds and marginal deployment costs even when broad profitability measures appear favorable.

5.2 Forecast accuracy and the consequences of decisions

Ridge regression reduced holdout error relative to persistence and gradient boosting. The result is consistent with the strong predictive role of lagged utilization and capacity in the features used. The modest difference in forecasting error nevertheless coincided with a larger modeled difference in where reserve was assigned. The two outcomes answer different questions. Average prediction error weighs every forecast, while the optimization loss responds to where workload crosses a planning threshold. Scenario allocation also reduced the modeled rural gap under the same minimum reserve floor. The observed rural improvement is conditional on the chosen floor and thresholds and cannot establish distributional equity within rural areas.

The national cost and demand series help explain why the capacity problem cannot be reduced to admissions alone. Longer stays absorb bed-days even if the number of episodes remains below an earlier level. In daily operations, staffing schedules, discharge timing and variation by service line determine whether annual capacity can meet short-run demand. The next evaluation should observe those processes directly. That would allow the health system to measure whether reserve deployment improves care access and outcomes under the constraints that actually bind.

Interoperability and staffing also matter when moving from retrospective analysis to a live pilot. In their review of Medicaid information exchange, Manyura et al. (2025) recommend common data standards, governance, and security controls. A hospital network could apply those principles by agreeing on definitions for staffed-bed availability and setting rules for access to shared operational data. Akinyelure's (2022) conference abstract on occupational therapy in Lagos illustrates a separate measurement issue by reporting functional outcomes such as attention and memory. A reserve pilot would need service-specific outcome measures alongside bed-day counts wherever reserve decisions affect access to specialist care.

5.3 Practical Implications and Recommendations

- Maintain separate ledgers for patient operating performance, other income, liquidity and reserve deployment. Set the reserve envelope from committed funds and validate staffing and supply costs at each hospital before authorization. Include travel or transfer costs where the network uses interfacility referrals.
- Test the forecasting and allocation system in a silent pilot using archived data vintages and prospectively defined reporting deadlines. Compare the model's decisions with the health system's existing allocation at the same staffing budget. Monitor forecast error, interval coverage, and allocation performance by hospital type, geography, and service line.
- Evaluate deployment in network clusters, with a randomized phased rollout where feasible. Track staffed-bed hours, ED boarding, transfer delay, canceled procedures, staff overtime, variable costs, and adverse outcomes. Use a prespecified access safeguard for smaller sites and measure whether observed improvements persist across high- and low-demand periods.
- Record overrides, infeasible allocations, realized activation costs, and unintended access changes. A financial evaluation should calculate the net budget impact over a defined follow-up horizon. It should identify which savings accrue to the hospital, the payer, and the wider network before attributing returns to the allocation policy.

5.4 Limitations and Future Research

Cost reports are administrative filings whose values may be revised. Facility changes, reporting conventions and exclusions affect which records are available for analysis. The cost reports include aggregate accounts and annual utilization. They offer no direct measures of unmet demand, daily staffing, patient acuity, unit-level costs, or clinical quality. AHRQ's weighted national estimates represent its community-hospital sampling frames, while HHS state-day statistics reflect a different reporting system and emergency-era coverage. The comparison across sources is contextual.

The policy design has two treated states and a balanced cohort that omits closures and irregular filings. State adoption was nonrandom, differential trends remain plausible, and the financing result is vulnerable to contemporaneous shocks and policy changes. The forecast used revised releases and one held-out target year. Selection into the paired cohort

may favor hospitals with more consistent reporting. The bootstrap intervals summarize variation across observed state clusters and omit future-year, intervention and structural model uncertainty. In the allocation model, one annualized reserve bed-day is treated as transferable budget capacity within a state. Actual beds, staffing credentials and geographic access cannot be pooled so simply. The model holds observed utilization fixed and assumes that each site can deploy its assignment. Validation with daily unit-level data, separate labor and cash budgets, geographic referral constraints and prospective patient outcomes is needed to estimate the effects of implementation.

6 CONCLUSION

Public CMS, HHS and AHRQ data supported a hospital finance and utilization panel, an observational policy comparison, and a temporal workload benchmark. Ridge regression improved the 2024 forecast relative to persistence. The Medicaid expansion analysis showed a relative shift in reported payer composition alongside substantial uncertainty about operating margins. With the same reserve budget, the scenario-based program placed capacity where it covered a larger share of recorded workload under the modeled planning rule. The framework's contribution is the explicit connection between the credibility of financing assumptions, forecast uncertainty, and the performance of a constrained allocation. The reported allocation gains are conditional on the annual scoring proxy. A funded, prospective evaluation would establish whether the approach improves service continuity and operating performance in practice.

STATEMENTS ON COMPLIANCE WITH ETHICAL STANDARDS

Disclosure of conflict of interest

The authors declare no conflicts of interest regarding this manuscript.

