From Gray Income to Maternal and Infant Health Gains: The Impact of Anti-Corruption Inspections in Chinese Hospitals

Ruiqi Ding, Aljoscha Janssen, Jessica Ya Sun, Xuan Zhang*

April 24, 2026

Abstract

We investigate how anti-corruption enforcement affects physician behavior and patient welfare. Exploiting the staggered rollout of anti-corruption hospital inspections in China, we analyze all obstetrics and gynecology discharge records from a central province between 2015 and 2017. We find that inspections lead to an increase in aggregate healthcare spending, driven by an expansion in the quantity of admissions with no change in average costs per admission. Moreover, this expansion was accompanied by significant improvements in maternal and infant health outcomes. Mechanism analyses reveal that inspections corrected incentive distortions: providers substituted away from pharmaceutical rents toward service provision, reduced lengths of stay, and increased clinically-meaningful prenatal admissions. The enforcement also induced the exit of rent-seeking physicians. A welfare calculation suggests that the monetary value of the health gains substantially exceeds the marginal financial costs, indicating that anti-corruption oversight in healthcare can generate large net social benefits. (JEL I11, I18, D73, J13)

Keywords: Anti-corruption inspection; informal payments; maternal and infant health; prenatal care; physician incentives; physician turnover

*Ding: Singapore Management University (email: ruiqi.ding.2024@phdecons.smu.edu.sg); Janssen: Singapore Management University (email: ajanssen@smu.edu.sg); Sun: Huazhong University of Science and Technology (email: jessicasunya@hust.edu.cn); Zhang: Singapore Management University (email: xuanzhang@smu.edu.sg)

1 Introduction

Informal patient payments, ranging from kickbacks and other off-schedule transfers between providers and firms, remain pervasive in hospital care and are widely viewed as corrosive to equity and efficiency in low- and middle-income countries (Lewis, 2007; Chereches̨ et al., 2013). Global health commentary emphasizes that such corruption absorbs substantial resources and impedes progress toward universal coverage, implying that effective control measures could yield significant welfare gains (García, 2019). China presents a particularly salient case, where pharmaceutical kickbacks to physicians persist despite regulatory prohibitions (Xu & Yuan, 2022; Fu et al., 2023). Yet, despite clear theoretical grounds to expect that gray income distorts provider effort and patient access, especially for maternal and infant care, a domain where severe maternal morbidity continues to rise globally (Chatterji et al., 2026) and where time-sensitive and capacity-constrained service delivery makes provider incentives particularly consequential, there is little causal evidence on policies that remove providers' informal revenue streams within hospitals at scale, or on the downstream effects on utilization and clinical outcomes (Vian, 2020).

The welfare implications of such interventions are theoretically ambiguous: curtailing informal income may induce providers to substitute toward legitimate, billable services, raising official medical expenditure even as health outcomes improve, leaving the net social return unclear. Existing economic work has primarily evaluated formal payment reforms and audit programs in high-income settings and focused on billed (official) revenues but lacks evidence for the curtailment of informal earnings (Dafny, 2005; Clemens & Gottlieb, 2014; Fang et al., 2021; Gupta, 2021; Shi, 2024; Chen & Dillender, 2025). We fill this gap by analyzing a large-scale anti-corruption campaign targeting gray income in a developing economy and tracing its downstream consequences on hospital behavior, maternal and infant health, and social welfare.

In this paper, we evaluate a large anti-corruption inspection campaign designed to dismantle gray income channels and enforce compliance within Chinese hospitals. To identify the causal effect of these inspections on physician behavior and patient well-being, we use hospital admission data from the obstetrics and gynecology (OBGYN) department across all hospitals in a central province. The OBGYN data not only include healthcare utilization, but also contain health outcomes for both mothers and infants, allowing for a more complete cost-benefit analysis. We exploit the staggered rollout of inspections across hospitals to estimate event-study and difference-in-differences (DiD) models. Our identification strategy relies on the variation in the timing of inspections, which is uncorrelated with short-run OBGYN demand shocks. This setting allows us to compare changes in hospital outcomes (admissions, intensive vs. extensive margins of spending) and clinical outcomes (infant and maternal health) in inspected hospitals relative to those not yet inspected. Our study offers two distinct innovations: first, we treat the campaign, which targeted commercial bribery and off-schedule payments, as a shock to the informal incentive structure itself, rather than a change to official price schedules; second, we comprehensively trace the effects from provider supply responses to health outcomes.

We document that inspections led to a substantial increase in OBGYN outcomes. Following an inspection, total monthly OBGYN expenditure per hospital rose by approximately RMB 401,944 (an 18 percent increase over the pre-inspection mean of RMB 2.25 million).[1] Of this aggregate growth, prenatal (antepartum) care sees the largest proportional rise, with spending surging by 79% relative to the pre-inspection average. Furthermore, this expansion occurred almost entirely at the extensive margin: the number of prenatal admissions increased by approximately 20 patients per month (a 77 percent rise from a baseline of 26), while medical expenditure per admission remained unchanged. Beyond OBGYN outcomes, we find that this expansion in access translated into tangible improvements in health outcomes. For newborns, the rate of NICU admissions fell by 6.9 cases per 1,000 deliveries, while the incidence of low birth weight and very low birth weight declined by 7.4 and 2.1 cases per 1,000 deliveries, respectively. We also observed a modest increase in average birth weight (approximately 17.7 grams). For mothers, the inspections led to a reduction in delivery complications, including significant declines in surgical/anesthesia complications, cervical lacerations, severe maternal morbidity, and hypertensive disorder of pregnancy and eclampsia.

We investigate the mechanisms driving these health improvements and isolate two primary channels: behavioral substitution by incumbents and compositional shifts via selective attrition. First, the crackdown on illicit 'gray income' (e.g., pharmaceutical kickbacks) induced physicians to substitute toward legitimate, service-intensive provision like prenatal care to offset income shocks. This interpretation is supported by three findings: (i) treatment effects are concentrated in hospitals with high ex-ante rent-seeking norms (proxied by baseline C-section rates) and among senior physicians with greater discretionary power; (ii) hospital expenditures exhibit a structural break, characterized by a sharp decline in drug spending and a compensatory rise in service fees; and (iii) hospitals accommodated supply expansion through intensive-margin efficiency gains, significantly reducing average length of stay. Second, we find supporting evidence that hospitals improved physician quality, as hospitals with physician turnover due to the inspection experience greater increases in maternal and infant health outcomes. Additionally, we rule out a competing mechanism that the results are driven by strategic patient cherry picking: we do not find that treated hospitals increase admissions of low-risk pregnancies after the inspection, and increased prenatal hospitalizations reduce preventable maternal conditions. Together, these findings attribute welfare gains to corrected financial incentives and the exit of rent-seeking physicians.

Finally, to benchmark the welfare implications of the policy, we perform a conservative back-

[1]All magnitudes are derived from administrative data. Monetary values are expressed in constant 2015 RMB, deflated using the Consumer Price Index (CPI).

of-the-envelope cost-benefit analysis. We weigh the aggregate increase in medical expenditures against the social value of improved infant and maternal health outcomes. Our calculations yield a net social benefit of about RMB 22.15 million (USD 3.18 million),[2] driven by the short-term medical savings from averted very low birth weight cases and the long-term human capital returns associated with increased birth weight. These estimates represent a lower bound on the true welfare surplus, as they exclude the economic value of reductions in infant mortality and long-term disability.

Our findings contribute to three strands of literature. First, a large literature documents that physicians respond to financial incentives by distorting the quantity of care, such as over-providing high-reimbursement procedures when income falls (Gruber & Owings, 1996; Yip, 1998; Wang, 2023) or expanding service volume when payment rates rise (Grant, 2009; Clemens & Gottlieb, 2014). We contribute to this literature by documenting a distinct form of distortion: corruption induces physicians to shift the composition of care toward rent-generating activities, i.e., excessive drug prescribing pre-inspection, at the expense of other clinically appropriate services. Anti-corruption inspection eliminates this prescribing-based gray income, and we observe a reallocation of care toward prenatal hospitalization without any change in per-admission costs. This mechanism differs from standard supplier-induced demand (Ellis & McGuire, 1986; McGuire, 2000; Eliason et al., 2018) in that it operates through an extralegal rent channel rather than formal reimbursement rates, and is corrected by enforcement rather than payment reform.

Second, we show that the reallocation of care toward prenatal hospitalization improves maternal and infant health outcomes. A central challenge in health economics is distinguishing whether provider-driven volume expansions reflect wasteful over-treatment or the fulfillment of latent clinical need (Currie et al., 2014; Chandra & Staiger, 2020). The existing literature on prenatal care focuses almost entirely on outpatient visit timing and intensity (Corman et al., 2019), and causal identification has relied on instruments such as insurance expansions, conditional cash transfers, and clinic availability (Currie & Gruber, 1996; Cygan-Rehm & Karbownik, 2022; Kose et al., 2024). The role of inpatient monitoring during pregnancy as a distinct margin of prenatal care has received virtually no causal treatment. We exploit the staggered rollout of anti-corruption inspections as an exogenous shifter of prenatal hospitalization rates and find significant improvements in both maternal and infant health outcomes. Importantly, we rule out hospital cherry-picking—a standard concern in the literature on provider responses to regulatory change (Einav et al., 2018; Gupta, 2021; Savva et al., 2023; Shi, 2024)—as an alternative explanation, strengthening the interpretation that health gains reflect genuine improvements in care intensity rather than changes in patient composition.

Finally, we contribute to the political economy literature on state capacity and corruption control

[2]approximately equivalent to fully funding the annual public health expenditure for over 20,249 citizens in 2017

in public service delivery. While enhanced monitoring effectively reduces leakage (Olken, 2007), strict enforcement risks unintended consequences, such as defensive behavior or a retreat into bureaucratic inefficiency (Gerardino et al., 2024; Shi, 2024; Chen & Dillender, 2025). Our study identifies a setting where anti-corruption enforcement circumvents these distortions by actively realigning incentives. We demonstrate that curtailing gray income does not paralyze the healthcare bureaucracy; rather, it forces a reallocation of labor toward formal markets and measurable service improvements. This underscores that targeted state capacity can eliminate rent-seeking while simultaneously enhancing public service provision.

The paper proceeds as follows. Section 2 details the institutional background and the inspection campaign. Sections 3 and 4 describe the data and identification strategy, respectively. Section 5 presents the baseline estimates, while Section 6 explores the mechanisms driving these results. Section 7 offers a cost-benefit analysis, and Section 8 concludes.

2 Institutional Background

2.1 Informal Payments, Hospital Governance, and Physician Incentives in China

Since the 1980s, China's healthcare system has transitioned alongside the broader shift toward a market-oriented economy. As the government reduced direct fiscal subsidies, public hospitals were forced to become financially self-sustaining. To balance this fiscal retrenchment with affordable access, policymakers suppressed prices for routine medical services while allowing hospitals a 15 percent markup on pharmaceutical sales. This asymmetric pricing created a severe principal-agent problem: protected by information asymmetry, physicians faced strong incentives to induce demand via excessive prescriptions and redundant diagnostics. Consequently, while hospital budgets balanced, allocative efficiency and patient welfare declined.

The 2009 healthcare reform aimed to correct these distorted incentives, primarily through the Zero Mark-up Drug Policy (ZMDP) and National Essential Drugs List (NEDL) during 2010s. By abolishing the pharmaceutical markup, capping drug expenditure shares, and standardizing reimbursement lists, policymakers sought to sever the link between provider compensation and drug sales. While evidence suggests these measures lowered drug prices and weakened the "drug-financed" hospital model (Yip et al., 2010; Wu, 2019; Fang et al., 2021), underlying structural imbalances persisted. Because physician base salaries remained suppressed relative to their human capital and workload, the marginal utility of supplementary income stayed high. This "high-intensity, low-reward" equilibrium preserved the high marginal utility of supplementary income, incentivizing unintended behavioral responses (Xu & Yuan, 2022).

With the legal markup eliminated, pharmaceutical rents reorganized rather than disappeared.

Firms substituted institutional-level profit-sharing with direct, illicit financial relationships with individual physicians, increasingly using kickbacks to incentivize prescriptions (Yip et al., 2010). Rents previously captured formally by hospitals thus transformed into physician-level gray income, typically structured as volume-based commissions. Concurrently, on the demand side, patients facing severe information frictions and quality uncertainty continued to offer informal payments, commonly known as "red packets," to secure priority access or better care. These dual mechanisms further entrenched informal financial flows within clinical practice.

Ultimately, this experience illustrates the hydraulic nature of financial incentives in regulated markets: eliminating a legitimate but distortional revenue source fails to align incentives if the underlying participation constraint, i.e., adequate physician compensation, remains unmet. Instead of elimination, rents simply shifted from formal hospital revenues to illicit physician income. This reallocation pushed economic activity underground, increasing the agency costs of corruption and leaving the net impact on healthcare quality and social welfare ambiguous.

2.2 Anti-corruption and the Shift toward Active Hospital Governance

In November 2012, at the 18th National Congress of the Chinese Communist Party (CPC), the leadership framed official corruption as a fundamental threat to regime legitimacy. This shift prompted the December 2012 promulgation of the "Eight-Point Regulation" (*baxiang guiding*), launching an anti-corruption campaign of unprecedented scale. By targeting both high-ranking officials and grassroots cadres, the initiative aimed to disrupt the rent-seeking mechanisms entrenched during decades of decentralized economic growth. Crucially, enforcement expanded beyond traditional party organs to include state-owned enterprises, the military, and public service sectors, marking a systemic change in bureaucratic discipline.

Led by the Central Commission for Discipline Inspection (CCDI), the campaign utilized a top-down strategy involving repeated inspections, unannounced audits, and expanded investigative powers. Hundreds of thousands of officials have been disciplined since 2013. As shown on the left axis of Figure A1, the volume of bribery-related court judgments increased eightfold between 2011 and 2016. Unlike previous episodic crackdowns, this sustained enforcement has fundamentally altered political incentives and bureaucratic norms.

A prominent feature of this post-2012 regime is its focus on the healthcare sector. Publicly disclosed criminal verdicts reflect this shift: conditional on prosecution, the share of corruption cases originating from healthcare rose from 2 percent in 2011 to 10 percent in 2018 (Figure A1, right axis). Within this sector, the campaign targeted the gray income networks linking hospitals, physicians, and pharmaceutical firms. By expanding regulatory oversight across the entire medical delivery chain, from drug procurement to prescriptions, the state transitioned from de facto tolerance to strict prohibition. Criminalizing these customary informal pay practices definitively shifted the incentive structures for administrators and physicians. With severe legal liabilities now enforced, routine inspections carry the threat of career termination, substituting traditional supplementary income strategies with acute regulatory risk.

2.3 The Large Hospital Inspection Program

In 2005, the Chinese government launched the "Large Hospital Inspection Work Plan,"[3] (hereafter, Hospital Inspection), establishing the first nationwide systematic oversight of public hospitals. This initiative responded to growing concerns over hospital governance and healthcare efficiency during a period of rapid health system expansion.

The program has undergone five rounds. The Ministry of Health centrally administered the first two (2005 and 2010). The third round, launched in 2015, introduced a dual-level oversight structure: the National Health Commission (NHC) directly inspected a subset of hospitals,[4] while provincial commissions managed the rest. Led by senior officials, each inspection spanned 5 to 10 days on a rotating schedule. This 2015 round institutionalized these inspections as a permanent regulatory mechanism, a structure maintained in subsequent rounds (2019–2022 and 2023–2026).

Inspection focus and enforcement intensity evolved significantly. Early rounds prioritized basic administrative compliance and imposed weak sanctions.[5] In contrast, the 2015 round marked a sharp discontinuity. Aligning with the post-2012 national anti-corruption campaign, it explicitly targeted the gray income ecosystem, including pharmaceutical kickbacks and informal patient payments (*hongbao*). Enforcement became punitive: severe violations triggered public disclosure, administrative removal, medical license revocation, or criminal prosecution. Hospitals with systemic failures faced coordinated cross-agency penalties, such as targeted reimbursement audits. Consequently, the program began imposing binding constraints on hospital behavior.

Procedurally, the 2015 round standardized a three-stage protocol. First, hospitals conducted internal self-assessments and submitted corrective plans. Second, multidisciplinary inspection teams, including government officials, medical experts, and Party disciplinarians, performed on-site evaluations via document reviews, interviews, and random audits. Third, authorities publicly released the results and mandated verified rectification measures. This protocol emphasized three core areas: strengthening Party Committee oversight over hospital leadership, strictly enforcing

[3]Formally titled *daxing yiyuan xuncha gongzuo fang'an* jointly issued by the National Health and Family Planning Commission (NHFPC, the predecessor of the National Health Commission) and associated ministries.

[4]China's health administration was restructured when the Ministry of Health was reorganized into the National Health and Family Planning Commission in 2013 and later into the National Health Commission in 2018. For simplicity, we use "NHC" to refer to the MOH, NHFPC, and NHC throughout the paper.

[5]In 2005, noncompliant hospitals received only rectification recommendations. By 2010, repercussions remained largely confined to internal warnings, with public reprimands or dismissals remaining rare.

the "Nine Prohibitions" against kickbacks and supply-induced demand, and auditing budgetary and procurement compliance.

