Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Does integrated county-level healthcare reform improve health resource allocation? Evidence from a difference-in-differences analysis of 121 counties in Hunan, China

Qianyue Su

ssquri@163.com

Hunan University of Chinese Medicine

Xinbin Xia

Hunan University of Chinese Medicine

Research Article

Keywords: integrated care, health resource allocation, difference-in-differences, rural health, China, health workforce

Posted Date: November 16th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11328368/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background Equitable allocation of health resources between urban and rural areas remains a central challenge in China. China's compact county medical community (CMC) reform, piloted nationally since 2019, vertically integrates county, township, and village providers under unified governance. Rigorous county-level evidence on whether this integration reallocates health resources remains scarce.
Methods Using a two-way fixed effects difference-in-differences (DID) design, we analyzed a balanced panel of 121 county-level units in Hunan Province over 2018–2024 (847 county-year observations): 19 pilot counties operational since 2020 versus 102 non-pilot counties, excluding 2022 second-batch counties from the comparison group. Outcomes were licensed (assistant) physicians per 1,000 population, hospital beds per 1,000 population, and the nurse-to-doctor ratio. Robustness checks included event-study dynamics, a 500-permutation placebo test, and propensity score matching DID (PSM-DID).Results CMC adoption was associated with a relative decline of 0.259 licensed (assistant) physicians per 1,000 population ($p < 0.05$; −0.243 with covariates), no significant change in hospital beds per 1,000 population (+0.012/+0.008), and a negative but non-significant change in the nurse-to-doctor ratio (−0.031/−0.028). Results survived placebo tests (only 2 of 500 pseudo-estimates as extreme) and PSM-DID. The decline was directionally larger in less-developed Greater Western Hunan, though underpowered.Conclusions Organizational integration alone may not increase county-level workforce supply; without complementary incentives, resource-sinking policies can coincide with physician drain. Policymakers pursuing integrated care in China and other low- and middle-income countries should pair structural integration with targeted workforce retention mechanisms.

Background

Achieving an equitable geographic distribution of health resources remains one of the most intractable problems facing health systems in low- and middle-income countries (LMICs). Across much of the world, hospitals, beds, and—most critically—skilled health workers concentrate in cities, while rural populations depend on thin networks of primary care facilities with limited capacity [1, 2]. The result is inequitable access, delayed treatment, and financial hardship for rural households, and the pattern has resisted decades of supply-side investment.

In response, integration has become a dominant international reform paradigm. The World Health Organization has promoted integrated people-centred health services as a core strategy for strengthening primary care and improving continuity across levels of the delivery system [3]. Vertical integration—linking hospitals with primary and community providers under shared governance, budgets, or information systems—has been tried from U.S. integrated delivery systems to England's integrated care systems. Yet evaluations are sobering: integration does not mechanically produce resource reallocation, cost savings, or better patient flow [4]. Whether integration reforms actually shift resources toward underserved areas, and under what institutional conditions, remains an open question of direct relevance to LMICs adopting similar county- or district-level models.

Health system reform in China and the search for a workable tiered delivery system

China's 2009 national healthcare reform achieved near-universal insurance coverage and substantially expanded service capacity, but it did not resolve the system's structural "inverted pyramid," in which high-quality resources concentrate in large urban hospitals while patient needs spread across the population [5]. To correct this mismatch, China established the hierarchical diagnosis and treatment system as a central institutional arrangement of health reform, articulating its goal in a sixteen-character formula: first contact at the primary level, two-way referral, separation of acute and chronic care, and vertical coordination between levels [6]. A 2018 "four separations" pathway (by region, urban–rural level, level of care, and acute versus chronic care) further framed the functional repositioning of county, township, and village facilities.

Implementation has been difficult: large public hospitals continue to attract patients and resources, while primary care institutions—constrained by limited capacity, drug formularies, and public trust—struggle to act as "gatekeepers." The stylized facts are persistent: the share of total visits delivered by primary care institutions was 53.1% in 2018 and 51.8% in 2023, hovering in a narrow 50%–53% band for years; drawing on national data, Zhang et al. [7] documented a decline from 56.44% in 2015 to 51.73% in 2022. Hunan Province, the setting of this study, exhibits the same pressures: in 2024, township health centers operated at a bed occupancy rate of 58.64% against 81.65% for county-level hospitals—a widening gap between the county anchor and the grassroots safety net.

Against this backdrop, county-level integration emerged as China's organizational fix: in 2017, the General Office of the State Council called for building medical alliances, including county-level medical communities; in 2019, the National Health Commission formally launched the national pilot of the compact county medical community (CMC) [8]; and in December 2023, ten national departments jointly issued guidance moving CMC construction from piloting to nationwide rollout [9]. Evaluating whether the reform delivers its promised resource reallocation is therefore timely and policy-relevant.

Institutional background: the compact county medical community

The CMC is the core organizational vehicle for restructuring county health delivery in China. Its defining feature is the "six unifications"—unified management of administration, personnel, finance, operations, pharmaceuticals, and information systems across county, township, and village facilities. Typically, a county public hospital serves as the anchor, township health centers as the pivot, and village clinics as the foundation; the three tiers are reorganized into a single legal entity or a compact management committee, transforming formerly competing institutions into a community of shared responsibility, management, service, and interest. This distinguishes the CMC from the earlier, loosely coupled medical alliance, in which collaboration rested on voluntary pairing assistance without unified authority.

Operationally, the model rests on four levers: unified governance aligns decision-making across the three tiers; resource sinking channels talent, technology, and equipment downward through county-employed, township-deployed personnel schemes, specialty co-construction, and equipment sharing; payment leverage uses global budget payment under social health insurance with surplus retention ("total prepayment with retained savings") to convert the insurance fund from hospital revenue into a health cost to be managed, inducing the anchor hospital to keep patients at the grassroots; and information integration, built on regional health information platforms, enables mutual recognition of test results and coordinated two-way referral.

Hunan Province was among the earliest central-China provinces to explore county medical communities. Nineteen of its counties (county-level cities and districts) were selected for the first national pilot batch in 2019, were formally launched in 2020, and completed their first full operational year in 2021. A second batch of Hunan counties joined in the 2022 expansion; because these counties were treated within our observation window, we exclude them from the comparison group rather than misclassify them as untreated—a choice reported explicitly for transparency. Complementing the statistical record, we coded eight provincial-level CMC policy documents issued in Hunan (146 clauses in total) using a policy-instrument framework; double coding yielded a Cohen's kappa of 0.832 at the secondary-category level and an observed agreement rate of 84.93%.

Theoretical framework

Our interpretation draws on two complementary theories. Holistic (whole-of-government) governance theory, developed to remedy the fragmentation left by new public management [10], explains how the six unifications break administrative barriers among the three tiers and how unified information platforms dissolve data silos—providing the theoretical basis for the resource-sinking and information-integration mechanisms. Stakeholder theory [11] treats CMC construction as a profound redistribution of interests among the county government, the medical insurance bureau, the anchor hospital, township health centers, village doctors, and patients; without incentive-compatible arrangements—above all global budget payment with surplus retention—an alliance easily degenerates into linkage without integration. Together, the two theories support an analytical framework in which the CMC policy shock is transmitted through three mechanisms—resource sinking, payment leverage, and information integration—to the allocation response observable in county-level resource stocks.

