Gaming the Medical Loss Ratio: Vertical Integration and Regulatory Arbitrage in Health Insurance

Lynn M. Hua*

Greenberg School of Risk Science, Georgia State University

March 2026

Preliminary Draft: Please Do Not Circulate

Abstract

When a regulated firm achieves nominal compliance without the behavioral change the regulation requires, consumer protection fails. We provide causal evidence using the Affordable Care Act's 80 percent medical loss ratio requirement. A triple-differences design exploiting cross-sectional variation in pre-regulation compliance gaps shows that organizational form determines how insurers respond: traditional insurers exhibit a strong dose-response to regulatory pressure; vertically integrated insurers show a response ninety percent smaller, decoupled from their compliance gap. The mechanism is internal transfer pricing: vertically integrated insurers satisfy the 80 percent threshold by inflating payments to owned providers, retaining revenue within the corporate family rather than reallocating it to enrollees. Decomposing the MLR numerator confirms this: the differential response runs entirely through medical claims, with near-zero coefficients on quality improvement and administrative costs, ruling out accounting reclassification. Vertically integrated insurers also maintained systematically lower recorded claims before the ACA, consistent with strategic suppression of transfer prices when there was no regulatory incentive to report high claims.

JEL Codes: I13, I18, L22, L51

Keywords: medical loss ratio, vertical integration, regulatory arbitrage, health insurance, transfer pricing

*Email: lhua@gsu.edu. I thank the National Association of Insurance Commissioners for data access and financial support through the NAIC Research Fellows Program. The views expressed are solely my own and do not represent the views of the NAIC. The author declares no other sources of financial support or conflicts of interest. All errors are my own.

1 Introduction

Can firms comply with a binding regulation without actually doing what the regulation requires? This question is central to the welfare consequences of regulatory policy: a firm that achieves nominal compliance through accounting manipulation rather than genuine behavioral change imposes no compliance cost while providing no benefit to the consumers the regulation was designed to protect. This paper provides causal evidence on this question using the Affordable Care Act's requirement that health insurers spend at least 80 percent of premium revenue on medical claims and shows that the answer depends critically on the organizational form of the regulated firm.

The medical loss ratio (MLR) provision, effective 2011, requires individual market insurers to provide rebates to enrollees when their ratio of medical claims to premiums falls below 80 percent. Cicala, Lieber, and Marone (2019) show that the regulation produced large, causal increases in medical spending: issuers below the threshold closed most of their compliance gap in the years following implementation, consistent with genuine care expansion. What their analysis cannot determine is whether this response was uniform across regulated firms or whether some firms achieved compliance through a different mechanism entirely. We show that organizational form generates a stark divide: vertically integrated insurers, which own the providers to whom they pay medical claims, achieve compliance through a mechanism almost entirely decoupled from the size of their compliance gap, while traditional insurers show the strong dose-response Cicala, Lieber, and Marone (2019) document.

The distinction matters because of transfer pricing. When an insurer owns its providers, the price it records as a medical claim is an internal accounting entry, not a true market transaction. A vertically integrated insurer facing an MLR shortfall

can close the compliance gap by adjusting the transfer prices it pays owned providers. This is a bookkeeping change that raises the MLR numerator without changing the volume or quality of care delivered to enrollees. A traditional insurer facing the same shortfall must negotiate higher payment rates in competitive provider markets or genuinely expand utilization. Both are costly and scale with the size of the gap to be closed. Consistent with this interpretation, VI issuers with larger compliance gaps maintained anomalously low recorded claims before the ACA—suggesting the same transfer pricing lever was being pulled in the opposite direction when there was no regulatory incentive to report high claims. If instead VI insurers comply through genuine delivery efficiency gains, their response should scale with gap size just as traditional insurers' does. The organizational difference generates a testable prediction: transfer pricing predicts a VI response independent of gap size, while genuine care expansion predicts proportionality.