Our empirical analysis focuses on the 2015 Hospital Inspection's anti-corruption mandate within a central Chinese province. During this round, 44 large public hospitals experienced a staggered inspection rollout: 15 in November 2015, 15 in September 2016, and 14 in September 2017. This setting provides an ideal quasi-experiment. Occurring amid the broader anti-corruption campaign, it subjected a historically autonomous sector to intense political scrutiny and explicitly linked inspection outcomes to hospital resource allocation and leadership evaluations. Ultimately, it acts as a distinct regulatory shock, sharply increasing the expected costs of rent-seeking and raising the marginal returns to compliance with formal standards of care for administrators and physicians.

3 Data

3.1 Data Source

Our analysis utilizes a database constructed by linking two administrative datasets from the Health Commission of a central Chinese province. The merged sample covers the universe of public secondary and tertiary hospitals from 2014 to 2017.[6]

The first dataset, the Health Statistics Yearbooks, is a hospital-level panel capturing institutional operations and capacity. It records total revenue and its composition—disaggregated by care setting (outpatient versus inpatient) and source (pharmaceuticals, consumables, medical services)—alongside capacity metrics such as bed counts, staffing levels, and IT infrastructure. We use these variables as baseline controls for hospital-level heterogeneity.

The primary data source consists of admission-level discharge records from the OBGYN departments of 164 hospitals. These administrative microdata capture all provincial OBGYN inpatient admissions during the 2014–2017 period, yielding approximately 1.09 million observations. Each record contains hospital and physician identifiers, admission and discharge dates, patient demographics, and itemized expenditure breakdowns for pharmaceuticals, diagnostics, and consumables.

Crucially, these discharge records provide highly granular clinical information. Each observation includes up to nine diagnoses (coded via ICD-10) and seven procedures (coded via ICD-9-CM3), tracking the patient's diagnostic trajectory from initial assessment to final discharge. For delivery admissions, the data further report birth weight and specific newborn diagnoses. These clinical variables enable us to precisely categorize admission types and construct the maternal and infant health outcomes detailed in the subsequent section.

[6] China classifies hospitals into three tiers. We restrict our sample to Tier 2 (secondary) and Tier 3 (tertiary) hospitals; Tier 1 (primary) hospitals function mainly as community clinics with limited inpatient capacity and rarely perform the complex procedures relevant to our study.

3.2 Outcome Construction

We construct two primary sets of outcome variables to capture the impacts of inspections on hospital supply behavior and patient welfare: hospital outcome metrics and patient health outcomes.

Hospital outcomes. To quantify medical resource utilization, we aggregate administrative discharge records to the hospital-month level. Our primary outcome is total monthly inpatient expenditure within the OBGYN department. To elucidate the mechanisms driving aggregate expenditure dynamics, we decompose the provider supply response along two margins: the extensive margin (total monthly admission volume) and the intensive margin (average expenditure per admission). Given the clinical heterogeneity inherent to OBGYN services, we further disaggregate these measures by clinical function, partitioning admissions into pregnancy-related and non-pregnancy-related cases. Pregnancy-related admissions are further partitioned into four mutually exclusive categories: pregnancy loss, prenatal care, delivery, and postnatal care admissions.[7] This disaggregation enables us to assess whether supply responses are uniform across service types or concentrated in areas characterized by greater physician discretion.

Maternal and Infant Health Outcomes. To evaluate whether changes in service provision translate into welfare effects, we construct a vector of health indicators using admission-level delivery records. For infant health, we employ both continuous and binary metrics. These include birth weight (in grams), indicators for Low Birth Weight (LBW) and Very Low Birth Weight (VLBW), and an indicator for Neonatal Intensive Care Unit (NICU) admission.[8] Following Best et al. (2017), we also generate a distributional measure indicating whether birth weight falls within one standard deviation of the sample mean, capturing convergence toward optimal health. For maternal health, we identify adverse maternal outcomes based on discharge diagnoses as well. Key indicators include obstetric trauma such as perineal and cervical lacerations (Sentilhes et al., 2019), hypertensive disorders of pregnancy and eclampsia, severe maternal morbidity such as acute respiratory distress syndrome (ARDS) (Centers for Disease Control and Prevention, 2024), and

[7] We employ a hierarchical classification algorithm to define mutually exclusive admission types. First, we classify an admission as *pregnancy-related* if any diagnosis code falls within ICD-10 Chapter XV (codes beginning with 'O') or a specific subset of Chapter XXI (codes Z32–Z39); otherwise, it is classified as *non-pregnancy-related*. Second, within the pregnancy-related sample, we identify *delivery admissions* based on the presence of a recorded birth weight, which serves as a definitive marker for a birth event. Third, for the remaining non-delivery pregnancy admissions, we assign categories based on diagnosis codes: (i) *pregnancy loss* includes admissions with codes O00–O08 (ectopic pregnancy, hydatidiform mole, and other abortive outcomes) or explicit diagnoses of abortion, ectopic gestation, or stillbirth; (ii) *prenatal admissions* are identified by codes O10–O48 (maternal disorders related to pregnancy) or Z32–Z36 (antenatal screening and supervision); and (iii) *postnatal admissions* comprise the residual category, primarily capturing complications of the puerperium.

[8] We define *Low Birth Weight* (LBW) and *Very Low Birth Weight* (VLBW) based on standard clinical thresholds of <2,500g and <1,500g, respectively. Lacking a direct administrative flag for intensive care utilization, we construct a proxy for *NICU admission* using a composite algorithm. Following established protocols, this measure identifies NICU admissions based on a birth weight threshold of <2,000g (Braun et al., 2020; Pursley & Zupancic, 2020) as well as the presence of high-severity diagnostic codes for infants identified by Vance et al. (2023).

surgical/anesthesia complications.[9]

3.3 Sample Construction and Descriptive Statistics

To harmonize the data with the institutional context of the inspection campaign and ensure measurement reliability, we impose a series of sample restrictions and Table A1 shows the sample size and sample characteristics after each restriction is applied. First, we restrict the study window to the period from January 2015 to December 2017, excluding 2014 data due to substantial missing values in key variables. This initial restriction retains 929,837 admission records across the original 164 hospitals, with 47 percent of these admissions occurring in hospitals eventually subject to inspection. Second, we exclude small hospitals, defined as those with fewer than 1,000 OBGYN discharges over the three-year period or fewer than 100 total beds. This step mitigates measurement errors stemming from idiosyncratic volatility in low-volume hospitals and removes approximately 2 percent of observations. Third, to ensure the validity of our health outcome measures, we drop hospitals that failed to report birth weight data for more than half of the study months. We further remove hospital-month observations with fewer than 10 admissions to prevent outlier bias driven by the small-denominator problem. Finally, due to data availability constraints, we exclude hospitals inspected during the third round in September 2017, as they lack a sufficient post-treatment observation period.

Our final analytical sample consists of 686,588 admission records from 100 hospitals in a central Chinese province, spanning January 2015 to December 2017. Within this sample, 25 hospitals experienced inspection (accounting for 326,428 admissions), while the remaining 75 were never inspected (contributing 360,160 admissions). Although it seems natural to regard inspected hospitals as the treatment group and all non-inspected hospitals as a control group, to account for potential spillover effects—specifically, the deterrence effect of the inspection campaign on non-targeted hospitals within the same administrative jurisdiction, we divide non-inspected hospitals into a spillover control group and a clean control group. The spillover control group consists of non-inspected hospitals located in districts with at least one inspected hospital, whereas the clean control group comprises non-inspected hospitals in districts with no inspected hospital, and serves as our main control group in the analysis.

Table 1 reports summary statistics for each hospital group in the baseline period, i.e., before

[9] With the exception of *surgical and anesthesia complications*, which are directly observed in the raw administrative data, we construct all other adverse maternal outcome indicators using ICD-10 diagnosis codes recorded during the delivery admission. Specifically, we define binary indicators for: (i) *perineal lacerations* (second-degree and above, ICD-10 codes O70.1–O70.3); (ii) *cervical lacerations* (ICD-10 code O71.3); (iii) *hypertensive disorders*, encompassing gestational hypertension and eclampsia (ICD-10 codes O14–O16 and related subcategories); and (iv) *acute respiratory distress syndrome* (ICD-10 codes J80, J95.1, J95.2, J95.3, J95.82x, J96.0x, J96.2x, J96.9x, R06.03, R09.2), which is defined in accordance with CDC guidelines.

the first round of inspection. Columns 1 through 3 stratify the sample by the treatment, spillover control, and uncontaminated control groups, respectively, while column 4 presents the difference in means between treatment group and clean control group. In the treatment sample (column 1), which contains 250 hospital-month observations before Nov. 2015, the average Cesarean section (C-section) rate is around 44%, the mean birth weight is 3,262 grams, and the average length of stay is 7.67 days. Severe adverse events are rare: the LBW and VLBW rates are 7.16% and 1.39%, respectively, and no maternal deaths occurred during the baseline period. Regarding hospital operational outcomes, the average hospital admits 293 OBGYN patients per month, 69 percent of which are pregnancy-related. The mean inpatient expenditure per admission is 6,162 RMB, with pharmaceutical expenses accounting for roughly 22.6 percent of the total cost. Cross-column comparisons reveal that inspected hospitals are larger than their non-inspected counterparts, aligning with the policy's explicit targeting of major public hospitals. Furthermore, inspected hospitals exhibit higher C-section rates and (V)LBW rates, alongside lower mortality rates. This pattern is consistent with the established public health literature (Yin et al., 2023), reflecting that larger, higher-tier hospitals handle a disproportionate share of complex pregnancies but possess the superior medical capacity required to manage them effectively.

Figure A2 illustrates the unconditional trends in aggregated medical expenditures at the hospital-month level across different admission types for inspected and non-inspected hospitals. Panels A through D plot trajectories for all OBGYN admissions, non-pregnancy admissions, pregnancy admissions, and prenatal admissions, respectively.[10] In the full sample (Panel A), expenditures exhibit parallel pre-treatment trends ($t < 0$) followed by a persistent upward divergence for the treated group post-inspection ($t > 0$). Decomposing this aggregate effect demonstrates that the expenditure increase is mainly driven by pregnancy-related admissions (Panel C), particularly prenatal care (Panel D), while non-pregnancy admissions (Panel B) show no discernible change. Although these raw graphical patterns do not adjust for patient case mix or unobserved time-varying shocks, they offer compelling motivating evidence for the formal empirical analysis.

4 Empirical Strategy

4.1 Identification Strategy

To identify the causal effects of anti-corruption inspections, we employ a staggered DiD design. By exploiting the staggered rollout of province-level oversight as a quasi-experiment, this framework compares outcomes in treated hospitals against the control group before and after an inspection. The

[10]The horizontal axis denotes the relative time in months centered around the inspection date ($t = 0$). To ensure comparability, we randomly assign placebo inspection dates to the control group, drawn from the temporal distribution of actual inspection timings.

validity of this design hinges on the assumption that inspection timing is exogenous to unobserved determinants of hospital outcomes. In our context, inspection schedules were dictated by top-down administrative mandates under the broader political agendas, rather than endogenous responses to hospital-level financial distress or transient clinical fluctuations. This top-down assignment renders the precise timing of treatment orthogonal to high-frequency changes in OBGYN operations, making the parallel trends assumption plausible: absent the regulatory shock, treated and control hospitals would have evolved along parallel trajectories.

4.2 Hospital-Level and Patient-Level Specifications

To estimate the impact of inspections on institutional responses, we specify the following two-way fixed effects (TWFE) model at the hospital-month level:

$$Y_{ht} = \beta_0 + \beta_1 D_{ht} + \mu_h + \gamma_t + \varepsilon_{ht} \tag{1}$$

where Y_{ht} represents the outcome for hospital h in month t, including total OBGYN admissions, aggregated inpatient expenditures, and average expenditure per admission. The treatment variable D_{ht} is an indicator equal to 1 if hospital h has been inspected by month t, and 0 otherwise. The coefficient β_1 captures the average treatment effect on the treated (ATT). We include hospital fixed effects, μ_h, to absorb time-invariant institutional characteristics (e.g., administrative rank and baseline capacity), and year-month fixed effects, γ_t, to account for province-wide temporal shocks. Standard errors are clustered at the hospital level to adjust for serial correlation within hospitals.

To evaluate individual-level infant and maternal health outcomes, we estimate a corresponding specification as follows:

$$Y_{iht} = \beta_0 + \beta_1 D_{ht} + \mu_h + \gamma_t + \varepsilon_{iht} \tag{2}$$

where Y_{iht} denotes the health outcome for patient i admitted to hospital h in month t. Maternal outcomes include indicators for delivery trauma and severe obstetric complications, while infant outcomes capture birth weight and NICU admission. The treatment indicator D_{ht} varies at the hospital-month level. Fixed effects and standard error clustering remain identical to the institutional-level model.

4.3 Event Study Specification and Discussion

To scrutinize the dynamic effects of the inspection and assess the validity of the parallel trends assumption, we extend Equation (1) and Equation (2) into an event-study framework. Specifically, for the hospital-month level specification, the event-study specification can be written as:

$$Y_{ht} = \beta_0 + \sum_{k \neq -1} \beta_k \mathbf{1}\{t - T_h = k\} + \mu_h + \gamma_t + \varepsilon_{ht} \tag{3}$$

where T_h denotes the month in which hospital h is inspected, and $\mathbf{1}\{t - T_h = k\}$ is an indicator equal to one if month t is k periods away from the inspection event. The period prior to inspection ($k = -1$) is omitted and serves as the reference category. The coefficients β_k trace the evolution of outcomes in treated hospitals relative to control hospitals over event time. Under the identifying assumption of parallel trends, coefficients on the pre-inspection indicators ($k < 0$) should be indistinguishable from zero.

Recent econometric advances demonstrate that under staggered treatment timing, canonical TWFE models can yield severely biased estimates if treatment effects are heterogeneous across cohorts or over time (De Chaisemartin & d'Haultfoeuille, 2020; Goodman-Bacon, 2021). To overcome this negative weighting problem, we complement our canonical event study with the group-time average treatment effect estimator developed by Callaway and Sant'Anna (2021) and the heterogeneity-robust estimator proposed by Sun and Abraham (2021). By estimating cohort-specific dynamic effects before aggregating them, this approach ensures our dynamic estimates remain unbiased by underlying heterogeneous treatment dynamics.

5 Baseline Results

This section presents the baseline estimates of impacts of the inspection on hospital-level resource utilization and individual clinical outcomes. To prevent bias from treatment spillovers, our preferred specifications rely on strictly uncontaminated control groups. We first estimate the policy's effect on aggregate inpatient expenditures and then decompose this into extensive and intensive margin responses. Finally, we evaluate the welfare implications by examining maternal and infant health outcomes.

5.1 Hospital-level Outcomes

Table 2 Panel A reports the baseline estimates for aggregate monthly inpatient expenditure. Following an inspection, OBGYN expenditure in treated hospitals increased significantly by 401,944 RMB, representing an 18% increase relative to the pre-treatment mean (column 1). Decomposing this effect confirms that pregnancy-related admissions predominantly drive the aggregate increase (columns 2–3). Further stratification into four clinical subgroups reveals substantial heterogeneity (columns 4–7): while expenditures rose across prenatal, delivery, and postnatal care, the effect is disproportionately concentrated in prenatal admissions, which surged by 79%.

We next decompose this revenue growth into extensive (admission volume) and intensive (cost per admission) margins. Panel B presents the extensive margin responses. Total OBGYN admissions increased by 51 cases (15%; column 1). Consistent with the expenditure results, this expansion is heavily concentrated in discretionary services: prenatal and postnatal admissions increased by 77% and 91%, respectively (columns 5 and 7). Crucially, we observe no significant change in the volume of pregnancy loss admissions (column 4). Because interventions for pregnancy loss are dictated by acute biological necessity rather than provider discretion, this null result serves as a natural placebo test. It strongly suggests that the volume expansions in prenatal and postnatal care are driven by physician-induced demand.

Finally, Panel C evaluates the intensive margin. We find no significant changes in the average cost per admission for the full OBGYN sample (column 1), a null result that holds uniformly across all admission types (columns 4–7). Taken together, these estimates demonstrate that post-inspection revenue growth was driven entirely by the extensive margin, especially the expansion of discretionary admissions, without any corresponding intensification of per-admission medical spending.[11]

To test the parallel trends assumption and capture dynamic treatment effects, we estimate the event-study specification detailed in Equation (3). Figure 1 plots the coefficient estimates and 95% confidence intervals for aggregate expenditure and admission volume. Panels A and B report results for the full OBGYN department, while Panels C and D examine prenatal care. The event window spans from 6 months prior to 10 months post-inspection; observations outside this window are dropped, and $k = -1$ serves as the omitted reference period.[12] To address potential biases in TWFE models under staggered treatment timing, we compare canonical TWFE estimates with the heterogeneity-robust estimators proposed by Sun and Abraham (2021) and Callaway and Sant'Anna (2021).

The graphical evidence confirms the validity of our empirical design. Across all specifications, pre-treatment coefficients are statistically indistinguishable from zero, verifying the absence of differential pre-trends. Moreover, the point estimates from the TWFE, Sun and Abraham (2021), and Callaway and Sant'Anna (2021) models closely align, mitigating concerns regarding negative weighting biases in the canonical model.

After inspection, we observe a persistent structural increase in both expenditures and admission volumes, rather than a transitory spike. Notably, these effects take place with a three-to-four-month lag. This temporal delay directly maps onto the institutional timeline of the inspection process. Following the initial on-site audit, the inspection team issues formal feedback, prompting the

[11] In Appendix Table A2, following Card et al. (2009), we reveal that the increases in admissions are slightly more for deferrable admissions, especially among prenatal admissions.