What is known: three strands of literature

The first strand evaluates integrated care and bundled payment internationally. Reviewing U.S. integrated delivery systems, Hwang et al. [12] found organizational integration broadly associated with quality improvement but limited cost reduction; Colla et al. [13] showed that clinical rather than financial integration drives spending reductions for complex patients. The closest international analog to "total prepayment with retained savings" is the Medicare Shared Savings Program: McWilliams et al. [14], using

a difference-in-differences design, estimated first-year savings of only about 1.4%, larger for independent primary care groups than for hospital-integrated groups, and Kahn and Sullivan [15] found net savings to Medicare negligible once shared-savings payments were netted out—evidence that incentive intensity determines savings. On gatekeeping, Greenfield et al. [16] documented the double-edged effects of general practitioner gatekeeping in England, and Vedsted and Olesen [17] cautioned that strict referral thresholds can delay cancer diagnosis. The common lesson is that integration reshapes patient flows only when primary care capacity, incentive intensity, and referral design are aligned.

A second strand evaluates China's hierarchical diagnosis and treatment reforms. Quasi-experimental evidence is more cautious than descriptive accounts. Wang et al. [18], exploiting the gradual 2004–2018 pilots in a difference-in-differences design, found that the policy's correction of the "inverted pyramid" was transient and faded over time, with public hospital visit rates falling significantly by 3.1% only where community health service capacity was strong—identifying grassroots capacity as a boundary condition and motivating our focus on the CMC as an organizational carrier. Patient-side studies document persistent upward-seeking behavior: willingness to make first contact at primary care institutions is limited and is shaped by personal income, perceived technical level, reimbursement rates, distance, and the presence of visiting specialists [19]; implementation difficulties have also been attributed to benefit imbalances and obstructed referral interfaces among providers.

A third strand measures the equity of health resource allocation, typically using inequality indices or concentration and decomposition methods to describe the distribution of beds, workforce, and financing across regions [20]. This literature documents levels and trends but rarely identifies the causal effect of a specific reform on county-level resource stocks.

Within the growing CMC literature, Hunan's experience remains case-based: Sangzhi County's "seamless" model built on eight unifications reportedly raised the in-county visit rate from 84.6% in 2019 to 91.5% in 2024 while turning the county insurance fund from deficit to an accumulated surplus of 577 million yuan [21]; and counties such as Fenghuang, Ningxiang, and Liling illustrate diverse governance arrangements, ranging from government-led integration with total budget control, surplus retention, and shared risk to a private-hospital-anchored variant. Policy-process studies have framed the reform as a payment-centered "Chinese solution" for integrated delivery, reconstructed the institutional entrepreneurship logic of local innovation, shown that policy formulation is dominated by health departments, insurance bureaus, and large public hospitals while grassroots institutions exert little influence, and warned that positive and negative spillovers coexist. Quasi-experimental evaluations remain few: Zhu et al. [22] found improvements in service utilization and insurance payment structure in one province's pilot; micro-level evidence suggests that integrated health management within county medical consortia can improve self-management among rural patients with chronic disease [23]; and a recent village-clinic-level analysis in Sichuan combined propensity score matching with DID [24], confirming DID's suitability for Chinese health policy evaluation [25].

Three gaps define our research space. First, quasi-experimental evidence on county-level resource stock allocation is missing: existing evaluations emphasize service utilization, costs, and insurance flows rather than the supply-side stock of beds and health workers. Second, negative or unintended effects are rarely identified, even though case evidence warns that anchor hospitals may siphon resources from the grassroots. Third, the "mechanism black box"—which of the resource, payment, and information levers actually transmits effects—remains unopened, and within-province regional heterogeneity has been neglected by a literature concentrated on the developed eastern coast.

The present study

This study provides the first province-wide quasi-experimental evaluation of the CMC reform's effect on county-level health resource allocation. Using a classical two-way fixed effects DID design, we compare the 19 first-batch pilot counties of Hunan with 102 non-pilot counties on a balanced 2018–2024 panel (847 county-year observations), excluding the 2022 second-batch counties from the comparison group. Outcomes are licensed (assistant) physicians per 1,000 population, hospital beds per 1,000 population, and the nurse-to-doctor ratio. We assess robustness with an event-study specification (window k = − 2 to + 4, base year 2019), a placebo test with 500 random permutations (pseudo-coefficients centered near zero with a standard deviation of 0.104; the true estimate of − 0.259 lies at roughly the 0.4th percentile of the null distribution, with only 2 of 500 draws as extreme), and PSM-DID (1:1 nearest-neighbor matching on per-capita GDP, resident population, and urbanization rate, with all standardized mean differences below 10%). Heterogeneity is examined across Hunan's four economic regions (Changsha–Zhuzhou–Xiangtan, Dongting Lake, Southern Hunan, Greater Western Hunan).

Our contributions are threefold. First, we supply the first DID assessment covering all 121 counties of a central-China province, extending causal evidence beyond the eastern coastal cases dominating the literature. Second, we robustly document an unintended negative effect—a relative decline in licensed (assistant) physicians per 1,000 population—and interpret it through the resource-sinking, payment-leverage, and information-integration mechanisms, combining quantitative checks with policy-text and case evidence. Third, we derive implications for other LMICs pursuing county- or district-level integration: organizational integration is not a substitute for workforce policy, and resource-sinking mandates require retention incentives to avoid perverse flows.

Two commitments to transparency deserve emphasis. The parallel-trends test has limited statistical power because only two pre-treatment years (2018 and 2019) are observable, so we read pre-trend diagnostics alongside the placebo and matching evidence. The regional heterogeneity analysis—in which the physician-density decline is directionally larger in the less-developed Greater Western Hunan region—offers directional evidence only, as subgroup cells are too small for adequate power. Likewise, the mechanism analysis is interpretive by design: where county-level data do not permit causal quantification (notably for payment leverage and information integration), we make no quantitative claims beyond what the data support.

The remainder of the paper proceeds as follows. Section 2 describes the data, variables, and empirical strategy. Section 3 reports baseline estimates, robustness checks, and regional heterogeneity. Section 4 interprets possible mechanisms. Section 5 discusses policy implications and limitations and concludes.

Methods

2.1 Study setting

Hunan is a large inland province in central China whose county-level health system has been reshaped since 2019 by the compact county medical community (CMC; □□□□□□□□) reform, a nationally piloted integration scheme that vertically consolidates county, township, and village providers under unified governance, resource sinking, payment leverage, and information integration. Four stylized facts from official statistics frame the empirical problem [26–28].

First, patient flows remained concentrated upward while the primary-care tier was under strain. In 2024 the province recorded 395.65 million outpatient and inpatient visits, 8.32 million fewer than in 2023. Tertiary hospitals delivered 87.04 million visits (22.00 percent of the total, up 5.49 million year on year), whereas visits to primary/secondary hospitals and primary-care institutions combined fell by 12.73 million. Within county systems, county hospitals handled 59.44 million visits (up 0.56 million) with a bed occupancy rate of 81.65 percent (up 4.33 percentage points), while township health centers handled 60.57 million visits (down 5.80 million) at an occupancy rate of only 58.64 percent (down 2.24 percentage points), and village clinics handled 73.45 million visits (down 3.69 million).

Second, the primary-care share of total service utilization stagnated. Nationally, the primary-care share of visits hovered within a narrow 50–53 percent band: 53.1 percent in 2018, 52.0 percent in 2019, a brief rebound to 53.2 percent in 2020, and 50.2, 50.7, and 51.8 percent in 2021–2023, respectively. The "upward concentration, pressured grassroots" pattern is thus a national trend rather than a Hunan-specific anomaly, so rigorous causal identification of the CMC effect is essential.

Third, payment reform began to restructure costs. By 2025, disease-based inpatient payment covered 3,974 designated providers (95.5 percent of all designated facilities) and accounted for 38.456 billion yuan in inpatient fund expenditure, or 90 percent of total settled inpatient spending. Within the disease-based payment scope, drug costs fell by 1.158 billion yuan relative to 2024, medical service fees rose by 1.433 billion yuan, patient out-of-pocket spending fell by 214 million yuan, and the out-of-pocket share dropped by 2 percentage points. Yet a steep fee gradient persisted across facility tiers: in 2024 the average outpatient visit cost 360.57 yuan in hospitals versus 128.50 yuan in community health centers and 86.71 yuan in township health centers, and the corresponding inpatient figures were 8,584.41, 2,459.24, and 1,715.56 yuan. Drugs accounted for roughly 67 percent of per-visit outpatient spending at primary-care institutions (67.27 and 67.34 percent), about twice the hospital share of 32.94 percent, indicating that lower-tier providers remained disproportionately dependent on drug revenue.