Using NAIC data on individual market insurers from 2001 to 2024, we implement a continuous-treatment triple-differences design that interacts each issuer's pre-regulation distance below the 80 percent threshold with post-ACA timing and vertical integration status. Vertically integrated insurers show a near-zero dose-response to regulatory pressure—an implied slope of 0.070, approximately 90 percent smaller than the traditional insurer estimate—and a flat post-ACA intercept of approximately 8 percentage points. Regardless of how far below the threshold they started, VI insurers increased their MLR by the same amount. In contrast, and consistent with Cicala, Lieber, and Marone (2019), traditional insurers close approximately 71 percent of their pre-regulation compliance gap. This tenfold difference in dose-responses is highly significant, stable across specifications, and persistent through 2024.

Four results support the transfer pricing interpretation. First, VI issuers with larger pre-regulation compliance gaps maintained anomalously low MLR before the

ACA, consistent with profit extraction through suppressed transfer prices when there was no regulatory incentive to report high claims. Second, decomposing the MLR numerator using supplemental regulatory filings confirms that VI issuers' attenuated response runs entirely through the claims channel, with near-zero coefficients on quality improvement expenses and administrative costs, ruling out the accounting reclassification strategies documented among traditional insurers in the prior literature (Born, Sirmans, and Steinorth, 2023; Eastman, Eckles, and Van Buskirk, 2021). Third, the event study shows that the VI–traditional gap in regulatory response widens persistently over thirteen post-regulation years rather than converging, consistent with gaming that deepens as integration expands. Fourth, a sign reversal in the VI triple-difference coefficient when issuer-specific linear trends are added reveals that VI issuers were already on upward MLR trajectories before the ACA that the simpler specification incorrectly attributes to regulatory response.

The paper makes three contributions. First, it provides the first causal evidence that organizational form determines the *mechanism* through which a regulated firm achieves compliance, not merely whether it complies. Prior work documents behavioral responses to the 80 percent threshold and identifies gaming channels available to traditional insurers (Abraham, Karaca-Mandic, and Simon, 2014; Born, Sirmans, and Steinorth, 2023; Callaghan, Plummer, and Wempe, 2020; Cicala, Lieber, and Marone, 2019; Eastman, Eckles, and Van Buskirk, 2021; Karaca-Mandic, Abraham, and Simon, 2015), but does not address how organizational form determines which compliance mechanisms a firm can access. We show that vertical integration expands that set by creating access to internal transfer pricing as a compliance technology unavailable to arms'-length competitors. This complements Gupta, La Forgia, and Sacarny (forthcoming), who show that ownership type shapes both the propensity to exploit payment rules and how resulting revenue is deployed. Prior work on health

care gaming has documented upcoding and risk manipulation across a range of settings (Dafny, 2005), but focuses on the existence and magnitude of gaming rather than the organizational conditions that enable it.

Second, the MLR numerator decomposition identifies which channel carries the effect. Precisely estimated zeros on quality improvement and administrative cost shares confirm that the full attenuation of VI issuers' response runs through claims alone, ruling out the accounting reclassification strategies documented in the prior literature (Born, Sirmans, and Steinorth, 2023; Callaghan, Plummer, and Wempe, 2020; Eastman, Eckles, and Van Buskirk, 2021) and identifying intra-organizational transfer pricing as the operative mechanism. This channel does not appear in insurer-reported data and is invisible to the compliance rule as written.

Third, these results add a regulatory arbitrage dimension to the welfare calculus of vertical integration (Dafny, Ho, and Lee, 2019; Gaynor, Ho, and Town, 2015; Koch, Wendling, and Wilson, 2021). Even where integration generates genuine efficiency gains, it simultaneously undermines spending mandates by giving integrated firms access to compliance mechanisms that do not require genuine reallocation toward patient care, a consideration most existing welfare analyses overlook.

The paper proceeds as follows. Section 2 describes the MLR provision and the role of vertical integration in health insurance markets. Sections 3 and 4 present the data, identification strategy, and main results. Section 5 presents mechanism evidence. Section 6 presents the event study. Section 7 presents robustness checks. Section 8 concludes.

2 Institutional Background

2.1 The ACA's Medical Loss Ratio Provision

The Affordable Care Act requires health insurers in the individual market to spend at least 80 percent of earned premiums on medical claims and quality improvement activities, remitting rebates to enrollees when spending falls short of the threshold. The provision took effect in 2011, with the first rebates distributed in August 2012. Cicala, Lieber, and Marone (2019) provide causal evidence of near-complete compliance convergence following implementation.