[12] While the sample begins in January 2015—theoretically allowing a 10-month pre-period before the first inspection wave in November 2015—severe data missingness across several hospitals from January to April 2015 restricts our reliable pre-treatment window to 6 months.

hospital to submit and execute a detailed "rectification plan." This administrative cycle typically spans one quarter. Consequently, substantive operational adjustments, such as capacity expansions or shifts in admission strategies, are implemented around the third or fourth month, precisely coinciding with the mandatory follow-up re-inspections conducted by the authorities.

Our results may be confounded by two concurrent policies. First, the implementation of the Zero Mark-up Drug Policy (ZMDP) partially overlapped with our study period and is known to alter physician incentives, encouraging the substitution of prescribed drugs with medical services (Fang et al., 2021).[13] Table A3 demonstrates that after controlling for ZMDP rollout, the extensive margin responses remain robust. Second, the 2016 Two-Child Policy (TCP) might increase maternal healthcare utilization and affect who had a child.[14] To address these potential confounding factors, we conduct two robustness checks. First, we restrict the treatment group to hospitals inspected during the first round in 2015, prior to the TCP implementation. Figure A3 presents the event study results, demonstrating a persistent increase in both aggregated expenditures and admission volumes, where the occurrence of the effects is consistent with the re-inspection timing but not the TCP timing. Moreover, for the number of prenatal admissions, we start seeing an immediate increase in December 2015, which is before the announcement of TCP. Second, to account for the "legalization" of second pregnancies, we stratify the sample into two maternal age cohorts: pregnant women aged $\geq$29 years and those aged <29 years.The underlying assumption is that women in the <29 cohort are predominantly experiencing their first pregnancy and therefore less likely to be affected by the TCP.[15] The subgroup analysis (Figure A4, Panel A) indicates that while both age groups experienced an increase in prenatal care provision, with the effect moderately larger among the younger cohort, alleviating the concern of TCP.

In sum, the hospital-level analysis establishes that anti-corruption inspections significantly increased aggregate OBGYN revenues. This growth was driven entirely by the extensive margin: specifically, a sustained expansion in discretionary, pregnancy-related admissions such as prenatal care. By contrast, non-discretionary admissions, exemplified by pregnancy loss, remained unaffected.

These dynamics align closely with the income compensation hypothesis. By curtailing illicit gray income (e.g., pharmaceutical kickbacks) without concurrently adjusting formal wages, the inspections imposed an acute negative income shock on physicians. To offset these losses, providers expanded the volume of legitimate, high-discretion services. Moreover, this behavioral response

[13] In the studied province, the ZMDP was initially piloted in rural county-level hospitals in 2012 and fully implemented in these hospitals by July 2015. The policy was subsequently expanded to municipal and provincial hospitals beginning in December 2016, achieving a province-wide elimination of drug markups across all hospital tiers by July 2017.

[14] Announced nationwide in January 2016, China's Universal TCP formally ended the decades-long one-child policy by allowing all couples to legally have up to two children.

[15] According to Zhou et al., 2024, the average age at first childbirth in China from 2016 to 2020 was 28.8 ± 5.1 years.

operated exclusively along the extensive margin, contrasting with the intensive-margin adjustments typically observed following formal fee schedule reforms (Fang et al., 2021).

However, the normative implications of this extensive-margin expansion remain theoretically ambiguous. It could reflect welfare-reducing physician-induced demand (the over-provision of unnecessary care) or the beneficial fulfillment of latent, unmet medical needs. To identify which mechanism dominates and to evaluate the policy's net welfare impact, we next examine effects of inspections on maternal and infant health outcomes.

5.2 Individual-level Health Outcomes

To determine whether the extensive-margin expansion in obstetric services yields actual health benefits, we estimate the admission level specification (Equation (2)) using delivery admission microdata.

Table 3 Panel A reports the effects on infant health across four metrics: NICU admissions, the incidence of LBW and VLBW, birth weight, and an indicator for birth weight within one standard deviation of the sample mean (following Best et al., 2017). After inspection, treated hospitals experience a decline in NICU admissions of 6.92 per 1,000 births.[16] The incidence of LBW and VLBW similarly dropped by 7.36 and 2.09 per 1,000 births, respectively. Average birth weight increases significantly by 17.70 grams. In addition, the proportion of newborns within one standard deviation of the population mean increases by 13.81 per 1,000 births. This distributional shift confirms that the aggregate weight gain reflects a convergence toward a healthy median rather than an elevated risk of macrosomia.

Panel B evaluates maternal health using delivery diagnoses for obstetric trauma (perineal and cervical lacerations), surgical or anesthesia complications, ARDS, and hypertensive disorders of pregnancy (e.g., eclampsia). We find broad reductions in adverse maternal outcomes after inspections. Perineal and cervical lacerations decrease by 2.57 and 6.20 cases, respectively, while hypertensive disorders fall by 8.76 cases per 1,000 deliveries. Although the absolute reductions for severe, rare events, such as ARDS (0.14 cases) and surgical/anesthesia complications (0.19 cases), are small in magnitude, they represent substantial proportional declines relative to pre-treatment baselines. Overall, these estimates indicate a systematic improvement in obstetric safety and quality of care.

Figure 2 plots event-study estimates for key maternal and infant health outcomes: NICU admissions (Panel A), LBW (Panel B), cervical lacerations (Panel C), and hypertensive disorders (Panel

[16]Considering the marginal significance and the potential over-rejection issue arising from the relatively small number of treated clusters ($N = 25$), we re-estimated the p-values using the Wild Cluster Bootstrap-t procedure (Cameron et al., 2008), implementing Webb weights with 9,999 replications. The wild-bootstrapped p-values are reported in Table 3 and remain significant, confirming that our findings are not driven by the small number of treated clusters.

D). Across all specifications, the pre-treatment coefficients ($t \leq -2$) are statistically indistinguishable from zero, validating the parallel trends assumption. Following the inspection ($t \geq 0$), the point estimates exhibit a sustained downward trajectory. This persistent decline confirms that the regulatory shock does not disrupt clinical care; instead, it induces persistent improvement in infant and maternal outcomes.

A standard concern with high-stakes regulatory oversight is that administrative burdens may distract medical staff and inadvertently compromise patient safety. To rule out this risk, we estimate Equation (2) on extreme adverse clinical outcomes: maternal and fetal mortality, alongside 21 indicators of severe maternal morbidity (SMM). Figure A5 reports the treatment effects. Across all critical safety metrics, ranging from mortality to severe emergency interventions like hysterectomy and mechanical ventilation, the point estimates are tightly bounded around zero and insignificant. These precisely estimated null effects confirm that the inspections did not induce operational disruptions, defensive medicine, or compromises in foundational obstetric safety.

Considering the TCP, we conduct similar robustness checks for infant and maternal health. First, restricting our event study to the early 2015 inspection cohort yields results (Figure A6) that closely mirror our baseline findings (Figure 2). Moreover, the positive improvement in infant health emerged since March 2016, only three months after the announcement of TCP, where these infants should be conceived several months before the TCP. Second, Figure A4 Panel B demonstrates consistent health gains among mothers aged under 29, which rules out the possibility that the observed health gains are attributable to the "legalization" of prenatal care rather than to the inspections themselves.

Taken together, our findings demonstrate that the anti-corruption inspections yield substantial improvement in maternal and infant health outcomes without crowding out care for severe, life-threatening conditions. This resolves the normative ambiguity surrounding the post-inspection expenditure growth. Rather than reflecting wasteful supply-induced demand aimed merely at recouping lost gray income, the inspection-induced expansion in discretionary services, particularly prenatal care, appears to satisfy latent unmet medical needs and may generate net welfare gains, which we examine in the cost-benefit analysis in Section 7.

6 Mechanism Analysis

Although the preceding analysis demonstrates that inspections improved clinical outcomes, two critical issues remain. First, an incentive compatibility puzzle arises: if expanding prenatal care is both revenue-generating and health-enhancing, why was it underprovided prior to the regulatory shock? Second, the mechanisms driving these health gains are unidentified. Improved average outcomes do not necessarily reflect genuine enhancement in care quality. Instead, they may stem from strategic patient selection (i.e., cherry-picking) or defensive medicine practices, as physicians seek to minimize risk and increase leisure time. By systematically turning away high-risk patients, hospitals could artificially inflate aggregate health metrics. Such compositional shifts would represent a distortion of medical access rather than a net welfare gain.

To distinguish between these competing mechanisms and quantify the welfare consequences of inspections, this section examines the channels underlying the observed health improvements.

6.1 Reduced Length of Stay and Increased Prenatal Treatment

If the income compensation hypothesis holds, whereby providers expand legitimate care to offset lost gray income, treatment effects should scale with a provider's pre-intervention reliance on informal payments. We test this mechanism by exploiting institutional heterogeneity at both the hospital and physician levels.

First, we proxy a hospital's baseline reliance on financial incentives using its pre-inspection C-section rate. Because gray income is hard to track, C-section rates serve as a well-established proxy for supplier-induced demand; extensive literature documents that physicians manipulate C-section provision to maximize revenue under asymmetric information (Gruber & Owings, 1996; Johnson & Rehavi, 2016). We calculate baseline rates using pre-treatment data (January–June 2015) and hypothesize that hospitals with higher baseline C-section rates are more responsive to financial incentives and, consequently, suffered a more severe income shock from inspections.

We stratify the sample into high C-section (baseline rate > 40%) and low C-section (baseline rate < 40%) hospitals.[17]. Table 4 reports the split-sample estimates. Consistent with our hypothesis, health improvements are concentrated in high C-section hospitals (Panel A). In this cohort, post-inspection birth weight increases by 25.7 grams, and the incidence of LBW drops by 12.1 per 1,000 births (columns 1–2). Maternal outcomes exhibit parallel improvements, with cervical lacerations and hypertensive disorders falling by 4.72 and 14.65 per 1,000 deliveries, respectively. By contrast, the low C-section cohort experiences no significant improvements across any domain (Panel B).

Second, we exploit within-hospital heterogeneity by physician seniority. Because senior physicians possess greater clinical and administrative authority, they are the primary targets for gray income (e.g., pharmaceutical kickbacks or device-related bribes). Consequently, the health gains driven by the anti-corruption shock should be most pronounced among patients treated by these senior physicians. To examine this in our discharge records, we identify physician seniority via title-surname linkage: attending physicians are classified as "senior" if their surname matches that of a Chief Physician or Department Head involved in the same hospitalization, and "junior" otherwise.

[17] China has long exhibited high C-section utilization by international standards, with national estimates ranging from 35% to 40% in the mid-2010s, and hospital-level data revealing substantial variation extending well into the mid-40% range (Li et al., 2017; Zhang et al., 2022)

Table 4 (Panels C and D) reports the heterogeneous health effects stratified by attending physician seniority. Consistent with our hypothesis, health improvements are concentrated among patients treated by senior physicians (Panel C). Following an inspection, birth weight in this cohort increases by 18.57 grams, and the incidence of LBW falls by 9.01 per 1,000 births (columns 1–2). Similarly, maternal cervical lacerations and hypertensive disorders decline by 4.46 and 12.31 per 1,000 deliveries, respectively (columns 3–4). By contrast, the estimates for junior physicians are small and indistinguishable from zero across most of these metrics (Panel D), indicating null health improvement. These findings align with our conceptual framework: the health gains were driven precisely by physicians most entrenched in gray income before inspections.

To map these health improvements directly to provider behavioral changes, we next test whether the inspection-induced expansion in healthcare utilization also increases with baseline reliance on informal payments. Columns 5–7 aggregate utilization metrics at the hospital-month level. Panels A and B demonstrate that while both high and low C-section hospitals experience comparable baseline growth in total OBGYN expenditures, the expansion in discretionary prenatal care, measured by both expenditures and admission volumes, is substantially larger in the high C-section cohort. Panels C and D document a parallel dynamic for physician seniority: senior physicians drive a disproportionately larger expansion in service provision, a pattern particularly evident in targeted prenatal care.

The surge in prenatal care raises a capacity puzzle: How can OBGYN departments accommodate this service influx given their fixed ward infrastructure? Theoretically, hospitals can expand short-term capacity via phantom admissions (billing without physical occupancy), ad hoc physical expansion (e.g., corridor beds), or accelerated bed turnover. Lacking data on the first two margins, we test the third by examining changes in the ALOS. Column 8 of Table 4 shows that ALOS decreases significantly after the inspection, with the reduction concentrated entirely in high C-section hospitals (-0.31 days, $p < 0.01$). Stratifying by physician seniority (Panels C and D) reveals that this decline is driven by patients admitted under junior physicians (-0.47 days, $p < 0.1$), whereas the estimate for senior physicians is smaller and insignificant. This differential reduction likely reflects a mechanical shift in case mix: as service volume expanded, complex, time-intensive cases were disproportionately directed to senior physicians, mechanically shortening the ALOS for the less severe patient pool managed by junior physicians.

Overall, these findings illuminate the micro-mechanisms: to compensate for the loss of informal income, providers that are most reliant on illicit payments ex ante expanded the supply of discretionary prenatal care by accelerating bed turnover. This capacity expansion enabled earlier identification and management of pregnancy risks, thereby averting severe delivery complications and yielding the documented improvements in maternal and infant health.

6.2 Structural Changes in Expenditure: Drug-Service Substitution

Given that anti-corruption inspections primarily target gray income driven by pharmaceutical kickbacks, we next examine whether this regulatory shock altered the composition of medical care. Although our baseline estimates at the intensive margin show no significant change in per-admission expenditure, this aggregate null effect may obscure underlying compositional shifts, as providers still have incentives to replace targeted drugs with unregulated medical services.

To test this, we decompose inpatient expenditure into four mutually exclusive categories: drugs, services, diagnostics, and consumables.[18] From Table 5 Panel A, we document a clear substitution pattern: after the inspection, average drug expenditure per admission falls significantly by 227 RMB, offset by a concurrent 216 RMB increase in medical service expenditure. By contrast, expenditures on diagnostics and consumables remain unchanged. This nearly one-to-one financial offset explains the stable intensive margin observed in our baseline findings, and corroborates the income compensation hypothesis. As heightened regulatory scrutiny increased the expected cost of pharmaceutical kickbacks, physicians substituted away from drug-centric treatments toward legitimate, labor-intensive medical services to recoup lost gray income.

However, a decline in per-admission drug expenditure could reflect concurrent pricing reforms, such as Zero Mark-up Drug Policy (a price effect), rather than a genuine behavioral reduction in prescribing (a quantity effect). The patterns of drug-service substitution remain robust after considering the ZMDP, as shown in Table 5 Panel B. Furthermore, to isolate the quantity effect, we examine the effects of inspections on incidence of hospital admissions with near-zero drug utilization, which are reported in Table 5 Panel C. While the increase in admissions with exactly zero drug expenditure is insignificant, the number of admissions with drug spending below 10 RMB, 50 RMB, and 100 RMB exhibit a significant expansion of over 20 percent. This compositional shift toward low-drug-intensity care confirms that the expenditure decline is not merely an artifact of price reforms, but rather reflects a genuine structural decrease in prescribing quantity.

If providers compensate for lost pharmaceutical rents by expanding per-admission medical services, health improvement should scale with the magnitude of the inspection-induced decline in drug expenditure. To corroborate that the drug-to-service substitution drives the observed health gains, we investigate this as a dose-response relationship by stratifying hospitals by the post-inspection reduction in drug expenditure. For robustness, we employ two classification criteria: above- versus below-median reductions in absolute drug expenditure per admission (Table 6, panels A and B), and above- versus below-median reductions in the drug-to-total revenue ratio, which accounts for baseline hospital scale differences (panels C and D).

Columns 1–4 report the heterogeneous treatment effects on individual-level health outcomes,

[18] Service include nursing, surgical and anesthesia fees; diagnostics consist of pathology, laboratory, and imaging fees; and consumables include single-use medical supplies for examinations, treatments, and surgeries.

from which we find that health improvements are concentrated in hospitals experiencing a sharper reduction in drug revenues. In the above-median absolute reduction cohort (Panel A), birth weight increases by 21.01 grams, and LBW incidence decreases by 8.87 per 1,000 births. By contrast, the below-median cohort (Panel B) exhibited no improvement in these neonatal metrics. For maternal health, although absolute reductions in cervical lacerations are comparable between the two groups (-5.91 vs. -5.06 per 1,000 deliveries), the proportional decline relative to the pre-treatment baseline is larger in the above-median group (26.4%) than in the below-median group (4.0%). This dose-response pattern holds firmly when stratifying hospitals by the drug-to-total revenue ratio. Hospitals with above-median ratio reductions (Panel C) achieve a 19-gram increase in birth weight and an 8.52 per 1,000 decrease in LBW incidence, whereas the below-median cohort (Panel D) experiences precisely estimated null effects.