Fourth, resource endowments divergently across regions. In 2024, Changsha had approximately 4.25 licensed (assistant) physicians per 1,000 population and 89,861 hospital beds, whereas Zhangjiajie had about 3.12 physicians per 1,000 and only 10,697 beds; the entire health technical workforce of Xiangxi Prefecture (22,782 persons) was less than one-fifth of Changsha's. Per-capita GDP in Changsha reached 144,525 yuan, while Xiangxi, Zhangjiajie, and similar areas fell below one-third of that level. These gradients may condition CMC implementation capacity and motivate the heterogeneity analysis below.

2.2 Research hypotheses

Guided by the framework of "policy shock → transmission mechanisms → tiered healthcare delivery outcomes," we tested the hypotheses stated in the Chinese manuscript:

- **H1.** CMC adoption significantly affected county-level health resource allocation. Specifically, relative to non-pilot counties, pilot counties experienced significant changes in hospital beds per 1,000 population, licensed (assistant) physicians per 1,000 population, and the nurse-to-doctor ratio; the direction and magnitude were determined by the baseline estimates rather than presumed ex ante.
- **H1a (regional heterogeneity).** The resource-allocation effect of CMC adoption differed significantly across regions at different stages of economic development; its direction was left to subgroup regressions, since two competing logics are plausible—stronger implementation capacity in richer areas versus larger marginal room for improvement in poorer ones.
- **H1b (leadership-model heterogeneity).** Conditional on adequate sample size, the resource-allocation effect differed between communities led by public county hospitals and those led by private hospitals; where private-led pilots were too few for subgroup regression, only descriptive case comparisons (e.g.Liling) were reported.

2.3 Data and sample

We constructed a balanced county-level panel for 2018–2024 from the county-by-county statistical tables of the *Hunan Statistical Yearbook*, 2019–2025 volumes [28]. Population and urbanization came from the county economic-and-social indicator tables; GDP and per-capita GDP from the county GDP tables; and health institutions, beds, health technical personnel, licensed (assistant) physicians, and registered nurses from the county health resource tables. All variables are reproducible from publicly available yearbook tables; no restricted insurance-claims or application-based data were used.

The treatment group comprised the 19 Hunan counties (county-level cities and districts) named in the first national CMC pilot list [29]: Ningxiang, Liling, Xiangxiang, Shaoshan, Xiangtan County, Qidong, Xinshao, Pingjiang, Linli, Li County, Jinshi, Sangzhi, Heshan District of Yiyang, Nan County, Lanshan, Linwu, Lianyuan, Tongdao Dong Autonomous County, and Luxi. The treatment year was set to 2020: the pilot list was published in 2019, the communities were formally inaugurated in 2020, and 2021 was the first full operating year. Post equals 1 for 2020 and thereafter, and 0 for 2018–2019. Counties admitted in the second (2022) expansion batch were excluded entirely to avoid contaminating the control group:

loosely organized medical alliances formed after 2017 meant that "not in the tight pilot" did not imply "untouched by integration."

The final sample is a balanced panel of 121 county units (19 treated, 102 controls) over 7 years, yielding 847 county-year observations. Municipal districts and province-administered units ineligible for county CMC pilots were dropped; sporadic missing values were linearly interpolated; and all monetary measures were deflated to 2018 prices.

2.4 Variables

Three outcome variables captured county health resource allocation, the dimension for which consistent county-level public statistics exist: (i) **hospital beds per 1,000 population** (Beds_pc), defined as the number of beds in county health institutions divided by the year-end resident population, measuring hardware agglomeration; (ii) **licensed (assistant) physicians per 1,000 population** (Doctors_pc), the key observable for whether high-quality resources sink toward the grassroots; and (iii) the **nurse-to-doctor ratio** (ND_Ratio), defined as registered nurses divided by licensed (assistant) physicians, capturing workforce structure, which the county-employed, township-deployed personnel scheme may shift. The core regressor was the interaction Treat_i × Post_t. Time-varying county controls comprised the logarithm of per-capita GDP (lnGDP_pc), capturing income effects on demand and fiscal capacity for health investment; the logarithm of year-end resident population (lnPop), capturing scale-driven demand pressure; and the urbanization rate (Urban). Fiscal health expenditure shares and insurance enrollment rates are unavailable in county-level yearbooks and were not entered directly; their influence is partially absorbed by the fixed effects. Descriptive statistics appear in Table 1 (Section 3).

2.5 Empirical strategy

Because all treated counties adopted the reform at the same time, we estimated the policy effect with a classical two-way fixed effects (TWFE) difference-in-differences model [30]:

$$Y_{it} = \alpha + \beta\,(Treat_i \times Post_t) + \gamma\, X_{it} + \mu_i + \delta_t + \varepsilon_{it} \quad (1)$$

where *i* indexes counties and *t* indexes years; *Y* is one of the three allocation outcomes; β is the parameter of interest, the net effect of CMC adoption; *X* is the vector of controls; μ_i and δ_t are county and year fixed effects; and ε is the disturbance. Standard errors were clustered at the county level to accommodate within-county serial correlation. With a common treatment date, the TWFE estimator is not subject to the bias arising from heterogeneous treatment timing, so heterogeneity-robust estimators were not required; we nonetheless note this literature for completeness [31, 32].

To assess parallel trends and effect dynamics directly, we estimated an event-study specification:

$$Y_{it} = \alpha + \Sigma_{k=-2,\, k\neq -1}^{+4} \beta_k D_{it}^k + \gamma\, X_{it} + \mu_i + \delta_t + \varepsilon_{it} \quad (2)$$

where $D_{it}^{k} = 1\{t - E_i = k\}$ indicates that county i is k years relative to its treatment year E. The window spans two pre-treatment periods through four post-treatment periods, with k = − 1 (the year 2019) as the omitted base period.

Two design limitations merit emphasis. First, pilot assignment was plausibly non-random. We therefore re-estimated the model on a propensity-score-matched sample, matching each treated county 1:1 to its nearest neighbor on baseline (2018) per-capita GDP, resident population, and urbanization rate; covariate balance was verified by standardized bias below 10 percent and the common-support condition. Second, to rule out unobserved shocks, we ran a placebo test with 500 random permutations, each drawing 19 pseudo-treated counties while keeping the real treatment timing. Additional robustness checks excluded five counties adjacent to the Changsha–Zhuzhou–Xiangtan metropolitan core to purge spillovers from large-city hospitals; replaced the outcomes with log beds, log health technical personnel, and log registered nurses; and computed wild cluster bootstrap p-values for county-clustered inference.

Identification rests on the parallel-trends assumption: absent the reform, treated and control counties would have followed common outcome trajectories. Because the panel begins in 2018, only two pre-treatment periods are observable, so the pre-trends test has inherently limited power; we interpret the event-study coefficients accordingly and treat causal claims with caution.

Results

3.1 Descriptive statistics

Table 1 reports descriptive statistics for the four counties selected for in-depth field study, which span the province's endowment gradient—from near-top-tier Liling (per-capita GDP 109,831 yuan, about 1.35 times the provincial average; 7.10 beds per 1,000 population) down to Sangzhi, where per-capita GDP (34,188 yuan) was less than half the provincial average and beds per 1,000 population (5.44) and health technical personnel per capita were the lowest of the four. Pingjiang and Lanshan occupied the middle-to-lower range, at roughly 65–69 percent of the provincial per-capita GDP average. This gradient provides the contextual variation exploited in the heterogeneity analysis below. The regression sample underlying all subsequent results is the balanced panel of 847 county-year observations (121 counties over 2018–2024) constructed in Section 2.