The regulatory formula defines the medical loss ratio (MLR) as incurred claims plus quality improvement (QI) expenses divided by earned premiums net of taxes. Administrative expenses, including claims processing, marketing, and profit, do not count toward the numerator. This structure creates two gaming channels relevant to our analysis. First, an insurer can inflate the claims component by recording higher prices paid to providers, even if care volume is unchanged; this is the channel we study. Second, an insurer can reclassify administrative expenses as QI activities; prior work documents evidence of this channel (Born, Sirmans, and Steinorth, 2023; Eastman, Eckles, and Van Buskirk, 2021).[1]

Enforcement is administered through annual NAIC financial filings at the state-market segment level. The regulation applies to fully insured coverage; self-insured employer plans are excluded.

[1]The large group market threshold is 85 percent.

2.2 Vertical Integration in Health Insurance Markets

Vertical integration between health insurers and health care providers takes several organizational forms in the US market. Provider-sponsored health plans—the form we study—are insurance entities whose parent organization is a health care delivery system. These include fully integrated delivery systems such as Kaiser Permanente, hospital-system-affiliated plans such as UPMC Health Plan and Geisinger Health Plan, and physician-group-affiliated plans that bundle primary care delivery with insurance coverage.

The central economic feature of insurer-provider integration is the elimination of arms'-length bargaining over provider payment rates. A traditional (non-integrated) insurer must negotiate prices with independent hospitals and physician groups in an external market, resulting in rates that reflect provider market power, network design, and utilization management. A vertically integrated insurer pays itself: the price at which the insurer arm compensates the provider arm is an internal transfer price set within the integrated organization rather than determined by market negotiation. This transfer price governs the allocation of combined profits between the insurance and delivery components of the balance sheet but does not affect the integrated enterprise's total income.

Under the MLR provision, this transfer pricing flexibility creates a regulatory arbitrage opportunity unavailable to traditional competitors. A traditional insurer seeking to increase reported claims spending must pay more to independent providers, requiring higher negotiated rates, expanded utilization, or both, each of which imposes real economic costs. A vertically integrated insurer, by contrast, can increase reported claims by marking up the internal transfer price it assigns to care delivered by owned facilities. The total cost to the integrated enterprise is zero: the higher claims

payment is offset dollar-for-dollar by higher revenue at the provider subsidiary, while the reported MLR rises and the rebate obligation falls.[2]

An important alternative interpretation is that vertical integration generates genuine efficiency improvements, including better care coordination, reduced avoidable hospitalizations, and lower administrative overhead, that increase medical spending relative to premiums through channels unrelated to gaming. Our triple-differences design distinguishes between these explanations by exploiting variation in regulatory pressure: if VI issuers' response reflects genuine efficiency gains, it should scale with gap size just as traditional insurers' does; if it reflects transfer pricing, it should be independent of gap size.

3 Data and Identification Strategy

3.1 Data

Our primary data source is the NAIC Exhibit of Premiums, Enrollment and Utilization, covering all individual comprehensive health insurance filers from 2001 through 2024. The exhibit reports premium revenue, medical claims, enrollment, ambulatory encounter rates, and inpatient admission rates at the issuer-state-year level. We restrict to the individual comprehensive market throughout, excluding Medicare supplement, dental, vision, and short-term limited duration policies.

We classify issuers as vertically integrated if their organizational parent is a health care delivery system rather than a standalone insurance company, using a combination

[2]The potential for transfer pricing to inflate reported MLR has been noted in health policy commentary (Angeles and Bailit, 2025; Argüello and Murphy, 2025). The NAIC framework does not require disclosure of intercompany transfer pricing methodology, making regulatory detection difficult. Our paper provides the first causal empirical evidence that this mechanism operates at scale.

of string-matching on NAIC company names and a manually verified issuer crosswalk, with ambiguous cases resolved by hand.[3] Our key outcome is the medical loss ratio. [4] We also construct claims and premiums per member month, claims per utilization unit (our proxy for the internal transfer price), and log enrollment as a size control. Additional controls include indicators for nonprofit status, HMO organizational form, and Blue Cross Blue Shield affiliation, and a state-level Herfindahl-Hirschman Index.