We further examine whether the hospitals facing greater drug revenue losses also exhibit stronger compensatory expansions in healthcare service provision. Columns 5–7 of Table 6 report hospital-month utilization outcomes stratified by the magnitude of drug expenditure reductions. Hospitals experiencing above-median absolute reductions in per-admission drug expenditure (Panel A) expand total monthly OBGYN expenditures by 425,493 RMB while hospitals with below-median reductions (Panel B) exhibit an insignificant change in total OBGYN expenditures. Discretionary care follows a similar gradient that more prenatal care is provided in hospitals with more reductions in drug utilization. Notably, the efficiency gains in bed turnover (ALOS reduction) occurs uniformly across all subgroups, ranging from -0.18 to -0.22 days (column 8). This uniform decline suggests that while bed turnover improves universally after inspection, the expansion of prenatal services is strictly contingent on the severity of the revenue shock.

Taken together, these results provide further support for the income compensation mechanism. The intensity of the informal income shock, which is proxied by the contraction in pharmaceutical revenues, determines the magnitude of a hospital's shift toward legitimate medical service provision, i.e., prenatal care in our context. This structural reallocation of medical effort, in turn, drives the documented improvements in maternal and infant health.

6.3 Physician Turnover and Improved Hospital Performance

As outlined in the institutional background, the 2015 inspection introduced stringent penalties for rent-seeking, including license revocation and criminal prosecution. This regulatory severity motivates a "selective attrition hypothesis": physicians heavily dependent on informal income may have strategically exited inspected hospitals to evade scrutiny. By relocating to uninspected hospitals (e.g., private or lower-tier hospitals) or opting for early retirement, these high-risk providers would mechanically shift patients to new hires or remaining staff. If the exiting physicians were indeed rent-seeking agents and primary drivers of supply-induced demand, their departure and the subsequent reallocation of care to less financially incentivized peers could independently drive the observed improvements in maternal and infant health.

To test this hypothesis, we first examine how inspections affect personnel composition by tracking attending physician presence using surname data from discharge records. We classify physicians into three mutually exclusive cohorts: (1) Exit, appearing pre-inspection but permanently vanishing thereafter; (2) Entry, appearing exclusively post-inspection; and (3) Stay, observed consistently throughout the sample period. Aggregating these categories to the hospital-month level allows us to measure the physician turnover (i.e., monthly exit and entry volumes). Panel A of Table 7 reports the impact on personnel turnover. Using a six-month window around the inspection (columns 1–2), we observe a significant increase in physician mobility: monthly exits and entries per hospital increase by 0.15 and 0.12, respectively. Narrowing the bandwidth to three months amplifies these estimates to 0.19 for exits and 0.20 for entries (columns 3–4), confirming an immediate and structural restructuring of human capital.

To determine whether this compositional shift contributes to the documented health gains, we exploit variation in physician attrition. If the departure of gray-income-reliant providers is the primary mechanism, treatment effects should be concentrated in hospitals that experienced personnel turnover. We test this by stratifying the sample based on whether a hospital recorded at least one physician exit within the six-month window—a strict temporal restriction designed to isolate inspection-induced departures from normal personnel changes. Intuitively, the presence of exiting physicians proxies for the intensity of the regulatory shock and the resulting ”purification” of the provider pool.

Panels B and C of Table 7 present the heterogeneity in health outcomes. Consistent with the attrition hypothesis, patient health improvements are exclusively driven by hospitals that experienced physician exits (Panel B). In this cohort, birth weight increases significantly by 20.59 grams, and LBW incidence falls by 10.52 per 1,000 births. Maternal outcomes improve commensurately, with cervical lacerations and hypertensive disorders declining by 7.08 and 11.26 per 1,000 deliveries, respectively (representing 16.0% and 16.8% reductions relative to baseline means). By contrast, hospitals with no observed exits during the event window (Panel C) exhibit small and insignificant changes across all health metrics.

Overall, the empirical evidence is consistent with the reallocation of human capital serving as a mechanism linking regulatory enforcement to improved maternal and infant health. The anti-corruption shock operated not only through the behavioral adaptation of incumbent staff, but also via a structural shift in workforce composition. By inducing the exit of providers most reliant on illicit rents, the policy reallocated patient volume toward less financially incentivized and more compliant peers, thereby fostering more rational healthcare provision.

6.4 Ruling Out Patient Cherry-Picking

Before attributing the observed health gains to expanded prenatal care and selective provider attrition, we must address a critical threat to internal validity: hospitals' cherry-picking of healthier patients. If physicians respond to inspections by preferentially admitting low-risk patients and diverting complex cases, the observed improvements would just reflect a shift in patient composition rather than genuine improvements in care delivery. Under such risk-selection, the apparent gains in health outcomes would mask a deterioration in aggregate welfare by deterring high-risk patients from accessing necessary tertiary care.

Providers operating under heightened scrutiny indeed have incentives to engage in such risk-selection. On the one hand, avoiding high-risk patients minimizes the probability of adverse clinical outcomes and subsequent malpractice disputes during a politically sensitive enforcement period. On the other hand, selecting healthier patients reduces clinical effort, allowing physicians to substitute leisure for their lost informal income. Hence, to test for this compositional shift, we estimate Equation (2) at the individual level to examine whether the ex-ante clinical risk profile of patients changed in treated hospitals.

Following Leonard et al. (2020), we utilize ICD-10 diagnostic codes to identify 27 specific obstetric comorbidities at the time of admission to measure risk profile.[19] Figure 3 solid circles plot the estimated treatment effects on this comorbidity composition among delivery admissions, which yield no evidence of cherry-picking. The changes in admission probabilities for nearly all severe obstetric comorbidities are statistically insignificant, indicating that treated hospitals did not selectively deter or transfer complex delivery cases. By contrast, the probability of admitting mothers of advanced maternal age increased after the inspection. Given that advanced maternal age is a well-documented risk factor for adverse perinatal outcomes, this positive coefficient indicates that, if anything, treated hospitals absorbed a higher-risk demographic.

Moreover, an alternative interpretation of increased prenatal care is that physicians engaged in revenue-driven over-provision by admitting low-risk patients during the prenatal stage. To rule out this compositional effect, we analyze the prevalence of obstetric comorbidities among prenatal admissions as well. Figure 3 reveals that treated hospitals experienced an increase in prenatal admissions for severe medical conditions. This confirms that the observed expansion in preventive care reflects genuinely unmet clinical needs rather than unnecessary financial-driven hospitalizations.

Overall, Figure 3 offers compelling evidence that the inspection-induced expansion of prenatal care improves health outcomes. We observe a distinct divergence in coefficients for treatable conditions between the prenatal and delivery samples. For example, the prevalence of anemia

[19] The identification is based on the 27 Comorbidity Diagnosis Groups outlined in Appendix 2 of Leonard et al. (2020). See https://cdn-links.lww.com/permalink/aog/b/aog_136_3_2020_07_02_leonard_20-592_sdc1.pdf for details.

exhibits an increase in prenatal admissions but a decrease at delivery, indicating that earlier detection and management effectively reduce its prevalence at birth. Conversely, non-modifiable conditions, such as advanced maternal age and previous C-section births, remain similar, reflecting a baseline demographic shift.

Taken together, these findings rule out the confounding mechanism that inspected hospitals cherry-pick healthier patients to improve performance. On the contrary, the documented health gains occur despite a stable, or even marginally riskier, patient case mix, as expanded preventive care successfully keeps severe conditions under control.

7 Cost-Benefit Analysis

While anti-corruption mandates primarily target illicit rent-seeking, their broader welfare implications dictate long-term policy viability. In this section, we outline a stylized cost-benefit framework. By monetizing the estimated treatment effects on patient outcomes, we evaluate whether the realized health gains justify the policy-induced shifts in medical expenditures.

7.1 Policy Cost and Benefit

We define the policy cost as the aggregate increase in medical expenditures across the treated hospitals. The sample comprises 25 treated hospitals observed over 515 post-treatment hospital-months (14 first-round hospitals with 25 months of exposure, and 11 second-round hospitals with 15 months).[20] Applying our baseline estimate of a 401,944 RMB increase in aggregate monthly expenditure per hospital, we calculate a total policy cost of approximately 207 million RMB.[21] This figure represents the comprehensive financial burden of tighter regulatory compliance and the induced structural shift toward more intensive clinical practice.

To quantify the corresponding monetary benefits, we evaluate two primary channels: infant health (incorporating short-term medical savings and long-term human capital accumulation) and maternal health (capturing avoided hospitalization costs and preserved labor productivity).

We first quantify the policy benefit based on infant health gains. Our baseline estimates indicate a 0.21 percentage point decline in the probability of VLBW. Applied to the 101,019 deliveries in treated hospitals post-inspection,[22] the hospital inspection decreased VLBW by 213 cases, which is equivalent to an estimated short-term medical savings between 31.95 million and 106.50 million

[20] Data for one first-round and four second-round hospitals are unavailable in the raw dataset.

[21] We exclude the administrative costs of the inspections, as these are absorbed within the existing operational budgets of the Disciplinary Inspection Team stationed at the Health Commission. Our cost definition therefore strictly captures the direct financial impacts on the healthcare system.

[22] Within the analyzed sample, treated hospitals accounted for a total of 163,520 deliveries, of which 62,501 occurred prior to the inspection and 101,019 occurred afterward.

RMB.[23] Beyond immediate medical savings, improved infant health generates long-term human capital returns (Black et al., 2007; Almond et al., 2018; Bharadwaj et al., 2018). Prior literature establishes that a one-standard-deviation increase in birth weight raises adult annual income by roughly 2.75% (Lambiris et al., 2022). To monetize our estimated 18-gram increase in mean birth weight, we project the present value (PV) of lifetime income gains using a standard formulation:

$$PV = [p \times w_0] \times \sum_{t=1}^{T} \frac{(1+g)^{t-1}}{(1+r)^t} \tag{4}$$

where we adopt a relatively conservative estimation strategy: $r = 3\%$ is the discount rate, $g = 1\%$ is the real wage growth rate, $T = 40$ denotes working years, w_0 is the baseline annual wage, and p scales the income premium by our estimated birth weight gain relative to its sample standard deviation.[24] Applying the non-private sector wage benchmark yields a per-child lifetime income gain of 1,937 RMB,[25] aggregating to a social benefit of 195.67 million RMB. A more conservative private-sector benchmark implies a per-child gain of 1,193 RMB and an aggregate benefit of 120.52 million RMB.

The second benefit channel stems from reductions in adverse maternal outcomes. We use the decline in cervical lacerations, a complication known to prolong hospital stays and delay postpartum recovery (Liu et al., 2015), as a representative metric. The estimated 0.62 percentage point reduction translates to 626 averted cases post-inspection. Assuming this avoids one to two days of hospitalization, aggregate medical savings range from 0.59 million to 1.19 million RMB.[26] Furthermore, if avoiding this trauma accelerates the return to work by five days, applying an average daily wage of 297 RMB yields an additional 0.93 million RMB in preserved maternal productivity.

7.2 Net Social Benefits and Discussion

Finally, we make use of the midpoints of the previously derived ranges to calculate the net social benefits. The estimated short-term infant medical savings average 69.23 million RMB, while the long-term human capital benefits average 158.1 million RMB. Similarly, after considering averted medical costs and preserved productivity, maternal health gains add approximately 1.82 million RMB. Aggregating these components yields a total monetized social benefit of 229.15 million RMB. Subtracting the estimated policy cost of 207 million RMB results in a net social surplus of

[23] Drawing on evidence from Chinese NICU, the medical cost of treating a VLBW infant ranges from 150,000 to 500,000 RMB (Zhu et al., 2020).

[24] $p = (\Delta_{BW}/\sigma_{BW}) \times 0.0275$, where Δ_{BW} indicates the increase in birth weight, and σ_{BW} indicates the sample standard deviation, which is equal to 516.

[25] Baseline annual wages (w_0) in 2017 were 74,318 RMB in the urban non-private sector and 45,761 RMB in the private sector (Ministry of Human Resources and Social Security of the People's Republic of China, 2018).

[26] In our sample, the average daily cost of a delivery-related stay is approximately 950 RMB.

22.15 million RMB (approximately 3.18 million USD), which is equivalent to fully funding the annual public health expenditure for over 20,249 citizens in 2017.[27]

It should be noted that this stylized accounting framework yields a strict lower bound of the true welfare gains. Unlike comprehensive assessments (Baran et al., 2025), our calculations omit the economic value of averting infant mortality, preventing lifelong childhood disabilities, and mitigating severe maternal morbidity. In high-income settings, mortality risk reduction alone often exceeds 1.3 million USD per VLBW case, constituting the vast majority of social benefits. We exclude these dimensions due to the scarcity of context-specific parameters for the Value of a Statistical Life and Quality-Adjusted Life Years in our setting. Nevertheless, incorporating these unobserved dimensions would substantially scale the estimated benefits, reinforcing our conclusion that the healthcare reforms induced by the anti-corruption mandates are highly cost-effective.

8 Conclusion

This paper provides novel empirical evidence on the healthcare utilization and welfare consequences of anti-corruption inspections within China's public hospital sector. Exploiting the staggered rollout of the 2015–2016 anti-corruption hospital inspections, we document a robust causal link between tighter regulatory oversight on gray income, the expansion of legitimate preventive care, and significant improvement in maternal and infant health outcomes.

Our analysis yields several insights into the mechanics of healthcare provision under regulatory scrutiny. First, the inspections significantly increased aggregate inpatient expenditures, driven entirely by the extensive margin (admission volume) rather than the intensive margin (per-admission spending). Among all types of admission, prenatal admissions increase the most. Moreover, this expansion in healthcare access translates into health gains, including significant increases in birth weight, reductions in (very) low birth weight incidence, a distributional shift of birth weights toward the clinical optimum, and fewer maternal delivery complications.

Heterogeneity and mechanism analyses reveal that these gains stem from a realignment of provider incentives. Health improvements are concentrated among hospitals and senior physicians with greater ex-ante reliance on informal income, suggesting that the intervention effectively targeted margins where agency problems were most acute. As inspections curtailed pharmaceutical kickbacks, providers substituted away from drug-centric revenue models toward legitimate, labor-intensive medical services, particularly prenatal care. This supply-side response was facilitated by improved bed turnover and reinforced by a compositional shift, i.e., the selective attrition of highly financially driven providers. Finally, our cost-benefit analysis demonstrates that this enforcement mandate generated positive net social surplus, as the long-term human capital and health gains outweigh the fiscal costs of increased medical expenditure.

[27] China's per capita Government Health Expenditure (GHE) was 1,093.88 RMB in 2017 (Yip et al., 2019).

These findings offer two salient policy implications for healthcare governance, especially for developing economies facing structural corruption and scarcity of healthcare resources. First, our results support the transition from episodic enforcement to institutionalized regulatory inspections. We find no evidence that sudden regulatory scrutiny compromised infant health or led to strategic risk-selection (cherry-picking). Instead, it improved hospital performance and generated net social gains. Establishing a persistent threat of detection can durably alter the risk-return calculus of rent-seeking, compelling healthcare providers to compete on clinical efficiency rather than pharmaceutical arbitrage.

Second, our mechanism analysis highlights the necessity of formal compensation reforms. The aggressive expansion of legitimate service supply following the loss of gray income underscores physician responsiveness to financial incentives. The pre-inspection prevalence of corruption was largely symptomatic of misaligned pecuniary incentives, where formal wages were suppressed below the market-clearing level. Therefore, sustainable anti-corruption efforts must decouple physician income from drug and device sales while simultaneously raising base salaries and medical service fees to reflect the true value of clinical human capital.

References

Almond, D., Currie, J., & Duque, V. (2018). Childhood circumstances and adult outcomes: Act ii. *Journal of Economic Literature*, *56*(4), 1360–1446.

Baran, C., Currie, J., Dursun, B., & Tekin, E. (2025). *Clean rides, healthy lives: The impact of electric vehicle adoption on air quality and infant health* (tech. rep.). National Bureau of Economic Research.

Best, K. E., Tennant, P. W., & Rankin, J. (2017). Survival, by birth weight and gestational age, in individuals with congenital heart disease: A population-based study. *Journal of the American Heart Association*, *6*(7), e005213.

Bharadwaj, P., Lundborg, P., & Rooth, D.-O. (2018). Birth weight in the long run. *Journal of Human Resources*, *53*(1), 189–231.

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2007). From the cradle to the labor market? the effect of birth weight on adult outcomes. *The Quarterly Journal of Economics*, *122*(1), 409–439.

Braun, D., Braun, E., Chiu, V., Burgos, A. E., Gupta, M., Volodarskiy, M., & Getahun, D. (2020). Trends in neonatal intensive care unit utilization in a large integrated health care system. *JAMA network open*, *3*(6), e205239.

Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of econometrics*, *225*(2), 200–230.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *The review of economics and statistics*, *90*(3), 414–427.

Card, D., Dobkin, C., & Maestas, N. (2009). Does medicare save lives? *The quarterly journal of economics*, *124*(2), 597–636.

Centers for Disease Control and Prevention. (2024). Identifying severe maternal morbidity (smm): Icd codes [https://www.cdc.gov/maternal-infant-health/php/severe-maternal-morbidity/icd.html (accessed March 8, 2026)].

Chandra, A., & Staiger, D. O. (2020). Identifying sources of inefficiency in healthcare. *The quarterly journal of economics*, *135*(2), 785–843.

Chatterji, P., Li, F., & Markowitz, S. (2026, April). *Exploring trends in severe maternal morbidity in the u.s.* (Working Paper No. 35062). National Bureau of Economic Research. https://doi.org/10.3386/w35062

Chen, Y., & Dillender, M. (2025). *Government monitoring of health care quality: Evidence from the nursing home sector* (tech. rep.). National Bureau of Economic Research.