Table 1
Descriptive statistics: key indicators of the four field-study counties, 2024

County (city)	Prefecture	Resident population (10,000)	GDP per capita (yuan)	Health institutions (n)	Beds (n)	Beds per 1,000 population	Health technical personnel (n)
Liling	Zhuzhou	86.39	109,831	626	6,135	7.10	6,541
Pingjiang	Yueyang	93.31	56,362	983	5,258	5.63	6,454
Sangzhi	Zhangjiajie	36.78	34,188	369	2,000	5.44	2,336
Lanshan	Yongzhou	31.82	53,376	393	2,545	8.00	2,492
Notes: Source: *Hunan Statistical Yearbook 2025*, county tables 22 − 1, 22 − 2, and 22 − 17; beds per 1,000 population computed against year-end resident population.							

3.2 Parallel trends

Figure 1 plots the event-study coefficients from Eq. (2), with 2019 (k = − 1) as the base period. The pre-treatment coefficient for 2018 (k = − 2) was statistically indistinguishable from zero, providing no evidence that treated and control counties diverged before the reform. Post-treatment coefficients grew in absolute magnitude over time but carried wide confidence intervals, indicating that a policy effect was present while statistical precision was limited by sample size and the length of the observation window. Because only two pre-treatment periods are observable, the power of this pre-trends test is limited, and we read it as supportive rather than definitive evidence.

3.3 Baseline difference-in-differences estimates

Table 2 reports the baseline TWFE-DID estimates. Columns (1)–(3) include only county and year fixed effects; columns (4)–(6) add the time-varying controls (log per-capita GDP, urbanization rate, log year-end resident population). All six specifications used the full 847 observations, with R^2 ranging from 0.786 to 0.918. Two patterns emerged.

First, CMC adoption was associated with a statistically significant relative decline in physician density. The Treat × Post coefficient on licensed (assistant) physicians per 1,000 population was − 0.259 (SE 0.118) without controls and − 0.243 (SE 0.112) with controls, both significant at the 5 percent level. Pilot counties thus experienced a relative decline of roughly 0.26 physicians per 1,000 population over the post-reform period. This pattern is consistent with communities concentrating township physicians at the lead county hospital, and with a denominator effect from resident-population outflow in treated counties.

Second—and equally informative—the estimated effects on the other two outcomes were null. The coefficient on hospital beds per 1,000 population was 0.012 (SE 0.024) without controls and 0.008 (SE 0.023) with controls, neither statistically significant, indicating that the reform did not expand hardware capacity in pilot counties during the sample window. The coefficient on the nurse-to-doctor ratio was −

0.031 (SE 0.022) and − 0.028 (SE 0.021), respectively, negative but insignificant. Taken together, the aggregate effect of CMC adoption on county health resource allocation over 2020–2024 was limited, and the anticipated release of "high-quality resources sinking to the grassroots" was not observed in the allocation totals.

Table 2
Baseline DID estimates of the CMC effect on county health resource allocation

Variable	(1) Beds per 1,000	(2) Physicians per 1,000	(3) Nurse-to-doctor ratio	(4) Beds per 1,000	(5) Physicians per 1,000	(6) Nurse-to-doctor ratio
Treat×Post	0.012	−0.259**	−0.031	0.008	−0.243**	−0.028
	(0.024)	(0.118)	(0.022)	(0.023)	(0.112)	(0.021)
Controls	No	No	No	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	847	847	847	847	847	847
R^2	0.892	0.915	0.786	0.895	0.918	0.789

Notes: County-clustered robust standard errors in parentheses; *,, and * denote significance at the 1, 5, and 10 percent levels. Columns (1) and (4) use hospital beds per 1,000 population; columns (2) and (5) use licensed (assistant) physicians per 1,000 population; columns (3) and (6) use the nurse-to-doctor ratio. Controls include log per-capita GDP, urbanization rate, and log year-end resident population.

3.4 Placebo test

Figure 2 displays the distribution of 500 placebo coefficients obtained by randomly designating 19 counties as pseudo-treated while holding the true treatment timing fixed. The pseudo-effects had a mean of approximately − 0.000 and a standard deviation of 0.104. The true estimate (− 0.259) fell near the 0.4th percentile of the left tail of the placebo distribution, and only 2 of the 500 pseudo-coefficients were as extreme in absolute value. The physician-density decline is therefore unlikely to be driven by random factors or unobserved county heterogeneity.

3.5 Propensity-score matching and further robustness

Because pilot selection may have responded to observable county characteristics, we re-estimated the model on a 1:1 nearest-neighbor matched sample using baseline per-capita GDP, resident population, and urbanization rate as matching variables; matching retained all 19 treated counties against 96 controls within the common-support range. Table 3 shows that the coefficient signs matched the baseline: the effect on physicians per 1,000 population remained negative (− 0.221, SE 0.128) and significant at the 10 percent level, while the effects on beds per 1,000 population (0.009, SE 0.026) and the nurse-to-doctor ratio (− 0.024, SE 0.023) remained insignificant. The slightly weaker significance

indicates that the negative physician effect survived the correction for selection on observables but should be interpreted prudently.

Table 3
PSM-DID robustness estimates (1:1 nearest-neighbor matching)

Matching method	Outcome	Coefficient	Standard error	Significance
Nearest neighbor	Beds per 1,000 population	0.009	(0.026)	Not significant
Nearest neighbor	Licensed (assistant) physicians per 1,000 population	−0.221*	(0.128)	$p < 0.1$
Nearest neighbor	Nurse-to-doctor ratio	−0.024	(0.023)	Not significant

Notes: County-clustered robust standard errors in parentheses; *,, and * denote significance at the 1, 5, and 10 percent levels.

Table 4 reports covariate balance after matching. Standardized biases for all three matching variables were below 10 percent (5.2, 3.8, and 6.1 percent), none of the post-matching t-tests were significant, and the common-support condition held, confirming that treated and matched control counties were comparable on observable characteristics.

Table 4
Covariate balance after propensity-score matching

Matching variable	Treated mean	Control mean	Standardized bias (%)	t-test p-value
Log GDP per capita	10.872	10.795	5.2	0.342
Log year-end resident population	4.231	4.187	3.8	0.418
Urbanization rate	0.452	0.438	6.1	0.287
Post-matching common-support range	0.082–0.915	–	–	–
Post-matching sample size	19	96	–	–

Notes: After matching, standardized biases were below 10 percent for all variables and no t-test was significant, indicating no significant differences in observable characteristics between treated and matched control counties; the common-support assumption was satisfied.

Three further checks left the conclusions unchanged: (a) excluding five counties adjacent to the Changsha–Zhuzhou–Xiangtan metropolitan core to remove spillovers from large-city hospitals preserved the sign of the core coefficients; (b) replacing the outcomes with log beds, log health technical personnel, and log registered nurses reproduced the baseline pattern; and (c) wild cluster bootstrap inference at the county level left the significance conclusions intact.