Our analysis sample restricts to issuer-state panels observed in 2010, the year immediately before the MLR provision took effect and the year in which we measure each issuer's pre-regulation compliance gap. The NAIC introduced the Qualified Improvement Expense reporting category in 2011, creating a small measurement discontinuity at the 2010–2011 boundary; year fixed effects absorb this change and our identification does not rely on MLR levels across this boundary. This drops 2,370 issuer-state-years (41%) consisting primarily of post-ACA entrants and issuers with missing 2010 filings. Dropped issuers are less VI-intensive (9.4% versus 15.2%), smaller (median enrollment 11,600 versus 20,500 life-years), and predominantly post-2014 entrants, a composition consistent with new market entry following ACA implementation rather than selective attrition of established issuers. The final sample contains 3,380 issuer-state-year observations covering 204 unique issuer-state pairs across 24 years.

3.2 Identification Strategy

Our identification exploits cross-sectional variation in issuers' pre-regulation distance from the 80 percent threshold. Following Cicala, Lieber, and Marone (2019), we

[3]We describe the process of identifying vertical integration status in our analysis sample in Appendix Table A1.

[4]The ACA calculates MLR rebate obligations using a three-year rolling average. Our analysis uses annual issuer-level MLR to capture year-to-year behavioral responses to regulatory pressure. The rolling average is a compliance mechanism rather than a measure of annual insurer behavior, and using it would attenuate our ability to detect the timing and magnitude of responses to the regulation.

define treatment intensity as:

$$dist_below80_i = \max\big(0,\ 0.80 - MLR_{i,2010}\big) \tag{1}$$

This variable equals zero for issuers already compliant in 2010 and increases with the size of the compliance gap. Because it is defined by 2010 MLR, it is time-invariant within issuer, entering the model only through interactions with post-ACA timing and VI status.

Our estimating equation is:

$$\begin{aligned} MLR_{ist} = {} & \beta_1(dist \times post)_{it} + \beta_2(dist \times post \times VI)_{it} + \beta_3(dist \times VI)_i \\ & + \beta_4(post \times VI)_{it} + \gamma_i + \lambda_t + \delta_i t + X'_{ist}\pi + \varepsilon_{ist} \end{aligned} \tag{2}$$

where γ_i are issuer-state fixed effects, λ_t are year fixed effects, and $\delta_i t$ are issuer-specific linear time trends. Standard errors are clustered at the issuer level.

The coefficient β_1 captures the traditional insurer dose-response to regulatory pressure, replicating Cicala, Lieber, and Marone in our extended panel. The triple-difference coefficient β_2 is our key parameter: a finding of $\beta_2 < 0$ means VI issuers' MLR increases less per unit of regulatory pressure than traditional issuers, the signature of a compliance mechanism decoupled from gap size. The coefficient β_4 captures the VI baseline post-ACA shift independent of distance, and β_3 captures differential pre-period MLR behavior among VI issuers with greater compliance exposure.

Our preferred specification includes issuer-specific linear trends, which absorb each issuer's own pre-existing MLR trajectory. This choice is motivated by a substantive asymmetry: VI issuers with greater regulatory exposure were on different pre-regulation MLR trajectories than traditional issuers, likely reflecting the organic

upward pressure on internal transfer prices that accompanies growing integration. Without absorbing these trajectories, the two-way fixed effects specification attributes pre-existing VI trend differences to the regulatory response—a misattribution confirmed by the sign reversal of $\hat{\beta}_2$ from $+0.315$ (insignificant) to -0.643 ($p<0.001$) when issuer trends are added. The issuer trends correction is not a robustness check but is necessary for correct inference.

The identifying assumption is that, conditional on issuer fixed effects, year fixed effects, and issuer-specific linear trends, post-ACA MLR changes would have been the same across issuers with high and low values of *dist_below*80, absent the regulation. We test this assumption using a pre-period placebo regression examining whether *dist_below*80 predicts differential MLR trends by VI status before 2011. The test passes under our continuous treatment specification ($p = 0.158$), supporting a causal interpretation of the post-2011 differential. Section 6 verifies the assumption visually through an event study that plots year-by-year coefficients separately for VI and traditional (non-integrated) issuers, normalized to the 2006–2009 pre-period mean.