Chereches, R. M., Ungureanu, M. I., Sandu, P., & Rus, I. A. (2013). Defining informal payments in healthcare: A systematic review. *Health Policy*, *110*(2–3), 105–114. https://doi.org/10.1016/j.healthpol.2013.01.010

Clemens, J., & Gottlieb, J. D. (2014). Do physicians' financial incentives affect medical treatment and patient health? *American Economic Review*, *104*(4), 1320–1349.

Corman, H., Dave, D., & Reichman, N. E. (2019). The effects of prenatal care on birth outcomes: Reconciling a messy literature. In *Oxford research encyclopedia of economics and finance*.

Currie, J., & Gruber, J. (1996). Saving babies: The efficacy and cost of recent changes in the medicaid eligibility of pregnant women. *Journal of political Economy*, *104*(6), 1263–1296.

Currie, J., Lin, W., & Meng, J. (2014). Addressing antibiotic abuse in china: An experimental audit study. *Journal of development economics*, *110*, 39–51.

Cygan-Rehm, K., & Karbownik, K. (2022). The effects of incentivizing early prenatal care on infant health. *Journal of Health Economics*, *83*, 102612.

Dafny, L. S. (2005). How do hospitals respond to price changes? *American Economic Review*, *95*(5), 1525–1547.

De Chaisemartin, C., & d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American economic review*, *110*(9), 2964–2996.

Einav, L., Finkelstein, A., & Mahoney, N. (2018). Provider incentives and healthcare costs: Evidence from long-term care hospitals. *Econometrica*, *86*(6), 2161–2219.

Eliason, P. J., Grieco, P. L., McDevitt, R. C., & Roberts, J. W. (2018). Strategic patient discharge: The case of long-term care hospitals. *American Economic Review*, *108*(11), 3232–3265.

Ellis, R. P., & McGuire, T. G. (1986). Provider behavior under prospective reimbursement: Cost sharing and supply. *Journal of health economics*, *5*(2), 129–151.

Fang, H., Lei, X., Shi, J., & Yi, X. (2021). Physician-induced demand: Evidence from china's drug price zero-markup policy. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3880231

Fu, H., Lai, Y., Li, Y., Zhu, Y., & Yip, W. (2023). Understanding medical corruption in china: A mixed-methods study. *Health Policy and Planning*, *38*(4), 496–508. https://doi.org/10.1093/heapol/czad015

García, P. J. (2019). Corruption in global health: The open secret. *The Lancet*, *394*(10214), 2119–2124.

Gerardino, M. P., Litschig, S., & Pomeranz, D. (2024). Distortion by audit: Evidence from public procurement. *American Economic Journal: Applied Economics*, *16*(4), 71–108.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of econometrics*, *225*(2), 254–277.

Grant, D. (2009). Physician financial incentives and cesarean delivery: New conclusions from the healthcare cost and utilization project. *Journal of health economics*, *28*(1), 244–250.

Gruber, J., & Owings, M. (1996). Physician financial incentives and cesarean section delivery. *RAND Journal of Economics*, *27*(1), 99–123.

Gupta, A. (2021). Impacts of performance pay for hospitals: The readmissions reduction program. *American Economic Review*, *111*(4), 1241–1283.

Johnson, E. M., & Rehavi, M. M. (2016). Physicians treating physicians: Information and incentives in childbirth. *American Economic Journal: Economic Policy*, *8*(1), 115–141.

Kose, E., O'Keefe, S., & Rosales-Rueda, M. (2024). Does the delivery of primary health care improve birth outcomes?: Evidence from the rollout of community health centers. *Journal of Human Resources*.

Lambiris, M. J., Blakstad, M. M., Perumal, N., Danaei, G., Bliznashka, L., Fink, G., & Sudfeld, C. R. (2022). Birth weight and adult earnings: A systematic review and meta-analysis. *Journal of Developmental Origins of Health and Disease*, *13*(3), 284–291.

Leonard, S. A., Kennedy, C. J., Carmichael, S. L., Lyell, D. J., & Main, E. K. (2020). An expanded obstetric comorbidity scoring system for predicting severe maternal morbidity. *Obstetrics & Gynecology*, *136*(3), 440–449.

Lewis, M. (2007). Informal payments and the financing of health care in developing and transition countries. *Health Affairs*, *26*(4), 984–997. https://doi.org/10.1377/hlthaff.26.4.984

Li, H.-T., Luo, S., Trasande, L., Hellerstein, S., Kang, C., Li, J.-X., Zhang, Y., Liu, J.-M., & Blustein, J. (2017). Geographic variations and temporal trends in cesarean delivery rates in china, 2008-2014. *Jama*, *317*(1), 69–76.

Liu, X., Landon, M. B., Cheng, W., & Chen, Y. (2015). Cesarean delivery on maternal request in china: What are the risks and benefits? *American journal of obstetrics and gynecology*, *212*(6), 817–e1.

McGuire, T. G. (2000). Physician agency. *Handbook of health economics*, *1*, 461–536.

Ministry of Human Resources and Social Security of the People's Republic of China. (2018). 2017 statistical bulletin on human resources and social security development [https://www.mohrss.gov.cn/wap/fw/rssj/201805/W020220325394406391270.pdf (accessed March 9, 2026)].

Olken, B. A. (2007). Monitoring corruption: Evidence from a field experiment in indonesia. *Journal of political Economy*, *115*(2), 200–249.

Pursley, D. M., & Zupancic, J. A. (2020). Using neonatal intensive care units more wisely for at-risk newborns and their families. *JAMA Network Open*, *3*(6), e205693–e205693.

Savva, N., Debo, L., & Shumsky, R. A. (2023). Hospital reimbursement in the presence of cherry picking and upcoding. *Management Science*, *69*(11), 6777–6799.

Sentilhes, L., Madar, H., Ducarme, G., Hamel, J.-F., Mattuizzi, A., & Hanf, M. (2019). Outcomes of operative vaginal delivery managed by residents under supervision and attending obstetricians: A prospective cross-sectional study. *American Journal of Obstetrics and Gynecology*, *221*(1), 59–e1.

Shi, M. (2024). Monitoring for waste: Evidence from medicare audits. *The quarterly journal of economics*, *139*(2), 993–1049.

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of econometrics*, *225*(2), 175–199.

Vance, A. J., Bell, S., Tilea, A., Beck, D., Tabb, K., & Zivin, K. (2023). Identifying neonatal intensive care (nicu) admissions using administrative claims data. *Journal of neonatal-perinatal medicine*, *16*(4), 709–716.

Vian, T. (2020). Anti-corruption, transparency and accountability in health: Concepts, frameworks, and approaches. *Global Health Action*, *13*(sup1), 1694744. https://doi.org/10.1080/16549716.2019.1694744

Wang, Y. (2023). Wealth shocks and physician behavior: Evidence from childbirth.

Wu, B. (2019). Physician agency in china: Evidence from a drug-percentage incentive scheme. *Journal of Development Economics*, *140*, 72–89.

Xu, H., & Yuan, M. (2022). The red packet phenomenon from the perspective of young chinese doctors: A questionnaire study. *BMC Medical Ethics*, *23*(1), 56. https://doi.org/10.1186/s12910-022-00793-w

Yin, S., Zhou, Y., Yuan, P., Wei, Y., Chen, L., Guo, X., Li, H., Lu, J., Ge, L., Shi, H., et al. (2023). Hospital variations in caesarean delivery rates: An analysis of national data in china, 2016-2020. *Journal of Global Health*, *13*, 04029.

Yip, W., Fu, H., Chen, A. T., Zhai, T., Jian, W., Xu, R., Pan, J., Hu, M., Zhou, Z., Chen, Q., et al. (2019). 10 years of health-care reform in china: Progress and gaps in universal health coverage. *The Lancet*, *394*(10204), 1192–1204.

Yip, W. C. (1998). Physician response to medicare fee reductions: Changes in the volume of coronary artery bypass graft (cabg) surgeries in the medicare and private sectors. *Journal of health economics*, *17*(6), 675–699.

Yip, W. C.-M., Hsiao, W., Meng, Q., Chen, W., & Sun, X. (2010). Realignment of incentives for health-care providers in china. *The Lancet*, *375*(9720), 1120–1130.

Zhang, Y., Betran, A., Li, X., Liu, D., Yuan, N., Shang, L., Lin, W., Tu, S., Wang, L., Wu, X., et al. (2022). What is an appropriate caesarean delivery rate for china: A multicentre survey. *BJOG: An International Journal of Obstetrics & Gynaecology*, *129*(1), 138–147.

Zhou, Y., Yu, H., Wang, A., Liu, Y., Li, H., & Liu, J. (2024). Temporal and geographic trends in maternal and paternal ages in china from 2016 to 2020. *Chinese J Reproductive Health*, *35*(4), 301–5.

Zhu, Z., Wang, J., Chen, C., & Zhou, J. (2020). Hospitalization charges for extremely preterm infants: A ten-year analysis in shanghai, china. *Journal of Medical Economics*, *23*(12), 1610–1617.

Table and Figures

Table 1: Summary Statistics of Hospital Characteristics and Clinical Outcomes

	(1) Treatment Group	(2) Spillover Group	(3) Clean Control Group	(4) Diff (1–3)
No. of Admissions	293.98	204.07	106.71	187.27***
	(226.48)	(154.87)	(69.94)	(10.83)
No. of Non-pregnancy Admissions	91.53	17.88	9.80	81.73***
	(77.38)	(16.72)	(11.44)	(3.45)
No. of Pregnancy Admissions	202.45	186.20	96.91	105.54***
	(164.79)	(148.63)	(65.87)	(8.31)
No. of Prenatal Admissions	24.81	20.01	20.86	3.94
	(28.55)	(18.88)	(33.55)	(2.46)
No. of Delivery Admissions	133.25	139.00	54.02	79.23***
	(133.88)	(115.67)	(60.26)	(6.98)
No. of Postnatal Admissions	1.83	1.76	0.73	1.10***
	(2.45)	(2.83)	(1.46)	(0.14)
Ave. Cost per Admission (RMB)	6162.19	6332.45	2674.46	3487.73***
	(2318.16)	(7780.06)	(954.34)	(117.84)
Ave. Drug Cost per Admission (RMB)	1393.62	926.24	450.54	943.08***
	(1001.23)	(722.06)	(261.77)	(46.74)
Ave. Service Cost per Admission (RMB)	2479.19	4310.95	1558.19	921.00***
	(671.81)	(6545.71)	(761.26)	(56.39)
Ave. Diagnosis Cost per Admission (RMB)	1348.86	531.68	445.30	903.55***
	(575.14)	(359.81)	(280.00)	(30.73)
Ave. Consumable Cost per Admission (RMB)	940.51	563.59	220.43	720.08***
	(558.76)	(522.59)	(169.37)	(26.64)
ALOS (days)	7.67	6.23	5.55	2.12***
	(1.31)	(1.60)	(1.43)	(0.11)
C-section Rate (%)	0.44	0.37	0.35	0.09***
	(0.14)	(0.11)	(0.18)	(0.01)
LBW Rate (expressed per 1,000)	71.67	24.07	22.49	49.18***
	(60.44)	(22.65)	(38.97)	(3.85)
VLBW Rate (expressed per 1,000)	13.91	2.84	1.64	12.27***
	(18.79)	(5.37)	(5.81)	(0.99)
Birth Weight (gram)	3262.02	3335.18	3346.27	-84.25***
	(131.52)	(77.56)	(147.15)	(11.31)
Maternal Mortality Rate (expressed per 100,000)	0.00	0.00	43.89	-43.89
	(0.00)	(0.00)	(634.56)	(40.15)
N	250	80	524	774

Notes: This table reports the summary statistics of key hospital-level variables in the baseline period, with standard deviations in parentheses for columns (1)–(3) and standard errors in parentheses for column (4). Column (1) restricts the sample to the treatment group. Columns (2) and (3) divide the control group into a "spillover" group (non-inspected hospitals in districts with at least one inspected hospital) and a "clean" group (non-inspected hospitals in districts without any inspected hospitals), respectively. Column (4) reports the difference in means between treatment group and clean control group. "No. of Admissions" represents the monthly volume of OBGYN patients, with subsequent rows detailing admission counts by specific diagnoses. Financial variables are expressed in RMB and represent the average total cost and its components (drug, service, diagnosis, and consumable costs) per admission. "ALOS" denotes the average length of stay in days. "C-section Rate" represents the percentage of deliveries performed via cesarean section. "LBW Rate" and "VLBW Rate" indicate the prevalence of low and very low birth weight cases, respectively, measured per 1,000 deliveries. "Birth Weight" denotes newborn birth weight in grams. "Maternal Mortality Rate" is scaled as the number of maternal deaths per 100,000 admissions. N denotes the total number of hospital-month observations used in the analysis.

Table 2: Effects of Anti-corruption Hospital Inspection on Hospital Outcomes

	Full	NonPregn	Pregn	PregLoss	Prenatal	Delivery	Postnatal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Aggregated Expenditure (RMB)							
D	401,944***	66,411	335,533***	3,362	67,849***	240,589***	13,102*
	(108,152)	(42,540)	(89,618)	(6,506)	(25,290)	(70,084)	(7,700)
Obs. (Hospital-Month)	2,947	2,947	2,947	2,947	2,947	2,947	2,947
R-squared	0.962	0.954	0.939	0.957	0.729	0.929	0.512
Pretreat Mean T	2,246,161	958,322	1,287,839	207,851	86,110	963,720	12,696
Magnitude	0.180	0.070	0.260	0.020	0.790	0.250	1.030
Panel B: Admission Number							
D	50.69***	7.63***	43.06***	0.30	19.75***	15.99*	2.11*
	(12.30)	(2.86)	(11.48)	(0.84)	(5.39)	(8.34)	(1.18)
Obs. (Hospital-Month)	2,947	2,947	2,947	2,947	2,947	2,947	2,947
R-squared	0.956	0.958	0.936	0.956	0.744	0.916	0.560
Pretreat Mean T	346.83	98.47	248.36	42.54	25.79	173.61	2.31
Magnitude	0.150	0.080	0.170	0.010	0.770	0.090	0.910
Panel C: Average Expenditure (RMB)							
D	-23.41	-27.73	129.67	132.65	208.55	-231.37	518.20
	(144.29)	(209.94)	(135.33)	(170.38)	(153.96)	(164.29)	(340.41)
Obs. (Hospital-Month)	2,947	2,947	2,947	2,947	2,947	2,947	2,947
R-squared	0.938	0.846	0.888	0.634	0.653	0.814	0.407
Pretreat Mean T	6,150.83	8,780.53	4,927.60	4,461.70	3,256.17	5,398.28	3,853.45
Magnitude	-0.004	-0.003	0.026	0.030	0.064	-0.043	0.134
Hospital FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table estimates the impact of hospital inspection on hospital outcome metrics at the hospital-month level. Panel A reports the effects on aggregated medical expenditure (RMB), Panel B on the number of admissions, and Panel C on the average expenditure per admission. Column (1) presents results for the full sample, which is decomposed into non-pregnancy (col. 2) and pregnancy-related (col. 3) admissions. Columns (4)–(7) further categorize pregnancy-related admissions into pregnancy loss, prenatal, delivery, and postnatal care. The classification of different types of OBGYN admissions is based on the ICD-10. Row D presents the DiD coefficients. Magnitude represents the percentage change relative to the pre-treatment mean of the treatment group (Pretreat Mean T). All specifications include hospital and time fixed effects. Standard errors are clustered at the hospital level and reported in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 3: Effect of Anti-corruption Hospital Inspection on Maternal and Infant Health Outcomes

	(1)	(2)	(3)	(4)	(5)
Panel A: Infant Health					
	NICU	LBW	VLBW	BirthWeight	BW_Concentration
D (expressed per 1,000)	-6.92*	-7.36*	-2.09**	17.70***	13.81***
	(3.79)	(4.20)	(0.97)	(6.50)	(5.05)
Wild Bootstrap p-value	[0.073]	[0.082]	[0.047]	[0.021]	[0.004]
Obs. (Delivery Admissions)	332,224	332,224	332,224	332,224	332,224
R-squared	0.039	0.039	0.014	0.034	0.014
Pretreat Mean T	51.68	83.81	16.38	3,229.98	719.38
Magnitude	-0.130	-0.090	-0.130	0.010	0.020
Panel B: Maternal Health					
	SurgAnesth_Comp	Perineal_Lac	Cervical_Lac	ARDS	HyperT_Disorder
D (expressed per 1,000)	-0.19*	-2.57**	-6.20**	-0.14**	-8.76**
	(0.10)	(1.07)	(2.96)	(0.06)	(4.36)
Wild Bootstrap p-value	[0.053]	[0.008]	[0.054]	[0.008]	[0.049]
Obs. (Delivery Admissions)	332,224	332,224	332,224	332,224	332,224
R-squared	0.002	0.012	0.110	0.000	0.030
Pretreat Mean T	0.02	4.74	39.26	0.13	64.19
Magnitude	-11.69	-0.542	-0.158	-1.113	-0.137
Hospital FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes

Notes: This table presents the estimated effects of hospital inspections on infant and maternal health outcomes, utilizing a sample restricted to delivery admissions. Panel A reports infant health outcomes. The dependent variables in columns (1)–(3) and (5) are binary indicators scaled to represent cases per 1,000 admissions: NICU (= 1 for Neonatal Intensive Care Unit admission), LBW (= 1 for low birth weight < 2,500 g), VLBW (= 1 for very low birth weight < 1,500 g), and BW_Concentration (= 1 if birth weight falls within one standard deviation of the sample mean). Column (4) reports continuous birth weight in grams. Panel B reports maternal health outcomes. All dependent variables in Panel B are indicators scaled per 1,000 admissions, capturing surgical or anesthesia complications, perineal laceration, cervical laceration, ARDS (Acute Respiratory Distress Syndrome), and hypertensive disorders of pregnancy and eclampsia. Row D reports the DID coefficients. Magnitude indicates the percentage change relative to the pre-treatment mean of the treatment group (Pretreat Mean T). All specifications include hospital and time fixed effects. Standard errors are clustered at the hospital level and reported in parentheses. Wild cluster bootstrap p-values, computed using Webb weights with 9,999 replications, are reported in brackets. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 4: Heterogeneity Analysis by C-section Rate and Attending Physician Seniority

	Health Outcomes at Admission Level				Hospital Outcomes at Hospital-Month Level			
Outcomes	BW	LBW	Cervical_Lac	HyperT_Disorder	Exp_Tot	Exp_Prenatal	AdmNum_Prenatal	ALOS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Hospitals with Pre-C-section Rate Above 40%								
D	25.70***	-12.07*	-4.72**	-14.65**	381,378**	76,229**	20.03***	-0.31***
	(8.41)	(6.09)	(2.20)	(5.69)	(152,809)	(36,566)	(6.84)	(0.10)
Observations	161,449	161,449	161,449	161,449	1,136	1,136	1,136	1,136
R-squared	0.035	0.040	0.030	0.028	0.961	0.688	0.685	0.904
Pretreat Mean T	3,234.45	85.27	17.43	59.91	2,489,314	89,000	25.12	7.76
Magnitude	0.008	-0.142	-0.271	-0.245	0.153	0.857	0.797	-0.040
Panel B: Hospitals with Pre-C-section Rate Below 40%								
D	2.63	1.85	-8.98	0.51	330,173***	46,780	16.92*	-0.19
	(9.54)	(2.61)	(7.59)	(5.00)	(83,590)	(28,095)	(9.82)	(0.15)
Observations	146,745	146,745	146,745	146,745	1,310	1,310	1,310	1,310
R-squared	0.031	0.036	0.147	0.032	0.959	0.840	0.845	0.899
Pretreat Mean T	3,221.15	80.92	82.34	72.64	1,759,855	80,330	27.13	7.31
Magnitude	0.001	0.023	-0.109	0.007	0.188	0.582	0.624	-0.025
Panel C: Senior Attending Physicians								
D	18.57**	-9.01*	-4.46**	-12.31**	211,445**	39,508**	10.08***	-0.21
	(7.40)	(4.60)	(2.11)	(5.30)	(104,570)	(19,864)	(3.30)	(0.15)
Observations	188,333	188,333	188,333	188,333	2,858	2,858	2,858	2,737
R-squared	0.038	0.041	0.122	0.030	0.944	0.691	0.710	0.728
Pretreat Mean T	3,226.58	85.19	46.52	65.23	1,548,690	54,736	16.65	7.63
Magnitude	0.006	-0.106	-0.096	-0.189	0.137	0.722	0.606	-0.027
Panel D: Junior Attending Physicians								
D	9.08	-0.72	-11.29**	0.25	189,913**	25,185**	8.39**	-0.47*
	(11.23)	(5.81)	(5.59)	(8.90)	(85,406)	(11,162)	(3.78)	(0.24)
Observations	132,292	132,292	132,292	132,292	2,858	2,858	2,858	2,586
R-squared	0.031	0.038	0.052	0.032	0.899	0.773	0.792	0.620
Pretreat Mean T	3,236.33	80.98	23.95	61.61	668,435	29,918	8.72	8.01
Magnitude	0.003	-0.009	-0.471	0.004	0.284	0.842	0.962	-0.059
Hospital FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table presents the heterogeneous policy effects on key hospital and health outcomes. Panels A and B report subgroup analyses stratified by pre-inspection cesarean section (C-section) rates, distinguishing between hospitals with high (above 40%) and low (below 40%) baseline rates. Panels C and D stratify the sample by attending physician seniority, presenting results for senior and junior physicians, respectively. Columns (1)–(4) report the estimated policy effects on key health outcomes observed at the individual level: birth weight, low-birth-weight incidence, cervical laceration, and hypertensive disorders of pregnancy and eclampsia. Columns (5)–(8) report key hospital-level outcomes aggregated at the hospital-month level: total monthly expenditure for OBGYN admissions, total monthly expenditure for prenatal admissions, monthly volume of prenatal admissions, and average length of stay. Observation counts in column (8) of Panels C/D vary from columns (5)–(7) because the ALOS is treated as missing for hospital-month observations with zero admissions for the relevant physician group. Row D reports the DID coefficients. Magnitude indicates the percentage change relative to the pre-treatment mean of the treatment group (Pretreat Mean T). All specifications include hospital and time fixed effects. Standard errors are clustered at the hospital level and reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Drug-Service Substitution

	(1)	(2)	(3)	(4)
Panel A: Component of Medical Expenditure per Admission				
	Exp_Drug	Exp_Service	Exp_Diagnosis	Exp_Consumable
D	-226.63**	216.26***	-31.52	18.49
	(104.79)	(45.92)	(32.48)	(33.32)
Obs. (Hospital-Month)	2,947	2,947	2,947	2,947
R-squared	0.825	0.813	0.957	0.951
Pretreat Mean T	1,363.11	2,486.71	1,360.00	941.00
Magnitude	-0.166	0.087	-0.023	0.020
Panel B: Component of Medical Expenditure per Admission (ZMDP)				
	Exp_Drug	Exp_Service	Exp_Diagnosis	Exp_Consumable
D	-170.94**	96.88**	-27.80	4.95
	(84.45)	(47.63)	(33.53)	(36.28)
D_ZMDP	-111.74**	239.53***	-7.47	27.18
	(50.95)	(63.07)	(36.45)	(33.59)
Obs. (Hospital-Month)	2,947	2,947	2,947	2,947
R-squared	0.827	0.817	0.957	0.951
Pretreat Mean T	1,363.11	2,486.71	1,360.00	941.00
Magnitude	-0.125	0.039	-0.020	0.005
Panel C: Quantity Effect of Lower Drug Utilization				
	#Exp_Drug_0	#Exp_Drug_10	#Exp_Drug_50	#Exp_Drug_100
D	1.65	3.97*	7.52*	11.38**
	(1.23)	(2.21)	(4.40)	(4.46)
Obs. (Hospital-Month)	2,947	2,947	2,947	2,947
R-squared	0.706	0.758	0.841	0.866
Pretreat Mean T	10.63	15.76	37.22	51.92
Magnitude	0.155	0.252	0.202	0.219
Hospital FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes

Notes: This table investigates the policy effects on drug-service substitution at the hospital-month level. Panel A examines the impact on the distinct components of average expenditure per admission. The dependent variables are the average expenditures on drugs (col. 1), medical services (col. 2), diagnostics (col. 3), and consumables (col.4). Panel B further reports the impact after considering the Zero Mark-up Drug Policy. The variable D_ZMDP is an interaction term defined as $\mathbb{1}(\text{MunicipalHospitals}) \times \mathbb{1}(\text{PostDec2016})$, where $\mathbb{1}(\text{MunicipalHospitals})$ is an indicator for municipal- or provincial-level hospitals and $\mathbb{1}(\text{PostDec2016})$ is an indicator for months after December 2016. Panel C presents the effects on the volume of drug utilization, measured by the prevalence of admissions with zero or very low drug expenditures. Specifically, columns (1)–(4) report the estimated effects on the number of admissions with drug expenditures equal to zero, below 10 RMB, below 50 RMB, and below 100 RMB, respectively. Row D reports the DID coefficients. Magnitude indicates the percentage change relative to the pre-treatment mean of the treatment group (Pretreat Mean T). All specifications include hospital and time fixed effects. Standard errors are clustered at the hospital level and reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Heterogeneity Analysis by Reduction in Drug Expenditure

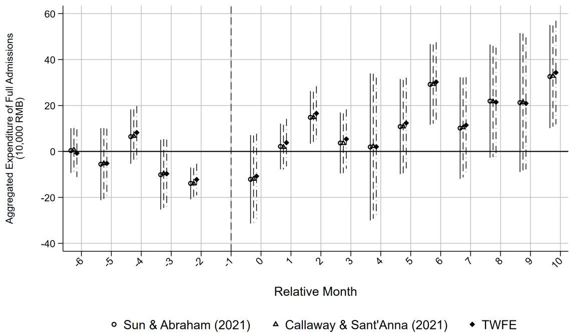
	Health Outcomes at Admission Level				Hospital Outcomes at Hospital-Month Level			
Outcomes	BW	LBW	Cervical_Lac	HyperT_Disorder	Exp_Tot	Exp_Prenatal	AdmNum_Prenatal	ALOS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Hospitals with Above-Median Reductions in Drug Expenditure per Admission								
D	21.01***	-8.87*	-5.91*	-7.97	425,493***	62,732**	16.84***	-0.18*
	(7.68)	(4.93)	(3.32)	(5.29)	(122,194)	(29,407)	(5.71)	(0.10)
Observations	204,473	204,473	204,473	204,473	1,445	1,445	1,445	1,445
R-squared	0.031	0.040	0.031	0.032	0.962	0.788	0.847	0.934
Pretreat Mean T	3,220.92	92.24	22.38	68.92	2,531,111	96,319	27.57	7.77
Magnitude	0.007	-0.096	-0.264	-0.116	0.168	0.651	0.611	-0.023
Panel B: Hospitals with Below-Median Reductions in Drug Expenditure per Admission								
D	9.03	2.06	-5.06*	-3.35	118,838	20,516*	8.46**	-0.21*
	(10.07)	(4.08)	(2.59)	(5.44)	(136,790)	(11,888)	(4.00)	(0.10)
Observations	118,906	118,906	118,906	118,906	1,389	1,389	1,389	1,389
R-squared	0.036	0.010	0.176	0.010	0.923	0.624	0.628	0.876
Pretreat Mean T	3,276.07	40.91	125.16	40.13	10,65,656	43,814	18.46	6.95
Magnitude	0.003	0.050	-0.040	-0.084	0.112	0.468	0.459	-0.030
Panel C: Hospitals with Above-Median Reductions in the Drug-to-Total Revenue Ratio								
D	19.00**	-8.52*	-5.76*	-7.95	374,966***	59,735**	16.49***	-0.22**
	(7.85)	(4.88)	(3.24)	(5.40)	(117,563)	(28,114)	(5.55)	(0.11)
Observations	199,913	199,913	199,913	199,913	1,406	1,406	1,406	1,406
R-squared	0.036	0.039	0.030	0.030	0.961	0.780	0.839	0.923
Pretreat Mean T	3,224.20	90.98	21.96	68.54	2,432,174	95,696	27.12	7.74
Magnitude	0.006	-0.094	-0.262	-0.116	0.154	0.624	0.608	-0.028
Panel D: Hospitals with Below-Median Reductions in the Drug-to-Total Revenue Ratio								
D	13.15	2.34	-5.40*	-5.24	184,772	22,620**	8.33**	-0.22*
	(9.46)	(4.36)	(2.74)	(4.39)	(160,545)	(10,341)	(3.68)	(0.11)
Observations	123,466	123,466	123,466	123,466	1,428	1,428	1,428	1,428
R-squared	0.027	0.010	0.176	0.011	0.927	0.634	0.637	0.895
Pretreat Mean T	3,263.29	42.45	139.05	39.09	1,092,884	26,671	17.56	6.79
Magnitude	0.004	0.055	-0.039	-0.134	0.169	0.848	0.474	-0.032
Hospital FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table presents heterogeneous policy effects on key hospital and health outcomes, stratified by post-inspection reductions in drug expenditures. Panels A and B report subgroup analyses based on the absolute reduction in drug expenditure per admission, comparing hospitals with above- and below-median decreases (measured in RMB). Panels C and D stratify the sample by the reduction in the drug-to-total revenue ratio, distinguishing between hospitals with above- and below-median decreases. Columns (1)–(4) report the estimated policy effects on key health outcomes observed at the individual level: birth weight, low-birth-weight incidence, cervical laceration, and hypertensive disorders of pregnancy and eclampsia. Columns (5)–(8) report key hospital-level outcomes aggregated at the hospital-month level: total monthly expenditure for OBGYN admissions, total monthly expenditure for prenatal admissions, monthly volume of prenatal admissions, and average length of stay. Row D reports the DID coefficients. Magnitude indicates the percentage change relative to the pre-treatment mean of the treatment group (Pretreat Mean T). All specifications include hospital and time fixed effects. Standard errors are clustered at the hospital level and reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

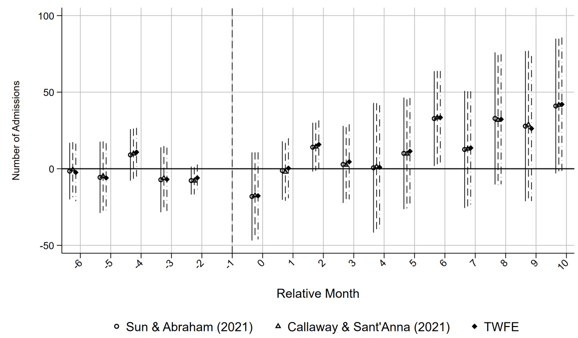
Table 7: Physician Turnover and Maternal and Infant Health Effects

	(1)	(2)	(3)	(4)
Panel A: Physician Turnover				
	Exit#$[-6, +6]$	Entry#$[-6, +6]$	Exit#$[-3, +3]$	Entry#$[-3, +3]$
D	0.15**	0.12*	0.19**	0.20**
	(0.06)	(0.07)	(0.08)	(0.09)
Obs. (Hospital-Month)	970	970	543	543
R-squared	0.225	0.253	0.253	0.311
Pretreat Mean T	0.09	0.26	0.04	0.16
Magnitude	1.762	0.454	4.691	1.248
Panel B: Health Effects of Hospitals W/ Exit Physicians within 6 Months before/after Inspection				
	BW	LBW	Cervical_Lac	HyperT_Disorder
D	20.59***	-10.52**	-7.08**	-11.26**
	(6.98)	(4.75)	(3.49)	(4.99)
Obs. (Delivery Admissions)	258,758	258,758	258,758	258,758
R-squared	0.039	0.044	0.110	0.031
Pretreat Mean T	3,218.99	88.96	44.39	66.99
Magnitude	0.006	-0.118	-0.160	-0.168
Panel C: Health Effects of Hospitals W/O Exit Physicians within 6 Months before/after Inspection				
	BW	LBW	Cervical_Lac	HyperT_Disorder
D	4.55	5.21	-1.50	-1.68
	(20.16)	(3.08)	(1.20)	(4.48)
Obs. (Delivery Admissions)	61,867	61,867	61,867	61,867
R-squared	0.014	0.011	0.009	0.013
Pretreat Mean T	3,304.65	48.07	2.44	43.58
Magnitude	0.001	0.108	-0.614	-0.038
Hospital FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes

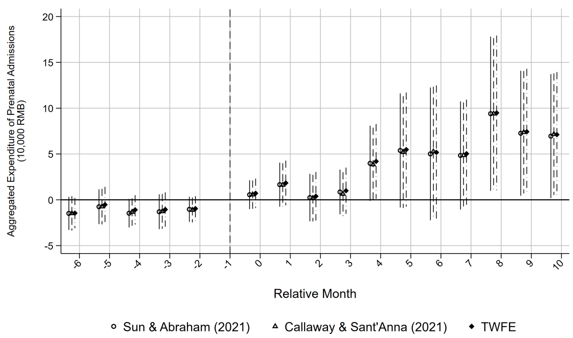
Notes: This table reports the impact of hospital inspections on physician turnover and health outcomes. Panel A presents the policy effects on the volume of departing (Exit#) and newly arriving (Entry#) attending physicians, identified by the last and first appearance of a physician's name in the dataset, respectively. Columns (1) and (2) restrict the study period to a six-month window before and after the inspection for each hospital, while columns (3) and (4) restrict the period to a three-month window. Panels B and C report the impact on key health outcomes, stratifying the sample into hospitals with at least one physician exit within the six-month window and hospitals with no physician exits during the same period, respectively. For Panels B and C, the dependent variables in columns (1)–(4) are continuous birth weight, low-birth-weight incidence, cervical laceration incidence, and the incidence of hypertensive disorders of pregnancy and eclampsia, respectively. Row D reports the DID coefficients. Magnitude indicates the percentage change relative to the pre-treatment mean of the treatment group (Pretreat Mean T). All specifications include hospital and time fixed effects. Standard errors are clustered at the hospital level and reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.