REFERENCES

[1] Agency for Healthcare Research and Quality (2026a). *HCUP Fast Stats national trends in emergency department visits through 2023* [Data set]. https://datatools.ahrq.gov/hcup-fast-stats/wp-content/uploads/sites/3/2026/02/FSNationalTrendsEDVisits-dy23Deliverable-28Jan2026-1.xlsx

[2] Agency for Healthcare Research and Quality (2026b). *HCUP Fast Stats national trends in inpatient stays through 2023* [Data set]. https://datatools.ahrq.gov/hcup-fast-stats/wp-content/uploads/sites/3/2026/02/FSNationalTrendsInpatientStays-dy23Deliverable-12Dec2025.xlsx

[3] Akinyelure, F. M. (2022). Music and dance therapy; the effectiveness of occupational therapy intervention in preventing depression amongst persons living with Alzheimer in Lagos Nigeria [Conference abstract]. *Alzheimer's & Dementia, 18*(S8), Article e069054. https://doi.org/10.1002/alz.069054

[4] Akinyelure, F. M. (2023). Bridging the gap: Integrating predictive analytics with culturally competent mental health care delivery in marginalized populations. *International Journal of Research in Psychiatry, 3*(2), 12–17. https://doi.org/10.22271/27891623.2023.v3.i2a.76

[5] Akinyelure, F. (2025). Leveraging behavioural health data for policy innovation: Closing the loop between community insights and public health decision-making. *International Journal of Innovative Science and Research Technology, 10*(7), 3458–3466. https://doi.org/10.38124/ijisrt/25jul1532

[6] Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., & Wager, S. (2021). Synthetic difference-in-differences. *American Economic Review, 111*(12), 4088–4118. https://doi.org/10.1257/aer.20190159

[7] Bertsimas, D., & Kallus, N. (2020). From predictive to prescriptive analytics. *Management Science, 66*(3), 1025–1044. https://doi.org/10.1287/mnsc.2018.3253

[8] Bertsimas, D., Boussioux, L., Cory-Wright, R., Delarue, A., Digalakis, V., Jacquillat, A., Kitane, D. L., Lukin, G., Li, M., Mingardi, L., Nohadani, O., Orfanoudaki, A., Papalexopoulos, T., Paskov, I., Pauphilet, J., Skali Lami, O., Stellato, B., Tazi Bouardi, H., Villalobos Carballo, K., … Zeng, C. (2021). From predictions to prescriptions: A data-driven response to COVID-19. *Health Care Management Science, 24*(2), 253–272. https://doi.org/10.1007/s10729-020-09542-0

[9] Blavin, F., & Ramos, C. (2021). Medicaid expansion: Effects on hospital finances and implications for hospitals facing COVID-19 challenges. *Health Affairs, 40*(1), 82–90. https://doi.org/10.1377/hlthaff.2020.00502

[10] Bracher, J., Ray, E. L., Gneiting, T., & Reich, N. G. (2021). Evaluating epidemic forecasts in an interval format. *PLOS Computational Biology, 17*(2), e1008618. https://doi.org/10.1371/journal.pcbi.1008618

[11] Bravata, D. M., Perkins, A. J., Myers, L. J., Arling, G., Zhang, Y., Zillich, A. J., Reese, L., Dysangco, A., Agarwal, R., Myers, J., Austin, C., Sexson, A., Leonard, S. J., Dev, S., & Keyhani, S. (2021). Association of intensive care unit patient load

and demand with mortality rates in US Department of Veterans Affairs hospitals during the COVID-19 pandemic. *JAMA Network Open, 4*(1), Article e2034266. https://doi.org/10.1001/jamanetworkopen.2020.34266

[12] Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics, 225*(2), 200–230. https://doi.org/10.1016/j.jeconom.2020.12.001

[13] Centers for Medicare & Medicaid Services (2020). *OK-21-0001* [Medicaid state plan amendment]. https://www.medicaid.gov/medicaid-spa/2020-12-04/104311

[14] Centers for Medicare & Medicaid Services (2021). *MO-21-0029* [Medicaid state plan amendment]. https://www.medicaid.gov/medicaid-spa/2022-08-26/139231

[15] Centers for Medicare & Medicaid Services (2026). *Hospital provider cost report* [Data set]. https://data.cms.gov/provider-compliance/cost-reports/hospital-provider-cost-report

[16] Collins, G. S., Moons, K. G. M., Dhiman, P., Riley, R. D., Beam, A. L., Van Calster, B., Ghassemi, M., Liu, X., Reitsma, J. B., van Smeden, M., Boulesteix, A.-L., Camaradou, J. C., Celi, L. A., Denaxas, S., Denniston, A. K., Glocker, B., Golub, R. M., Harvey, H., Heinze, G., … Logullo, P. (2024). TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ, 385*, Article e078378. https://doi.org/10.1136/bmj-2023-078378

[17] Elmachtoub, A. N., & Grigas, P. (2022). Smart "predict, then optimize". *Management Science, 68*(1), 9–26. https://doi.org/10.1287/mnsc.2020.3922