3.6 Regional heterogeneity

Following Hunan's "one core, two sub-centers, three belts, four zones" spatial development framework, we split the sample into four regional plates—the Changsha–Zhuzhou–Xiangtan, Dongting Lake, Southern Hunan, and Greater Western Hunan regions—and re-estimated the model by subgroup (Table 5). Point estimates for physician density were negative in all four plates, indicating that the direction of the workforce effect was not confined to any single region. The DID coefficients for the Dongting Lake and Southern Hunan regions shared the sign of the full-sample estimates but were not statistically significant. In Greater Western Hunan, which includes the minority autonomous counties of Sangzhi, Tongdao, and Luxi, the decline in physician density (− 0.342, SE 0.165, significant at 5 percent) was larger than the provincial average, plausibly reflecting stronger concentration of grassroots physicians at the prefecture-seat hospitals under community integration. The Changsha–Zhuzhou–Xiangtan estimate lacked precision because the plate contains only a handful of pilot counties (Ningxiang, Liling, Xiangxiang, Shaoshan, and Xiangtan County). One interpretation is that weaker fiscal capacity and limited talent-retention pull in Greater Western Hunan amplified the upward concentration of physicians, whereas better-endowed plates saw more balanced two-way flows. We emphasize that with 19 pilot counties divided across four plates, each subgroup contains few treated units; these results are directional evidence rather than definitive tests of between-group differences, consistent with hypothesis H1a being only partially supported.

Table 5
Regional heterogeneity: DID estimates by regional plate

Variable	Changsha–Zhuzhou–Xiangtan	Dongting Lake	Southern Hunan	Greater Western Hunan
Treat×Post (beds per 1,000)	0.021	0.008	0.015	0.003
	(0.035)	(0.031)	(0.029)	(0.038)
Treat × Post (physicians per 1,000)	−0.192	−0.235	−0.218	−0.342**
	(0.186)	(0.152)	(0.148)	(0.165)
Treat×Post (nurse-to-doctor ratio)	−0.018	−0.025	−0.022	−0.041
	(0.032)	(0.028)	(0.027)	(0.031)
Controls	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes
N	119	252	294	182
R^2	0.921	0.908	0.895	0.912

Notes: County-clustered robust standard errors in parentheses; *,, and * denote significance at the 1, 5, and 10 percent levels. Regional plates follow Hunan's provincial spatial development divisions: Changsha–Zhuzhou–Xiangtan, Dongting Lake, Southern Hunan, and Greater Western Hunan. Subgroup power is limited; estimates are directional evidence only.

Discussion

Principal findings

This study provides quasi-experimental evidence on how China's compact county medical community (CMC) reform reshaped county-level health resource allocation, using a balanced panel of 121 counties (19 pilot and 102 comparison counties) in Hunan Province over 2018–2024 (847 county-year observations), with 2020 as the treatment year. We found that CMC adoption was associated with a statistically significant relative decline in licensed (assistant) physicians per 1,000 population (− 0.259 in the baseline specification and − 0.243 with covariates; both $p < 0.05$). In contrast, we detected no significant effect on hospital beds per 1,000 population (0.012 and 0.008; both $p > 0.1$) and a negative but non-significant change in the nurse-to-doctor ratio (− 0.031 and − 0.028; both $p > 0.1$). Taken together, these findings suggest that organizational integration did not mechanically translate into resource increments at the county level; instead, it coincided with a measurable erosion of physician density in pilot counties.

Comparison with previous work

Our results both complement and diverge from the emerging evaluation literature on China's county-level integration reforms. The closest precedent is a difference-in-differences (DID) analysis of village clinics in rural Sichuan, which examined the CMC policy's effects on healthcare resource allocation and service utilization at the village level [24]. Three differences are worth noting. First, we analyze a county-level panel covering all 121 counties of one province rather than village-level facilities, so our estimates capture aggregate county-system resource stocks rather than frontline utilization. Second, whereas the Sichuan study emphasized service utilization outcomes, our primary outcomes are resource stocks—physicians, beds, and staffing structure. Third, and most importantly, we document a negative effect on physician density, a finding not reported in prior work, which points to a possible unintended consequence of integration reforms that village-level analyses may be unable to detect.

Internationally, our findings resonate with evaluations of integrated care initiatives in high-income settings, where organizational integration—such as integrated care systems in England or accountable care organizations in the United States—has frequently produced modest or mixed short-run effects on utilization, spending, and workforce distribution [33, 34]. This literature cautions that integration changes incentives and governance before it changes observable resource stocks, a sequence our results appear to mirror. Our findings also extend evidence on China's health system reform since 2009, which documented gains in coverage and capacity alongside persistent urban–rural workforce maldistribution [35, 36]; identifying a relative decline in county physician density under a reform explicitly designed to sink resources toward the grassroots adds a cautionary data point to that literature.

Interpretation: possible mechanisms

The CMC reform bundles four design elements—unified governance, resource sinking, payment leverage, and information integration. Guided by holistic governance and stakeholder theory, and following the mediation-analysis conventions recommended for causal inference research [37], we estimated only the effect of CMC adoption on mechanism-related variables (the X→M path) and deliberately refrained from regressions that would place mediators and treatment in the same model. Where county-level settlement or information-system data were unavailable in public sources, we relied on policy documents and publicly reported case evidence. The three pathways below should therefore be read as an interpretive framework for the observed decline in physician density, not as empirically identified causal mechanisms.

Resource sinking and reverse flows. The resource-sinking pathway rests on holistic governance: unified management of administration, personnel, and finance is intended to channel talent, technology, and equipment from county hospitals to township and village facilities through schemes such as the county-employed, township-deployed personnel arrangement, specialty co-construction, and paired assistance. Our estimates indicate that this pathway did not operate in the intended direction during the sample period. One plausible interpretation is that, in the early implementation stage, unified personnel management facilitated an upward flow of grassroots physicians toward lead county hospitals—a siphoning effect—rather than the downward flow policymakers envisioned. The negative but non-significant coefficients on the nurse-to-doctor ratio are consistent with the view that the extension of

nursing resources to the grassroots had likewise not yet materialized. Qualitative evidence from Zhejiang's integrated county healthcare consortia likewise highlights collaboration barriers and personnel incentive problems during implementation [38].

Payment leverage and labor-cost compression. Under stakeholder theory, the global budget payment reform—"global prepayment with retained surplus"—converts insurance funds from hospital revenue into a cost of population health, thereby creating incentives for lead hospitals to control costs and refer stabilized patients downward. Policy documents confirm that global budget rules have been widely established in Hunan (we note a scope caveat below: insurance authorities report 25 county-level jurisdictions under global payment, a different accounting scope from the roughly 24 counties involved in the second-batch pilot expansion). Because county-level data on the share of insurance funds spent at the grassroots and case-based payment coverage are not publicly available, we do not report quantitative tests of this pathway. Interpretively, however, a global budget that makes surplus retention contingent on cost control may also incentivize lead hospitals to compress labor costs and consolidate staffing upward, which is consistent with—though does not by itself prove—the observed decline in county-level physician density. Existing case research suggests that such payment incentives are conditional on surplus scale and on how retained funds are distributed within the alliance [39]; in several pilot counties, bundled payment only began in 2025, so any incentive effects likely lie beyond our observation window.

Information integration and redundancy reduction. The information-integration pathway posits that a unified regional health information platform lowers referral friction and information search costs through interoperable electronic medical records and mutual recognition of test results. County-level data on platform coverage and teleconsultation volumes are likewise unavailable, so we again rely on qualitative evidence. Case materials suggest that while physical connectivity has been achieved in some pilot counties—for example, a unified regional hospital information system in one case county and "Internet + CMC" teleconsultation in another—the effective utilization of teleconsultation and mutual recognition of results remains low, a pattern prior multi-case research describes as "digital suspension" (platforms built but not used). To the extent that interoperability does function, it may reduce the need for redundant staffing at duplicated service points, offering a further interpretation of why integration did not expand the measured physician stock. We do not rank the relative importance of the three pathways, and their potential complementarities, which we could not test with interaction terms, remain theoretical conjecture rather than empirical conclusion.