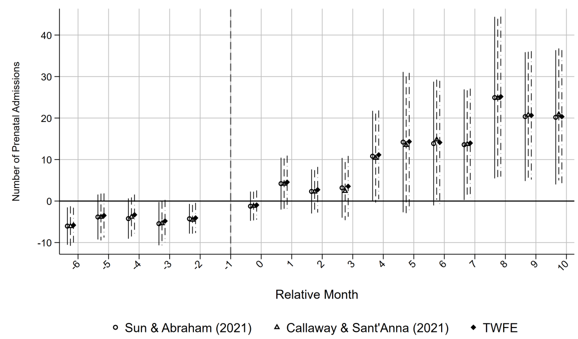
Panel A: Aggregated Expenditure



Panel B: Admission Number



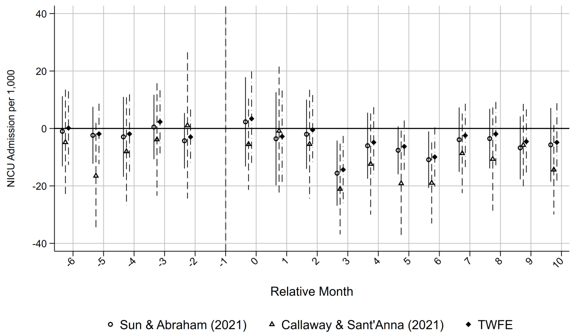
Panel C: Prenatal Expenditure



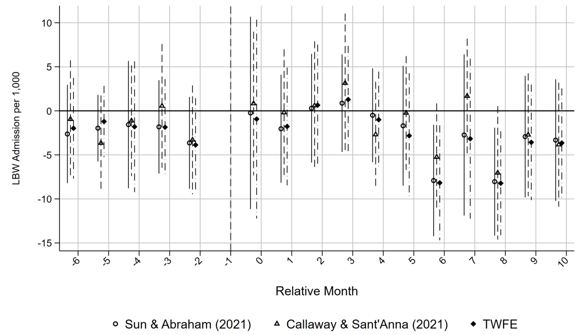
Panel D: Prenatal Number

Figure 1: Event Study of Hospital Outcomes

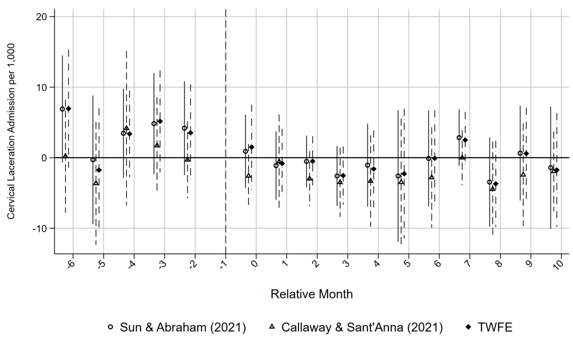
Notes: This figure plots event study estimates illustrating the dynamic impacts of hospital inspections on key hospital outcomes, specifically aggregate expenditures (measured in 10,000 RMB) and admission volumes. Panels A and B present the estimated effects on aggregate expenditures and the total number of admissions for all OBGYN patients, respectively. Panels C and D report the corresponding effects on aggregate expenditures and admission volumes for prenatal care. To ensure robustness against potential biases in standard difference-in-differences models with staggered timing, the figure compares results from three distinct specifications: (1) the standard Two-Way Fixed Effects (TWFE) model (solid diamonds), (2) the interaction-weighted estimator by Sun and Abraham (2021) (hollow circles), and (3) the estimator by Callaway and Sant'Anna (2021) (triangles). The horizontal axis represents the month relative to the inspection month ($t = 0$), and vertical bars represent 95% confidence intervals. The event window spans from 6 months prior to 10 months post-inspection; observations outside this window are dropped, and $t = -1$ serves as the omitted reference period. All specifications control for hospital and time fixed effects, with standard errors clustered at the hospital level.



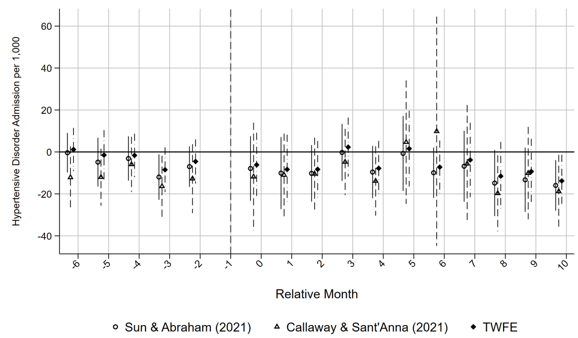
Panel A: Inferred NICU Admission



Panel B: Low-Birth-Weight Admission



Panel C: Cervical Laceration Admission



Panel D: Hypertensive Disorders

Figure 2: Event Study of Health Outcomes

Notes: This figure plots event study estimates illustrating the dynamic effects of hospital inspections on key infant and maternal health outcomes. Panels A through D present the estimated effects on the incidence of inferred infant NICU admissions, low-birth-weight cases, cervical lacerations, and hypertensive disorders of pregnancy and eclampsia, respectively. To ensure robustness against potential biases in standard difference-in-differences models with staggered timing, the figure compares results from three distinct specifications: (1) the standard Two-Way Fixed Effects (TWFE) model (solid diamonds), (2) the interaction-weighted estimator by Sun and Abraham (2021) (hollow circles), and (3) the estimator by Callaway and Sant'Anna (2021) (triangles). The horizontal axis represents the month relative to the inspection month ($t = 0$), and vertical bars represent 95% confidence intervals. The event window spans from 6 months prior to 10 months post-inspection; observations outside this window are dropped, and $t = -1$ serves as the omitted reference period. All specifications control for hospital and time fixed effects, with standard errors clustered at the hospital level.

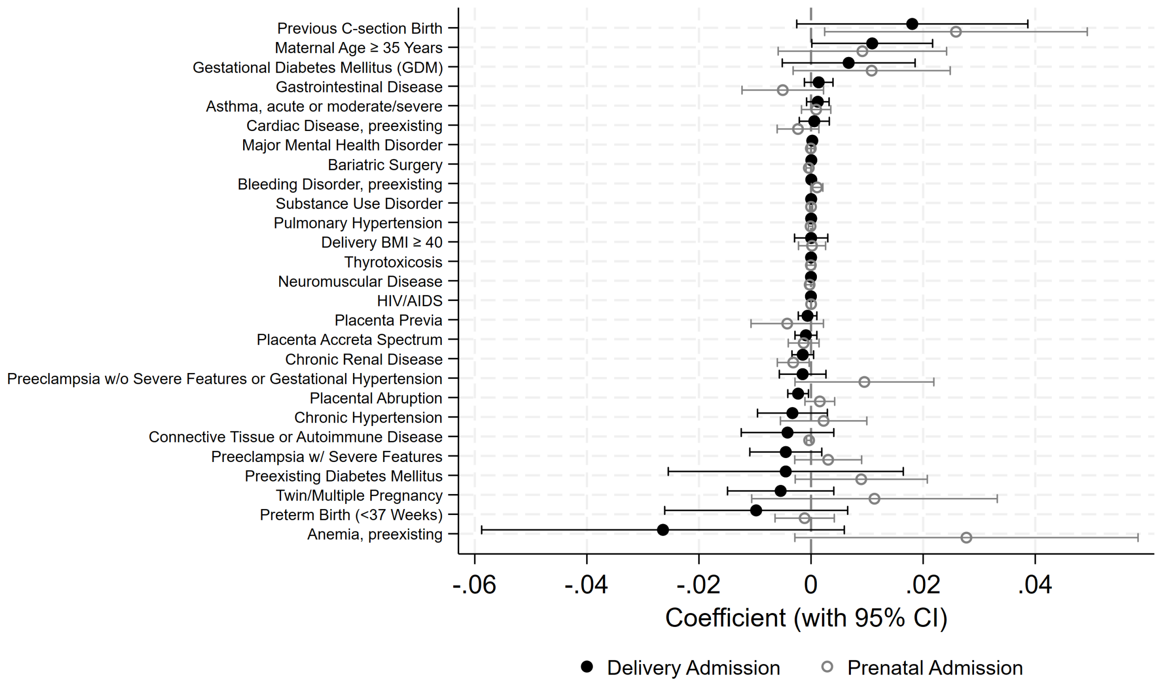


Figure 3: Impact on the Composition of Obstetric Comorbidities

Notes: This figure plots the estimated regression coefficients for the likelihood of 27 specific obstetric comorbidities among prenatal admissions ($N = 64,000$) and delivery admissions ($N = 332,224$). The dependent variable in each regression is a binary indicator equal to one if the patient is diagnosed with the specific comorbidity listed on the y-axis during their admission, and zero otherwise. The set of 27 comorbidities is selected following the clinical risk classification by Leonard et al. (2020) as predictors of severe maternal morbidity. The dots represent the point estimates of the treatment effect, and the horizontal lines indicate the 95% confidence intervals. All specifications control for hospital and time fixed effects, with standard errors clustered at the hospital level.

A Appendix Tables and Figures

Table A1: Sample Composition

	(1) Full Sample	(2) 2015–2017	(3) No Small Hosp.	(4) No Hosp. W/O CB	(5) No Outlier Month	(6) No 3rd-Round Hosp.
Demographic Information						
Age at Adm.	31.57	31.36	31.41	31.21	31.21	31.00
Share Female	0.99	0.99	0.99	0.99	0.99	0.99
Share Married	0.94	0.94	0.94	0.94	0.94	0.94
Share UEBMI	0.15	0.15	0.16	0.16	0.16	0.16
Share URBMI and NCMS	0.52	0.53	0.53	0.52	0.52	0.53
Clinical Information						
Share ED Adm.	0.14	0.14	0.14	0.15	0.15	0.15
Share Delivery Adm.	0.48	0.49	0.49	0.59	0.59	0.59
Length of Stay	6.64	6.52	6.54	6.46	6.46	6.36
Per-adm. Cost	5,003.22	5,003.22	5,051.01	5,014.24	5,014.98	4,854.64
Per-adm. Drug Cost	966.65	966.65	978.59	886.74	886.79	832.61
Per-adm. Service Cost	2,411.24	2,411.24	2,426.47	2,473.12	2,473.44	2,431.12
Per-adm. Diagnosis Cost	952.13	952.13	962.61	966.49	966.71	942.09
Per-adm. Consumable Cost	673.20	673.20	683.35	687.90	688.05	648.83
N	1,091,888	929,837	910,578	759,343	758,972	686,588
% of Sample (col. [1])	100%	85.16%	83.39%	69.54%	69.51%	62.88%
Share Treatment	0.54	0.47	0.48	0.53	0.53	0.48

Notes: This table displays summary statistics across different sample restrictions, with each column adding a further constraint to the raw OBGYN discharge records. Column (1) reports all OBGYN admissions. Column (2) limits the sample period to January 2015–December 2017. Column (3) removes small hospitals based on bed count and average monthly admissions, while column (4) drops hospitals missing birth weight data for over half the study months. Column (5) excludes outlier months with fewer than 10 admissions. Finally, column (6) removes the seven hospitals inspected during the September 2017 inspection round. The final treatment and control groups are built from this restricted sample. “Demographic Information” includes patient age, gender, marital status, and insurance type at admission, while “Clinical Information” comprises the emergency department (ED) admission rate, the share of deliveries among OBGYN admissions, length of stay, and per-admission expenditure (drugs, services, diagnostics, and consumables). “N” represents the number of remaining admissions. “% of Sample” is the percentage of the raw data retained, and “Share Treatment” is the percentage of total admissions occurring in inspected hospitals.

Table A2: Effect of Anti-corruption Hospital Inspection on Deferrable and Non-deferrable Admissions

	Full	NonPregn	Pregn	PregLoss	Prenatal	Delivery	Postnatal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Number of Non-deferrable Admissions							
D	22.63***	0.92***	21.70***	1.04	4.03***	15.24***	0.55***
	(5.75)	(0.35)	(5.62)	(0.74)	(1.40)	(5.07)	(0.19)
Obs. (Hospital-Month)	2,947	2,947	2,947	2,947	2,947	2,947	2,947
R-squared	0.898	0.728	0.893	0.866	0.780	0.869	0.622
Pretreat Mean T	63.45	3.64	59.81	7.84	5.39	44.95	0.77
Magnitude	0.357	0.253	0.363	0.133	0.749	0.339	0.712
Panel B: Number of Deferrable Admissions							
D	28.06**	6.71**	21.36*	-0.74	15.71***	0.75	1.56
	(12.30)	(2.74)	(12.13)	(1.00)	(4.48)	(8.07)	(1.07)
Obs. (Hospital-Month)	2,947	2,947	2,947	2,947	2,947	2,947	2,947
R-squared	0.949	0.957	0.921	0.945	0.702	0.900	0.469
Pretreat Mean T	283.38	94.82	188.55	34.70	20.41	128.66	1.54
Magnitude	0.099	0.071	0.113	-0.021	0.770	0.006	1.011
Hospital FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table estimates the impact of hospital inspection on the number of non-deferrable and deferrable admissions at the hospital-month level. Panel A reports the effects on the number of non-deferrable admissions while Panel B on the number of deferrable admissions. Column (1) presents results for the full sample, which is decomposed into non-pregnancy (col. 2) and pregnancy-related (col. 3) admissions. Columns (4)–(7) further categorize pregnancy-related admissions into pregnancy loss, prenatal, delivery, and postnatal care. The classification of different types of OBGYN admissions is based on the ICD-10. Row D presents the DiD coefficients. Magnitude represents the percentage change relative to the pre-treatment mean of the treatment group (Pretreat Mean T). All specifications include hospital and time fixed effects. Standard errors are clustered at the hospital level and reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A3: Effects on Admission Numbers (Robustness to the ZMDP)

	Full	NonPregn	Pregn	PregLoss	Prenatal	Delivery	Postnatal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
D	40.74***	3.03	37.70***	0.18	15.69***	15.89*	1.62*
	(10.49)	(2.23)	(9.82)	(1.02)	(4.60)	(8.13)	(0.87)
D_ZMDP	19.97**	9.22***	10.75	0.23	8.13**	0.21	0.98
	(9.34)	(3.01)	(8.45)	(0.97)	(3.73)	(6.19)	(0.71)
Obs. (Hospital-Month)	2,947	2,947	2,947	2,947	2,947	2,947	2,947
R-squared	0.957	0.958	0.936	0.956	0.746	0.916	0.563
Hospital FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pretreat Mean T	346.83	98.47	248.36	42.54	25.79	173.61	2.31
Magnitude	0.117	0.031	0.152	0.004	0.608	0.092	0.699

Notes: This table reports the robustness checks for the impact of hospital inspection on the volume of admissions after controlling for the Zero Mark-up Drug Policy. Column (1) presents results for the full sample, which is decomposed into non-pregnancy (col. 2) and pregnancy-related (col. 3) admissions. Columns (4)–(7) further categorize pregnancy-related admissions into pregnancy loss, prenatal, delivery, and postnatal care. The variable D_ZMDP is an interaction term defined as $\mathbb{1}(\text{MunicipalHospitals}) \times \mathbb{1}(\text{PostDec2016})$, where $\mathbb{1}(\text{MunicipalHospitals})$ is an indicator for municipal- or provincial-level hospitals and $\mathbb{1}(\text{PostDec2016})$ is an indicator for months after December 2016. Row D presents the DiD coefficients. Magnitude represents the percentage change relative to the pre-treatment mean of the treatment group (Pretreat Mean T). All specifications include hospital and time fixed effects. Standard errors are clustered at the hospital level and reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

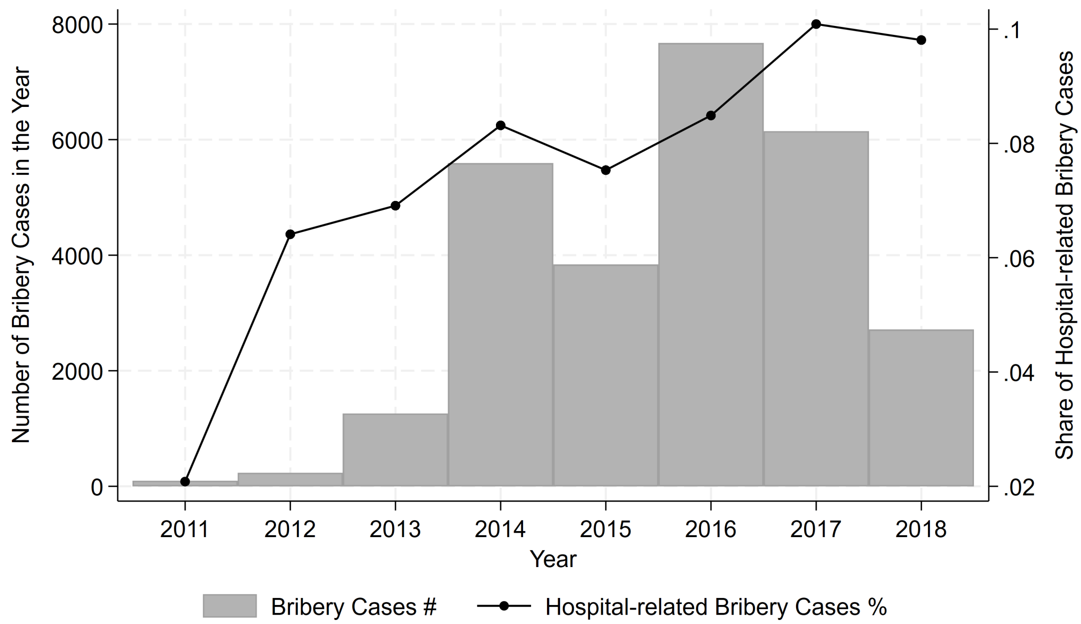
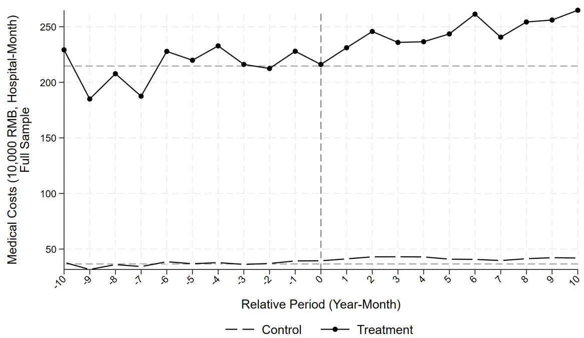
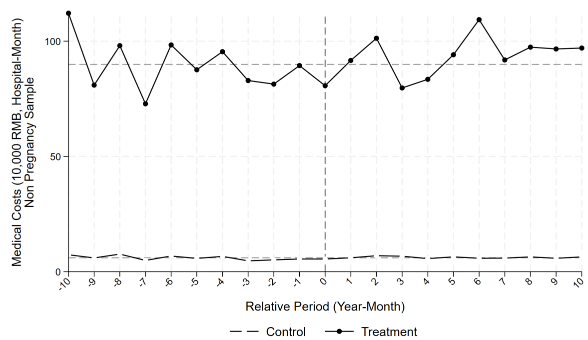


Figure A1: Annual Trends in Corruption Cases and the Hospital-Related Share

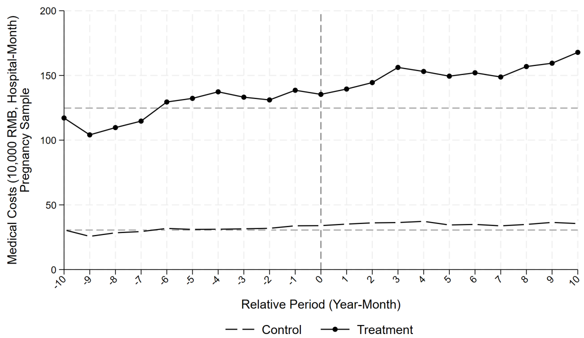
Notes: This figure plots the annual evolution of bribery litigation in China from 2011 to 2018 using data sourced from China Judgements Online. The gray histogram (left vertical axis) displays the total number of bribery cases per year, while the solid line (right vertical axis) indicates the annual share of these cases that are hospital-related. A case is defined as ”hospital-related” if textual analysis of the case description identifies the explicit involvement of a hospital or physicians. The figure highlights the temporal relationship between aggregate corruption cases and the relative prominence of healthcare-sector corruption.