[18] French, G., Hulse, M., Nguyen, D., Sobotka, K., Webster, K., Corman, J., Aboagye-Nyame, B., Dion, M., Johnson, M., Zalinger, B., & Ewing, M. (2021). Impact of hospital strain on excess deaths during the COVID-19 pandemic—United States, July 2020–July 2021. *Morbidity and Mortality Weekly Report, 70*(46), 1613–1616. https://doi.org/10.15585/mmwr.mm7046a5

[19] Grinsztajn, L., Oyallon, É., & Varoquaux, G. (2022). Why do tree-based models still outperform deep learning on typical tabular data? *Advances in Neural Information Processing Systems, 35*, 507–520. https://proceedings.neurips.cc/paper_files/paper/2022/hash/0378c7692da36807bdec87ab043cdadc-Abstract-Datasets_and_Benchmarks.html

[20] Hernán, M. A., Wang, W., & Leaf, D. E. (2022). Target trial emulation: A framework for causal inference from observational data. *JAMA, 328*(24), 2446–2447. https://doi.org/10.1001/jama.2022.21383

[21] Kadri, S. S., Sun, J., Lawandi, A., Strich, J. R., Busch, L. M., Keller, M., Babiker, A., Yek, C., Malik, S., Krack, J., Dekker, J. P., Spaulding, A. B., Ricotta, E., Powers, J. H., Rhee, C., Klompas, M., Athale, J., Boehmer, T. K., Gundlapalli, A. V., … Warner, S. (2021). Association between caseload surge and COVID-19 survival in 558 U.S. hospitals, March to August 2020. *Annals of Internal Medicine, 174*(9), 1240–1251. https://doi.org/10.7326/m21-1213

[22] KFF (2026). *Status of state action on the Medicaid expansion decision*. State Health Facts. Retrieved November 10, 2026, from https://www.kff.org/affordable-care-act/state-indicator/state-activity-around-expanding-medicaid-under-the-affordable-care-act/

[23] MacKinnon, J. G., & Webb, M. D. (2020). Randomization inference for difference-in-differences with few treated clusters. *Journal of Econometrics, 218*(2), 435–450. https://doi.org/10.1016/j.jeconom.2020.04.024

[24] Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2022). M5 accuracy competition: Results, findings, and conclusions. *International Journal of Forecasting, 38*(4), 1346–1364. https://doi.org/10.1016/j.ijforecast.2021.11.013

[25] Manyura, S., Alo, A., Saaka, D. K., & Ruth, K. N. (2025). A methodological framework for developing AI-enabled health information interoperability solutions to improve secure health information exchange across Medicaid ecosystems. *World Journal of Advanced Research and Reviews, 28*(2), 2667–2679. https://doi.org/10.30574/wjarr.2025.28.2.3712

[26] Mehrotra, S., Rahimian, H., Barah, M., Luo, F., & Schantz, K. (2020). A model of supply-chain decisions for resource sharing with an application to ventilator allocation to combat COVID-19. *Naval Research Logistics, 67*(5), 303–320. https://doi.org/10.1002/nav.21905

[27] Moghtaderi, A., Pines, J., Zocchi, M., & Black, B. (2020). The effect of Affordable Care Act Medicaid expansion on hospital revenue. *Health Economics, 29*(12), 1682–1704. https://doi.org/10.1002/hec.4157

[28] Moons, K. G. M., Damen, J. A. A., Kaul, T., Hooft, L., Andaur Navarro, C., Dhiman, P., Beam, A. L., Van Calster, B., Celi, L. A., Denaxas, S., Denniston, A. K., Ghassemi, M., Heinze, G., Kengne, A. P., Maier-Hein, L., Liu, X., Logullo, P., McCradden, M. D., Liu, N., … van Smeden, M. (2025). PROBAST+AI: An updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods. *BMJ, 388*, Article e082505. https://doi.org/10.1136/bmj-2024-082505

[29] Rambachan, A., & Roth, J. (2023). A more credible approach to parallel trends. *Review of Economic Studies, 90*(5), 2555–2591. https://doi.org/10.1093/restud/rdad018

[30] Roth, J., Sant'Anna, P. H. C., Bilinski, A., & Poe, J. (2023). What's trending in difference-in-differences? A synthesis of the recent econometrics literature. *Journal of Econometrics, 235*(2), 2218–2244. https://doi.org/10.1016/j.jeconom.2023.03.008

[31] Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics, 225*(2), 175–199. https://doi.org/10.1016/j.jeconom.2020.09.006

[32] U.S. Department of Health and Human Services (2024). *COVID-19 reported patient impact and hospital capacity by state timeseries (raw)* [Data set]. https://healthdata.gov/Hospital/COVID-19-Reported-Patient-Impact-and-Hospital-Capa/g62h-syeh

[33] Wang, Y., Bai, G., & Anderson, G. (2022). COVID-19 and hospital financial viability in the US. *JAMA Health Forum, 3*(5), Article e221018. https://doi.org/10.1001/jamahealthforum.2022.1018