Policy implications

Our findings carry implications that extend beyond Hunan to other settings pursuing integrated rural healthcare delivery. First, organizational integration should be paired with explicit physician retention incentives. If unified personnel management enables upward concentration of staff, schemes such as county-employed, township-deployed arrangements need dedicated performance allowances, career advancement channels that credit grassroots service, and closed-loop training pipelines (targeted training, service obligations, tuition reimbursement) for shortage specialties such as general practice

and chronic disease management. Without such counterweights, integration risks consolidating rather than redistributing the rural workforce [40].

Second, payment reform and personnel or remuneration systems should be reformed jointly. A global budget with retained surplus can discipline costs, but its incentive compatibility with tiered healthcare delivery depends on how surplus funds are shared with township facilities and individual providers. Composite payment designs—for example, capitation combined with performance rewards tied to the share of care delivered at the grassroots, standardized two-way referral, and chronic disease control targets—may better align cost containment with service quality than blunt expenditure ceilings. Risk-adjustment funds that compensate alliances for catastrophic case costs can further mitigate incentives to avoid or deflect severely ill patients.

Third, information infrastructure investment should shift from platform coverage to business enablement: bringing eligible telemedicine services into insurance reimbursement and building grassroots staff's digital capacity may help close the gap between "built" and "used."

Fourth, integration strategies should be differentiated by regional conditions. In our heterogeneity analysis, the decline in physician density was directionally larger in the less-developed Greater Western Hunan region (an intra-provincial economic division including several minority counties) than the provincial average, though the small number of pilot counties within each region means this evidence is directional only. For underdeveloped regions, priority should go to guaranteeing basic service accessibility through tilted fiscal, staffing, and equipment support, whereas better-resourced regions can pursue deeper vertical integration. This pattern suggests that a one-size-fits-all rollout may widen rather than narrow intra-provincial inequities in health workforce distribution.

Limitations

Several limitations qualify our findings. First, this is a single-province study; Hunan's economic gradient and implementation models may not generalize to other provinces or countries. Second, the observation window ends in 2024, while CMC construction entered a deepening phase from 2024 onward and several payment reforms only took effect in 2025, so longer-run effects remain unobserved. Third, the treatment group comprises only 19 counties, which limits statistical power, particularly for subgroup analyses. Fourth, with only two pre-treatment periods, the parallel-trends and event-study tests have limited power to detect divergent pre-trends. Fifth, mechanism variables are constrained by the public availability of county-level data: insurance settlement structures, platform utilization, and grassroots-tier staffing flows are not published, so two of the three proposed mechanisms could only be discussed qualitatively, and the mechanism measures available are proxies that may diverge from their theoretical constructs. Sixth, the second-batch pilot list could be confirmed for only five counties through public channels, and the insurance-authority scope of global payment (25 county-level jurisdictions) does not coincide with the pilot-expansion scope, which constrains the precision of scope-sensitive analyses. Finally, although we employed placebo tests (500 random permutations), event-study dynamics, and

propensity-score-matched DID, county selection into the pilot may still correlate with unobserved local reform capacity or fiscal strength, so residual selection bias cannot be fully excluded.

Conclusions

Using a difference-in-differences design with 121 counties in Hunan Province, China, over 2018–2024, this study found that the compact county medical community reform was associated with a relative decline of approximately 0.26 licensed (assistant) physicians per 1,000 population in pilot counties, with no significant effects on hospital beds or the nurse-to-doctor ratio. Organizational integration alone did not increase county-level resource stocks; interpretive evidence is consistent with early-stage upward concentration of physicians, conditionality of payment incentives, and under-utilized information platforms. Integration reforms in rural health systems should therefore be accompanied by workforce retention mechanisms, payment designs that share surplus with grassroots providers, and regionally differentiated implementation. Longer observation windows and finer-grained county-level data are needed to verify whether these patterns persist as the reforms mature.

Declarations

Ethics approval and consent to participate:

This study is based on publicly available county-level aggregate statistics from published statistical yearbooks and publicly issued government policy documents. No individual-level or identifiable data were collected; ethics approval and consent to participate were therefore not applicable.

Consent for publication:

Not applicable.

Competing interests:

The authors declare that they have no competing interests.

Authors' information

Su Qianyue (2000-), female, master's student, School of Humanities and Management, Hunan University of Chinese Medicine, mainly engaged in research on health policy and medical community management.

Funding:

2025 Graduate Research Innovation Project of Hunan University of Chinese Medicine, "Research on the Operational Efficiency and Effectiveness of Compact County Medical Communities" (No. 2025CX118);

2026 Policy Research Project of the Monitoring and Statistics Center of the National Administration of Traditional Chinese Medicine: "Research on Pathways and Safeguard Mechanisms for Promoting the Implementation of Hierarchical Diagnosis and Treatment in County-Level Medical Communities from the Perspective of Policy Tools" (No. 2026ZCYJZG1)

Author Contribution

Su was responsible for manuscript drafting and data organization, while Xia was responsible for review and revision. All authors are accountable for this manuscript.

Data Availability

The datasets analysed during the current study are derived from publicly available sources: the Hunan Statistical Yearbook (2019–2025 editions) and the statistical yearbooks/statistical communiqués of Hunan's prefecture-level cities. The list of pilot counties is from publicly issued policy documents of the National Health Commission of China and the Hunan Provincial Health Commission. The compiled county-level panel dataset and the analysis code are available from the corresponding author on reasonable request.

References

1. Xu H, Zhang W, Gu L, Qu Z, Sa Z, Zhang X, et al. Aging village doctors in five counties in rural China: situation and implications. Hum Resour Health. 2014;12:36. 10.1186/1478-4491-12-36.
2. Yu H, Yu S, He D, Lu Y. Equity analysis of Chinese physician allocation based on Gini coefficient and Theil index. BMC Health Serv Res. 2021;21(1):455. 10.1186/s12913-021-06348-w.
3. World Health Organization. Framework on integrated, people-centred health services: report by the Secretariat. Geneva: World Health Organization; 2016. Document A69/39.
4. Stokes J, Lau YS, Kristensen SR, Sutton M. Does pooling health & social care budgets reduce hospital use and lower costs? Soc Sci Med. 2019;232:382–8. 10.1016/j.socscimed.2019.05.038.
5. Yip WC, Hsiao WC, Chen W, Hu S, Ma J, Maynard A. Early appraisal of China's huge and complex health-care reforms. Lancet. 2012;379(9818):833–42. 10.1016/S0140-6736(11)61880-1.
6. General Office of the State Council. Guiding opinions on promoting the construction of the hierarchical diagnosis and treatment system: Guo Ban Fa 〔2015〕 No. 70. Beijing: State Council of the People's Republic of China; 2015. [Policy document in Chinese].
7. Zhang B, Duan L, Li Y, et al. Study on patient visits to primary healthcare institutions in China and their associated factors. Health Econ Res. 2025;42(12):36–40. [Article in Chinese].
8. National Health Commission, National Administration of Traditional Chinese Medicine. Notice on advancing the construction of compact county medical communities: Guo Wei Ji Ceng Han 〔2019〕

No. 121. Beijing; 2019. [Policy document in Chinese].