Risk selection differences between VI and non-VI issuers would affect MLR levels but not the differential dose-response to regulatory pressure that our triple-differences design identifies. Our key parameter captures whether VI issuers' MLR increases less per unit of compliance gap than non-VI issuers - a comparison that is robust to time-invariant differences in enrollee health status.

4 Main Results

Vertically integrated insurers respond to MLR regulation through a compliance mechanism almost entirely decoupled from the size of their pre-regulation gap. Figure 1 provides descriptive motivation: the MLR distribution shifts rightward and bunches

just above 0.80 post-2011 for both traditional and VI issuers, confirming the regulation is binding for both groups. The mechanism through which they achieve this shift, however, differs sharply.

Table 1 presents the main results. Column 1 replicates Cicala, Lieber, and Marone (2019) in our extended 2001–2024 panel: traditional issuers closed approximately 76 percent of their pre-regulation compliance gap, a precisely estimated dose-response stable across all specifications. Column 2 adds our triple-difference terms under the baseline two-way fixed effects specification. The VI intercept is positive and significant (0.064, $p = 0.047$), indicating that VI issuers increased MLR post-ACA independent of their compliance gap, but the triple-difference coefficient is insignificant (+0.315, $p = 0.170$).

Column 3, our preferred specification, adds issuer-specific linear trends. The triple-difference coefficient reverses sign and becomes highly significant: −0.643 ($p<0.001$). The implied VI dose-response is 0.070, roughly one-tenth the traditional insurer estimate. The VI intercept strengthens to 0.083 ($p<0.001$). Columns 4 and 5 confirm robustness to adding market concentration controls and excluding the COVID year. The VI triple-difference is identified from 29 issuer clusters; the event study in Section 6 provides visual evidence that no single issuer's trajectory drives the finding, and the widening post-period gap is robust to the confidence band width implied by this sample size.

The sign reversal from +0.315 to −0.643 when issuer trends are added is not a specification detail but the central identification result. VI issuers were on systematically upward pre-regulation MLR trajectories, and the standard specification incorrectly attributes this pre-existing trend to regulatory response. Once each issuer's own trajectory is absorbed, the true pattern emerges: VI issuers deviate less from their own trend post-regulation than traditional issuers do from theirs, because

their trend was already rising. The issuer trends correction is necessary for correct inference, not merely preferred.

The two key coefficients from Column 3 together characterize two distinct compliance regimes. Traditional issuers close their compliance gaps proportionally: the further below 80 percent they started, the larger their post-ACA MLR increase, consistent with genuine and costly care expansion through arms'-length provider contracts. VI issuers increase MLR by approximately 8 percentage points post-ACA regardless of starting position, consistent with a low-cost internal adjustment that does not scale with gap size. Figure 3 illustrates this contrast: the predicted response lines cross at approximately 0.13 distance units, below which VI issuers increase MLR more than traditional issuers, and above which the traditional dose-response dominates as genuine care expansion scales with the gap while the VI response remains flat.

A third finding reinforces the transfer pricing interpretation. The pre-period coefficient on $dist \times VI$ is large, negative, and highly significant (-0.990, $p<0.001$): VI issuers with larger pre-regulation compliance gaps maintained anomalously low MLR before the ACA, consistent with strategic profit extraction through suppressed transfer prices when there was no regulatory incentive to report high claims. Figure 2 corroborates this: VI issuers show substantially more pre-ACA mass below 0.80 than their post-ACA compliance behavior would predict.

5 Mechanism Evidence

The main estimates establish that vertically integrated insurers exhibit a significantly attenuated MLR response relative to traditional insurers. Understanding why requires distinguishing among the channels through which an insurer can move its reported

MLR toward the 80 percent threshold: increasing incurred claims, inflating quality improvement (QI) expenses, reclassifying administrative costs, or expanding utilization. The transfer pricing interpretation generates a specific prediction: VI insurers should differ from traditional insurers on the claims channel but not on QI expenses or administrative costs. Because VI insurers can adjust internal prices paid to owned providers, they can increase reported claims without changing care delivery, without arms'-length contract renegotiation, and without accounting reclassification. This section tests that prediction in two steps, using utilization data from the main NAIC Exhibit (Table 2) and a full decomposition of the MLR numerator from the NAIC Supplemental Health Care Exhibit (Table 3).