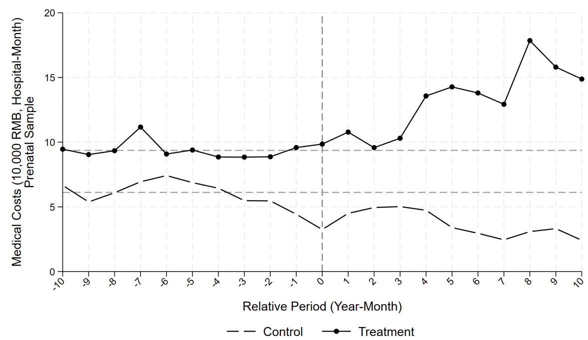
Panel A: Full Admissions



Panel B: Non-pregnancy Admissions



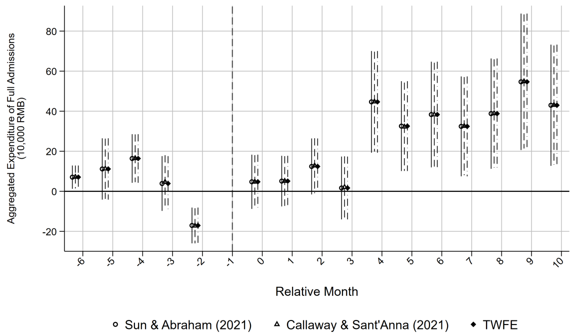
Panel C: Pregnancy Admissions



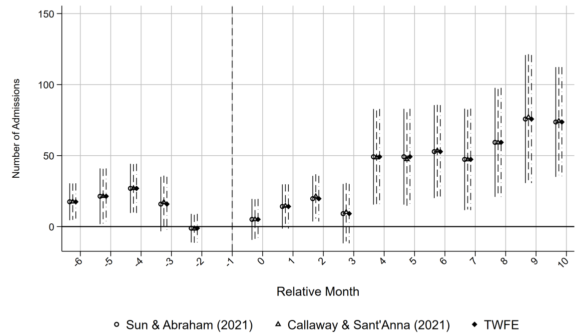
Panel D: Prenatal Admissions

Figure A2: Raw Trend of Aggregated Expenditure by Admission Types

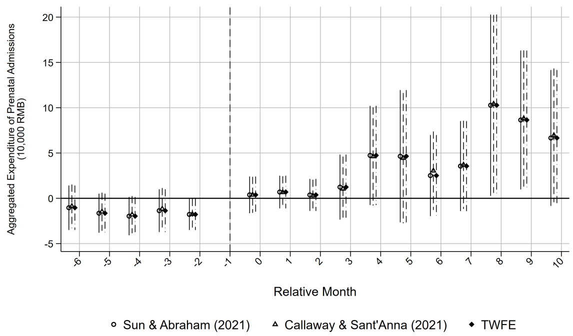
Notes: This figure illustrates raw trends in aggregate medical expenditures at the hospital-month level across different admission types. Panels A–D plot the trends for all OBGYN, non-pregnancy, pregnancy, and prenatal admissions, respectively. The vertical axis measures expenditures in 10,000 RMB. Solid lines depict the trajectories for inspected hospitals (the treatment group), while dashed lines represent non-inspected hospitals (the control group). The horizontal dashed lines indicate the pre-inspection means for the treatment and control groups. The horizontal axis indicates the relative period (in months) centered on the inspection date ($t = 0$). For the control group, placebo inspection dates were randomly assigned based on the temporal distribution of actual inspections to allow for a direct comparison. The figure plots the evolution of expenditure from 10 months prior to 10 months following the inspection (or placebo) date.



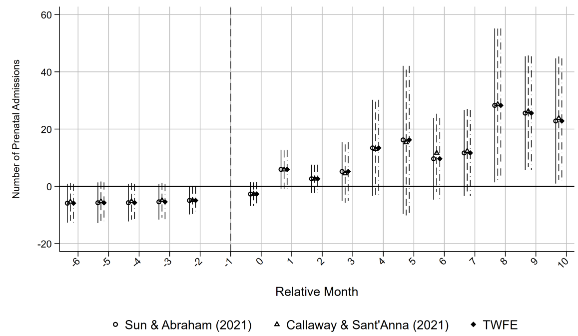
Panel A: (Hospital Outcomes) Aggregated Expenditure



Panel B: (Hospital Outcomes) Admission Number



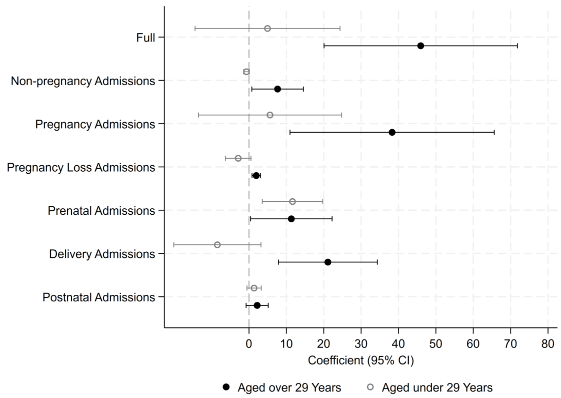
Panel C: (Hospital Outcomes) Prenatal Expenditure



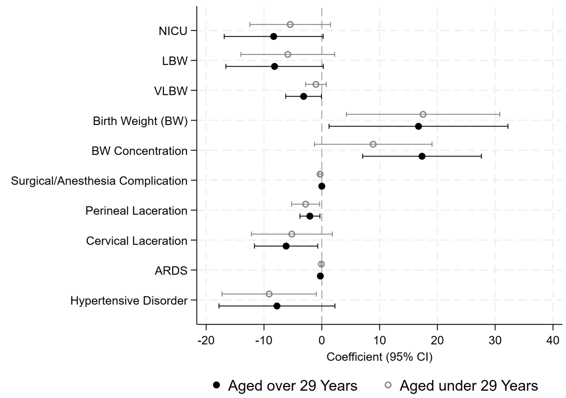
Panel D: (Hospital Outcomes) Prenatal Number

Figure A3: Event Study of Hospital Outcomes (Robustness to the Two-Child Policy)

Notes: This figure plots event study estimates illustrating the dynamic impacts of hospital inspections on key hospital outcomes—specifically aggregate expenditures (measured in 10,000 RMB) and admission volumes—after excluding treated hospitals inspected after the implementation of two-child policy. Panels A and B present the estimated effects on aggregate expenditures and the total number of admissions for all OBGYN patients, respectively. Panels C and D report the corresponding effects on aggregate expenditures and admission volumes for prenatal care. To ensure robustness against potential biases in standard difference-in-differences models with staggered timing, the figure compares results from three distinct specifications: (1) the standard Two-Way Fixed Effects (TWFE) model (solid diamonds), (2) the interaction-weighted estimator by Sun and Abraham (2021) (hollow circles), and (3) the estimator by Callaway and Sant'Anna (2021) (triangles). The horizontal axis represents the month relative to the inspection month ($t = 0$), and vertical bars represent 95% confidence intervals. The event window spans from 6 months prior to 10 months post-inspection; observations outside this window are dropped, and $t = -1$ serves as the omitted reference period. All specifications control for hospital and time fixed effects, with standard errors clustered at the hospital level.



Panel A: Effect on Admission Number



Panel B: Effect on Health Outcomes

Figure A4: Effect of Hospital Inspection by Age Group

Notes: This figure illustrates the heterogeneous effects of hospital inspections, with estimates partitioned by maternal age (under 29 vs. 29 and older). Panel A presents the estimated coefficients from Equation (1) regarding admission volumes, while Panel B displays the results from Equation (2) for maternal and infant health outcomes. Filled circles denote point estimates for the older age group, and hollow circles denote point estimates for the younger age group. Horizontal lines represent 95% confidence intervals. All specifications control for hospital and time fixed effects, with standard errors clustered at the hospital level.

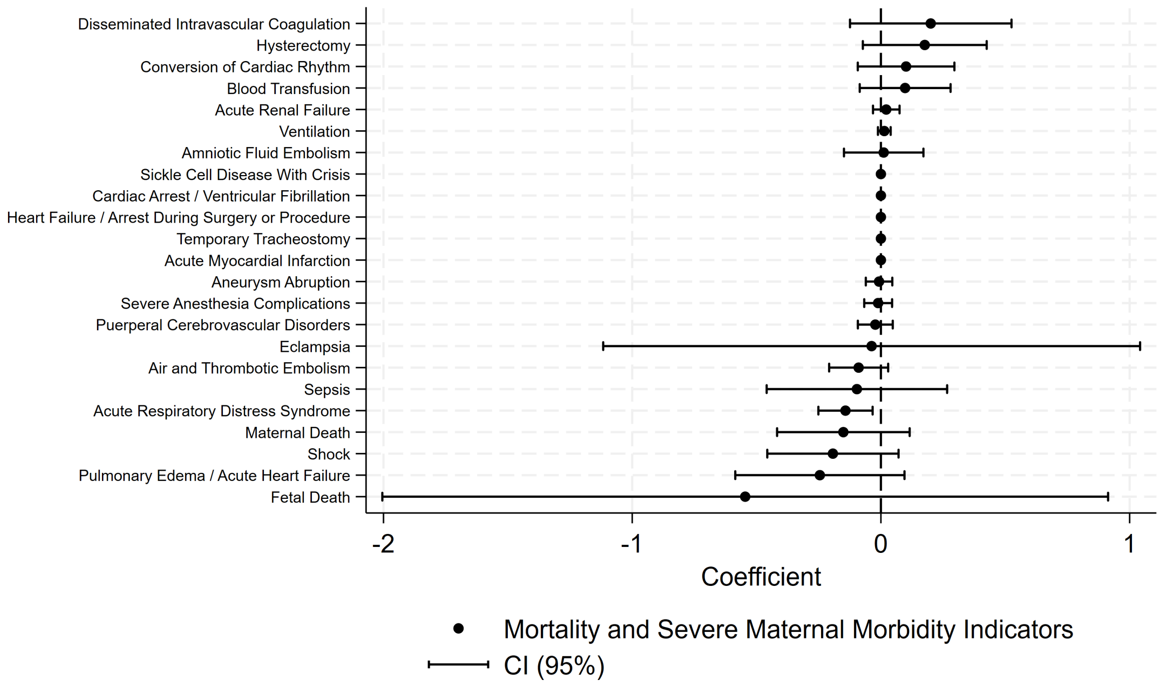
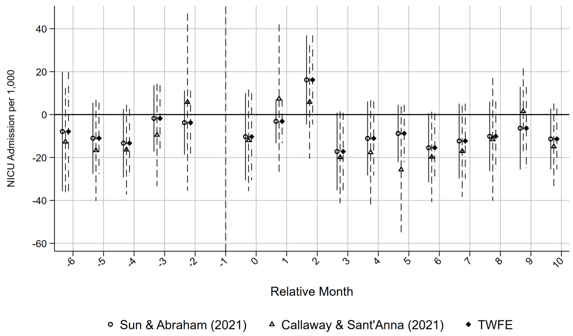
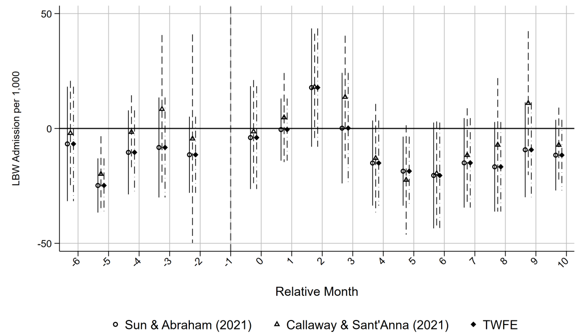


Figure A5: Impact on Other Maternal and Infant Health Outcomes

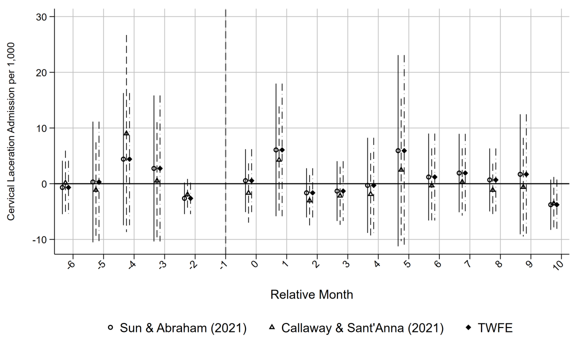
Notes: This figure plots the estimated regression coefficients capturing the effect of the anti-corruption inspections on the incidence of maternal death, fetal death, and 21 specific indicators of severe maternal morbidity (SMM) (Centers for Disease Control and Prevention, 2024), scaled per 1,000 births. In each separate regression, the dependent variable is a binary indicator that equals one if the patient was diagnosed with the corresponding condition listed on the vertical axis, and zero otherwise. Solid markers denote the point estimates of the treatment effect, and the horizontal error bars represent the 95% confidence intervals. All specifications control for hospital and time fixed effects, with standard errors clustered at the hospital level.



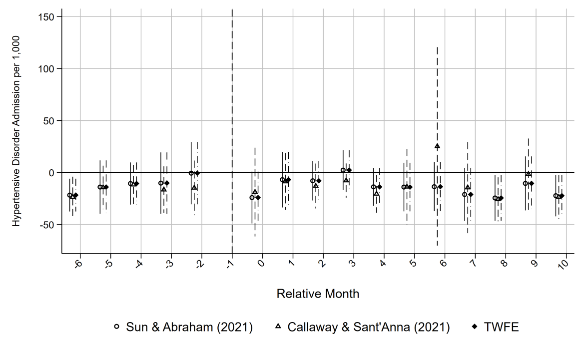
Panel A: Inferred NICU Admission



Panel B: Low Birth Weight



Panel C: Cervical Laceration Admission



Panel D: Hypertensive Disorders

Figure A6: Event Study of Health Outcomes (Robustness to the Two-Child Policy)

Notes: This figure plots event study estimates illustrating the dynamic effects of hospital inspections on key maternal and infant health outcomes after excluding treated hospitals inspected after the implementation of two-child policy. Panels A through D present the estimated effects on the incidence of inferred infant NICU admissions, low-birth-weight cases, cervical lacerations, and hypertensive disorders of pregnancy and eclampsia, respectively. To ensure robustness against potential biases in standard difference-in-differences models with staggered timing, the figure compares results from three distinct specifications: (1) the standard Two-Way Fixed Effects (TWFE) model (solid diamonds), (2) the interaction-weighted estimator by Sun and Abraham (2021) (hollow circles), and (3) the estimator by Callaway and Sant'Anna (2021) (triangles). The horizontal axis represents the month relative to the inspection month ($t = 0$), and vertical bars represent 95% confidence intervals. The event window spans from 6 months prior to 10 months post-inspection; observations outside this window are dropped, and $t = -1$ serves as the omitted reference period. All specifications control for hospital and time fixed effects, with standard errors clustered at the hospital level.