9. National Health Commission, National Development and Reform Commission, Ministry of Finance, et al. Guiding opinions on comprehensively advancing the construction of compact county medical and health communities: Guo Wei Ji Ceng Fa 〔2023〕 No. 41. Beijing et al. 2023. [Policy document in Chinese].
10. Christensen T, Lægreid P. The whole-of-government approach to public sector reform. Public Adm Rev. 2007;67(6):1059–66. 10.1111/j.1540-6210.2007.00797.x.
11. Freeman RE. Strategic management: a stakeholder approach. Boston: Pitman; 1984.
12. Hwang W, Chang J, LaClair M, Paz H. Effects of integrated delivery system on cost and quality. Am J Manag Care. 2013;19(5):e175–84.
13. Colla C, Yang W, Mainor AJ, et al. Organizational integration, practice capabilities, and outcomes in clinically complex Medicare beneficiaries. Health Serv Res. 2020;55(S3):925–35.
14. McWilliams JM, Hatfield LA, Chernew ME, Landon BE, Schwartz AL. Early performance of accountable care organizations in Medicare. N Engl J Med. 2016;374(24):2357–66. 10.1056/NEJMsa1600142.
15. Kahn JG, Sullivan KR. Promise vs. practice: the actual financial performance of accountable care organizations. J Gen Intern Med. 2021;36(11):3623–5. 10.1007/s11606-021-06898-1.
16. Greenfield G, Foley K, Majeed A. Rethinking primary care's gatekeeper role. J R Soc Med. 2018;111(7):279–80. 10.1177/0141076818785232.
17. Vedsted P, Olesen F. Are the serious problems in cancer survival partly rooted in gatekeeper principles? An ecologic study. Br J Gen Pract. 2011;61(589):e508–12. 10.3399/bjgp11X588484.
18. Wang C, Li X, Wu R, et al. Evaluation of the corrective effect of hierarchical diagnosis and treatment on the inverted pyramid of healthcare-seeking order: a quasi-natural experiment based on the gradual pilots. Chin J Health Policy. 2021;14(3):13–20. [Article in Chinese].
19. Song H, Zuo X, Cui C, Meng K. The willingness of patients to make the first visit to primary care institutions and its influencing factors in Beijing medical alliances: a comparative study of Beijing's medical resource-rich and scarce regions. BMC Health Serv Res. 2019;19(1):361. 10.1186/s12913-019-4184-0.
20. Zhang Y, Wang Q, Jiang T, Wang J. Equity and efficiency of primary health care resource allocation in mainland China. Int J Equity Health. 2018;17(1):140. 10.1186/s12939-018-0851-8.
21. Wang X, Sun F, Du Y. Mechanism innovation and practical pathway of seamless county medical community construction in Sangzhi County. Chin Health Econ. 2025;44(12):28–31. [Article in Chinese].
22. Zhu F, He Q, Wang Z. Study on the impact of county medical community construction on healthcare service utilization: an analysis based on the compact county medical community pilot of one province. J Manage World. 2025;41(10):130–53. [Article in Chinese].

23. Peng M, Li L, Shi X, Wang Z. Does integrated health management within a county medical consortium improve rural type 2 diabetic patients' self-management behavior and quality of life? An empirical analysis from Eastern China. BMC Public Health. 2024;24(1):1439. 10.1186/s12889-024-18885-0.
24. Li M, Fu L, Mo A, Li L, Lin Y. Impact of the compact county medical service community policy on healthcare resource allocation and healthcare service utilization in village clinics: a difference-in-differences analysis in rural Sichuan, China. Front Public Health. 2026;14:1826414. 10.3389/fpubh.2026.1826414.
25. Ding S, Zhou Y. County medical community, medical insurance package payment, and hierarchical diagnosis and treatment—empirical analysis of the impact of the pilot project of compact county medical communities in Sichuan Province. PLoS ONE. 2024;19(4):e0297340. 10.1371/journal.pone.0297340.
26. Hunan Provincial Health Commission. Statistical communiqué on the development of health services in Hunan Province in 2024. Changsha; 2026. [Statistical report in Chinese].
27. Hunan Provincial Healthcare Security Administration. Statistical communiqué on the development of healthcare security in Hunan Province in 2025. Changsha; 2026. [Statistical report in Chinese].
28. Hunan Provincial Bureau of Statistics. Hunan Statistical Yearbook (2019–2025 editions). Changsha: China Statistics; 2019–25. [Statistical yearbook in Chinese].
29. General Office of the National Health Commission, General Office of the National Administration of Traditional Chinese Medicine. Notice on issuing the list of pilot provinces and pilot counties for the construction of compact county medical communities: Guo Wei Ban Ji Ceng Han 〔2019〕 No. 708. Beijing; 2019. [Policy document in Chinese].
30. Bertrand M, Duflo E, Mullainathan S. How much should we trust differences-in-differences estimates? Q J Econ. 2004;119(1):249–75. 10.1162/003355304772839588.
31. Callaway B, Sant'Anna PHC. Difference-in-differences with multiple time periods. J Econom. 2021;225(2):200–30. 10.1016/j.jeconom.2020.12.001.
32. Goodman-Bacon A. Difference-in-differences with variation in treatment timing. J Econom. 2021;225(2):254–77. 10.1016/j.jeconom.2021.03.014.
33. McWilliams JM, Hatfield LA, Landon BE, Hamed P, Chernew ME. Medicare spending after 3 years of the Medicare Shared Savings Program. N Engl J Med. 2018;379(12):1139–49. 10.1056/NEJMsa1803388.
34. Abrams R, Spiers J, Maben J, Grosvenor W, Touray M, Gage H. Implementing and sustaining dementia care coordinators across integrated care systems: a realist evaluation. BMC Med. 2024;22(1):584. 10.1186/s12916-024-03806-0.
35. Yip W, Fu H, Chen AT, Zhai T, Jian W, Xu R, et al. 10 years of health-care reform in China: progress and gaps in Universal Health Coverage. Lancet. 2019;394(10204):1192–204. 10.1016/S0140-6736(19)32136-1.

36. Yip W, Fu H, Jian W, Liu J, Pan J, Xu D, et al. Universal health coverage in China part 2: addressing challenges and recommendations. Lancet Public Health. 2023;8(12):e1035–42. 10.1016/S2468-2667(23)00255-4.
37. Jiang T. Mediating effects and moderating effects in empirical research on causal inference. China Industrial Econ. 2022;(5):100–20. [Article in Chinese].
38. Gao Q, Ma Y, Zhu P, Chen D. Healthcare professionals' views of the Integrated County Healthcare Consortium in Zhejiang, China. Int J Integr Care. 2022;22(2):25. 10.5334/ijic.5690.
39. Cui L, Yin G, Dai GLX, Tao H. Does global capitation prospective payment promote integrated delivery networks? Evidence from China's compact county medical communities. Int J Equity Health. 2025;24(1):164. 10.1186/s12939-025-02490-7.
40. Zhang S, Zhang W, Zhou H, Xu H, Qu Z, Guo M, et al. How China's new health reform influences village doctors' income structure: evidence from a qualitative study in six counties in China. Hum Resour Health. 2015;13:26. 10.1186/s12960-015-0019-1.