A. Utilization Responses

The strongest evidence in Table 2 comes from pre-regulation behavior. Across all five outcomes—claims per member month, premiums per member month, ambulatory encounters, inpatient admissions, and claims per utilization unit—the pre-period coefficient on $dist \times VI$ is large, negative, and highly significant. VI issuers with larger pre-regulation compliance gaps maintained systematically lower claims, lower utilization, and lower premiums before the ACA, relative to their own issuer fixed effects. This pattern is consistent with strategic pre-regulation profit extraction: VI issuers suppressed internal transfer prices when there was no regulatory incentive to report high claims, then adjusted them upward post-2011. The consistency of this signature across five independent outcomes makes it difficult to attribute to sampling variation or specification choice.

Post-regulation, neither VI nor traditional issuers show significant utilization responses under the preferred specification: ambulatory encounters and inpatient ad-

missions do not increase commensurately with reported claims for either group. For traditional issuers, this is consistent with the regulatory response operating primarily through negotiated price increases with independent providers rather than volume expansion, a reading reinforced by the positive claims-per-utilization coefficient in Column 3 of Table 2. For VI issuers, the absence of a utilization response alongside a large MLR increase is the signature of the transfer pricing channel: reported claims rise through internal price adjustments rather than through changes in care delivered.[5]

B. MLR Numerator Decomposition

Table 3 decomposes the MLR numerator into its three components—incurred claims, QI expenses, and administrative costs—using the NAIC Supplemental Health Care Exhibit (SHCE), which reports these items separately starting in 2010. We estimate the triple-difference specification from equation (2) replacing the MLR outcome with each component's share of adjusted premiums. The prior literature documents two primary gaming channels available to traditional insurers facing binding MLR constraints: inflation of incurred claims through more generous provider contracts or relaxed utilization management (Abraham, Karaca-Mandic, and Simon, 2014; Cicala, Lieber, and Marone, 2019), and strategic inflation of QI expenses, which count toward the MLR numerator and exhibit discontinuous increases at the regulatory threshold (Born, Sirmans, and Steinorth, 2023). A third channel, reclassification of administrative expenses, has been documented in related regulatory contexts (Eastman, Eckles, and Van Buskirk, 2021). We test whether VI issuers' attenuated response operates

[5]Several post-regulation triple-difference coefficients attenuate under the preferred specification relative to the two-way fixed effects specification (Appendix Table A2), consistent with the differential pre-existing trends that motivate the issuer trends correction. The NAIC Exhibit does not separately identify quality improvement expenses, which prior work has documented as a distinct gaming channel (Born, Sirmans, and Steinorth, 2023); we address this in the next subsection.

differentially through any of these three channels.

The results point unambiguously to the claims channel. The triple-difference coefficient on claims share (−1.061, SE 0.095, $p<0.01$, Column 2) is nearly identical in magnitude to the coefficient on overall MLR (−1.052, SE 0.091, $p<0.01$, Column 1), indicating that the full attenuation of VI issuers' gaming response runs through claims rather than through other components. The implied traditional insurer dose-response slope on claims share is 1.149, meaning a 10 percentage point shortfall below the 80 percent threshold is associated with an 11.5 percentage point increase in claims share among traditional issuers. Among VI issuers, the net slope is $1.149 - 1.061 = 0.088$, roughly one-twelfth as large. By contrast, the triple-difference coefficients on QI share (0.008, $p = 0.545$, Column 3) and admin share (−0.014, $p = 0.923$, Column 4) are small and statistically indistinguishable from zero.