Figures

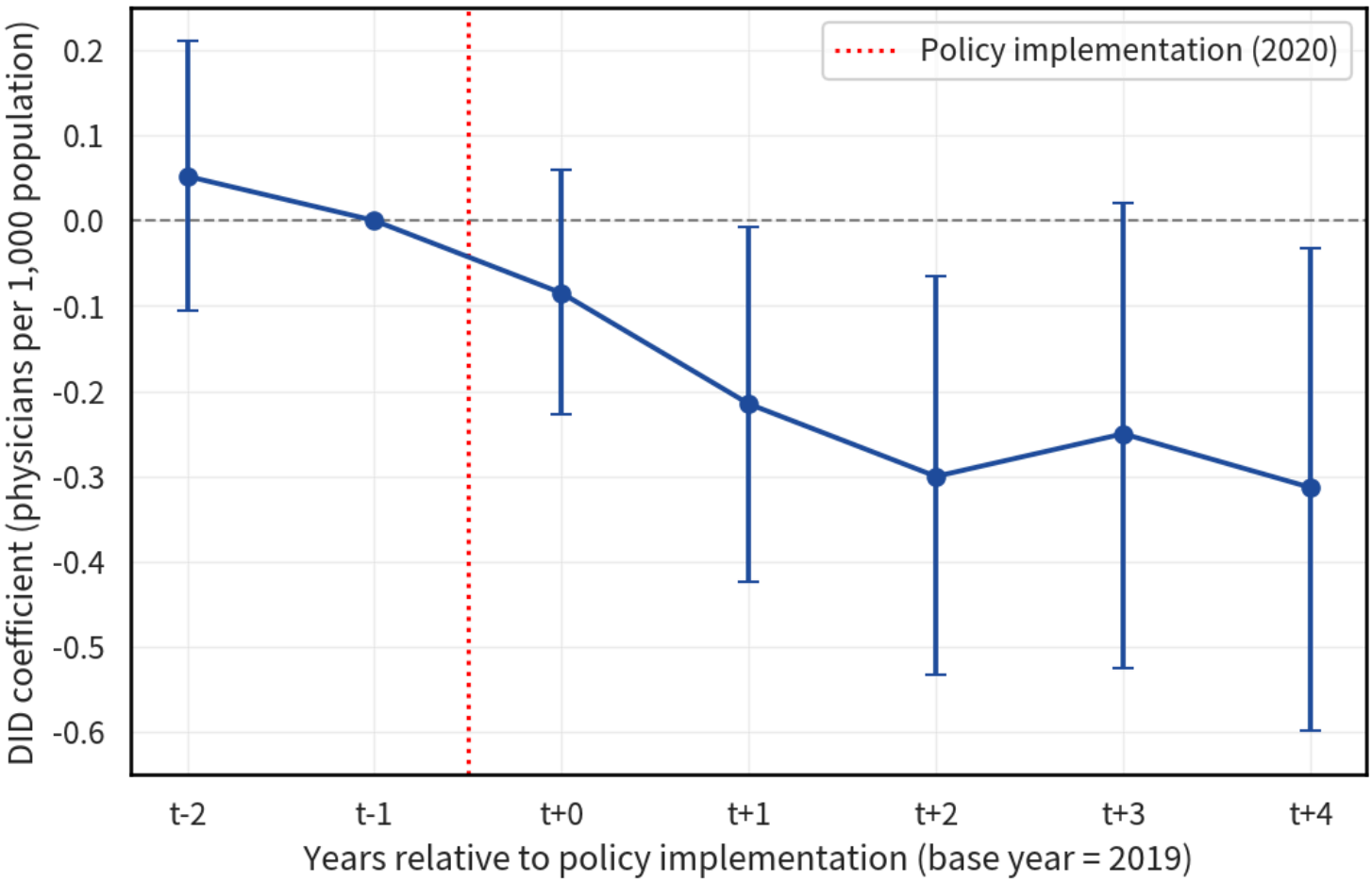


Figure 1

Event-study estimates of the CMC effect on licensed (assistant) physicians per 1,000 population, 2018–2024. The figure plots the coefficients β_k from equation (2) estimated on the balanced panel of 121 Hunan counties (19 treated, 102 controls; 847 county-year observations), with relative period k = −1 (the year 2019) as the omitted base period and a window spanning k = −2 to k = +4. Whiskers denote 95 percent confidence intervals based on county-clustered standard errors. The insignificant pre-treatment coefficient at k = −2 offers no evidence of differential pre-reform trends; post-treatment coefficients grow in absolute size with wide intervals.

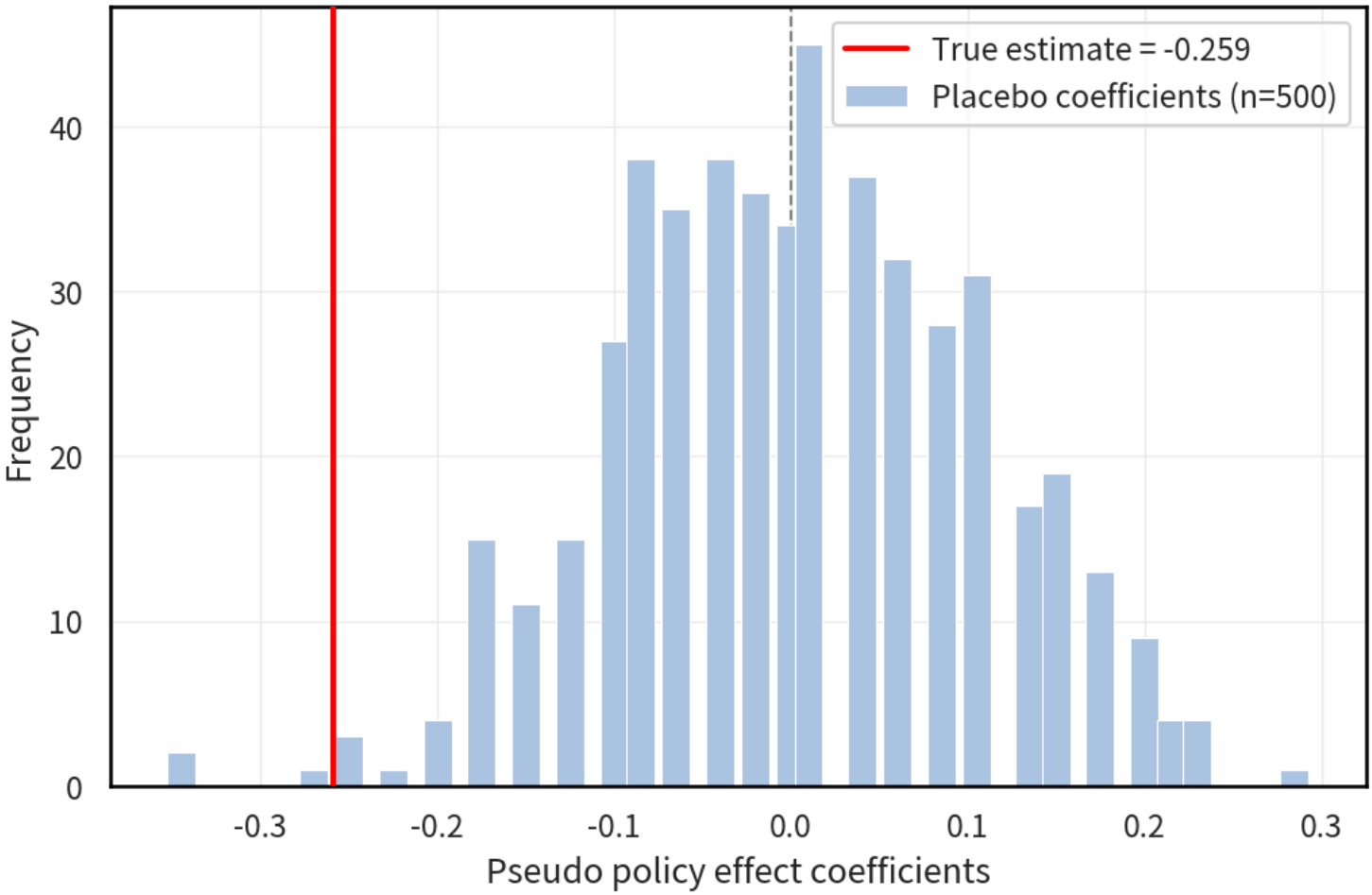


Figure 2

Placebo-test distribution of pseudo-policy effects (500 random permutations). Each permutation randomly drew 19 pseudo-treated counties from the 121-county sample while holding the true 2020 treatment timing fixed, and re-estimated the baseline DID model for licensed (assistant) physicians per 1,000 population. The 500 placebo coefficients have a mean of approximately −0.000 and a standard deviation of 0.104. The vertical line marks the true estimate (−0.259), which lies near the 0.4th percentile of the left tail; only 2 of 500 pseudo-coefficients were equally or more extreme in absolute value.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- fig42en.png
- fig41en.png