These findings have a natural interpretation within the vertical integration framework. Traditional insurers facing binding MLR constraints must rely on arms'-length strategies—renegotiating provider contracts, relaxing utilization management, or inflating QI expense classifications—to move their reported MLR upward. VI insurers, by contrast, can adjust internal transfer prices paid to owned providers, directly increasing reported claims without requiring external negotiation or accounting reclassification. The absence of a differential QI or admin response among VI issuers is precisely what this mechanism predicts: if the transfer pricing channel is available and low-cost, VI insurers need not resort to the QI expense inflation or administrative reclassification strategies documented among traditional insurers in the prior literature (Born, Sirmans, and Steinorth, 2023; Eastman, Eckles, and Van Buskirk, 2021). In other words, vertical integration expands the regulatory response toolkit in ways that shift gaming behavior toward less-visible channels, specifically inflated internal payments that satisfy MLR compliance on paper while retaining revenue within the

corporate family.

The SHCE sample is restricted to issuers appearing in the main NAIC panel, yielding 3,762 issuer-state-year observations with a single pre-regulation year (2010). Because parallel trends cannot be formally tested in this limited pre-period, the SHCE results should be interpreted as decomposing the mechanism underlying the main panel estimates rather than as independent causal evidence. The directional consistency between the SHCE MLR coefficient (-1.052) and the preferred main panel estimate (-0.643, Table 1, Column 3) confirms that the SHCE subsample captures the same underlying phenomenon; the larger SHCE magnitude reflects selection of larger, more established filers into the SHCE reporting universe.

A second limitation concerns the precision with which we can characterize the transfer pricing mechanism. While our SHCE decomposition confirms the differential response runs entirely through claims rather than quality improvement or administrative costs, we cannot directly observe internal transfer prices between integrated insurers and their owned provider networks. The extent to which regulatory, tax, or internal governance constraints limit transfer pricing in practice, and whether these constraints vary across states or organizational structures, is an important question for future regulatory and empirical work.

6 Event Study

Figure 4 plots year-by-year coefficients separately for VI and traditional issuers, normalized to the 2006–2009 pre-period mean, and reveals two distinct regulatory dynamics.[6]

[6]We use 2009 as the omitted baseline year rather than 2010 to avoid Ashenfelter's dip: treatment intensity is defined by 2010 MLR, making 2010 a local trough for treated issuers driven by transitory claims variation rather than reflecting structural compliance levels. Normalizing to the 2006–2009 mean provides a more stable visual baseline.

Traditional issuers show flat pre-period coefficients clustered around zero through 2009, supporting the parallel trends assumption visually. Post-2011, their response builds gradually to approximately 1.0 per unit of regulatory pressure by 2015–2016 and remains stable through 2024. The persistence of this response over thirteen post-regulation years extends the temporal scope of Cicala, Lieber, and Marone's finding and confirms that the individual market did not revert toward pre-regulation MLR levels as the regulation matured.

VI issuers tell a different story. Their post-2011 response is larger than traditional issuers from the outset and accelerates over time, reaching approximately 2.0–2.5 per unit of regulatory pressure by 2020–2024. Whether this acceleration reflects growing VI market share, cumulative ratcheting of internal transfer prices, or post-2014 Marketplace expansion is a question we leave for future work. More importantly, the gap between VI and traditional issuers widens persistently rather than converging. Simple regulatory compliance would predict convergence: once both groups reach 80 percent MLR, further increases should slow and the gap should close. The widening gap is inconsistent with this prediction and consistent instead with compounding gaming: as vertical integration deepens and transfer pricing arrangements mature, VI issuers extract increasing value from the regulation relative to their traditional counterparts.

The pre-period VI upward trajectory, visible in Figure 4 from 2007 onward, is the differential pre-trend that motivates the issuer-specific trends specification in Section 3. The preferred specification absorbs this trajectory by construction, identifying the post-regulation response as a deviation from each issuer's own trend rather than from a common baseline.[7]

[7]The VI event study is estimated from 29 issuer clusters; pre-period VI confidence intervals are suppressed in Figure 4 due to their width in this subsample. The widening post-period gap is visually robust to the confidence band width.

7 Robustness

Table 4 stress-tests the main findings across twelve specifications. The traditional insurer dose-response is stable throughout, ranging from 0.51 to 1.01 and significant at the 1 percent level in every column, confirming that genuine care expansion in response to regulatory pressure is a robust feature of the individual market. The VI triple-difference result is equally stable under the preferred issuer trends specification and reveals important pre-existing trend differences under the baseline specification.

Under the preferred specification (Table 5), the key finding holds across the two most demanding checks. Excluding the COVID year yields -0.617 ($p = 0.001$) and restricting to pre-2014 entrants yields -0.643 ($p<0.001$), identical to the full-sample benchmark. The VI intercept is significant in all four Table 5 checks, ranging from 0.076 to 0.100.

The large-issuer and balanced-panel checks (Columns 1 and 2) show attenuation toward zero under the preferred specification rather than the large positive values that appear under the baseline two-way FE specification. This contrast directly confirms the identification argument in Section 3: larger and longer-tenured issuers are disproportionately VI and disproportionately characterized by pre-existing upward MLR trajectories that the simpler specification cannot absorb. The issuer trends correction eliminates this bias; attenuation toward zero, rather than the sign reversal seen in Table 4, is the expected pattern once pre-existing trends are controlled. The 3+ pre-period years check shows the same pattern.

Two additional checks address alternative explanations. Two-way clustering by issuer and state leaves the point estimate unchanged while producing modestly larger standard errors, confirming that within-state correlation does not drive our inference. Replacing the 80 percent threshold with 85 percent yields a traditional insurer dose-

response of 0.706 ($p<0.001$), confirming that the findings are not an artifact of the specific regulatory threshold.

8 Conclusion

Whether regulation achieves genuine behavioral change or merely accounting compliance depends on the tools available to regulated firms. In health insurance markets, vertical integration creates access to internal transfer pricing, a compliance mechanism unavailable to traditional competitors and largely invisible to regulators. We refer to this as a compliance technology: a mechanism through which a firm satisfies a regulatory requirement without undertaking the behavioral change the requirement was designed to induce. This paper provides the first causal evidence that this technology is used at scale in the US health insurance market, and that it fundamentally alters the welfare consequences of a major federal consumer protection.

Vertically integrated insurers respond to the ACA's 80 percent medical loss ratio requirement through a mechanism almost entirely decoupled from the size of their compliance gap. Their implied dose-response is one-tenth that of traditional competitors, they increased MLR by approximately the same amount regardless of starting position, and they maintained anomalously low claims spending before the regulation, a pattern consistent with suppressed transfer prices when there was no regulatory incentive to report high claims. The event study reveals that the gap between VI and traditional insurer regulatory responses widens persistently through 2024 rather than converging, consistent with gaming that deepens as integration expands. Decomposing the MLR numerator confirms that this differential response runs entirely through the claims channel, with precisely estimated zeros on quality improvement expenses and administrative costs, ruling out the accounting reclassification strategies avail-

able to traditional insurers and identifying internal transfer pricing as the operative mechanism. Together these findings characterize two distinct compliance regimes operating simultaneously within the same regulated market.

These results change how aggregate MLR compliance statistics should be interpreted. Individual market insurers have been nearly universally compliant with the 80 percent threshold since 2013 (Cicala, Lieber, and Marone, 2019), and total rebates have exceeded $13 billion since 2012 (Ortaliza and Cox, 2024), figures sometimes cited as evidence that the provision is working as intended. Our findings suggest this interpretation overstates the regulation's consumer protection: VI issuers achieve compliance through internal accounting rather than genuine medical spending, retaining revenue within the corporate family rather than reallocating it to enrollees as the regulation intends. The MLR formula relies on insurer-reported claims data without independent verification of whether payments to owned providers reflect arms'-length prices, a gap that our results suggest is being systematically exploited. As vertical integration in health insurance continues to expand, aggregate MLR compliance statistics are likely to become an increasingly imperfect indicator of genuine consumer protection.

These results also add a regulatory arbitrage channel to the welfare calculus of vertical integration in health care markets. Prior work has documented mixed effects of insurer-provider integration on premiums, quality, and utilization. Our findings suggest that even where integration generates genuine delivery efficiency gains, it simultaneously undermines spending mandates designed to protect consumers by giving integrated firms access to compliance mechanisms that do not require genuine reallocation toward patient care. The net welfare effect of integration therefore depends not only on delivery efficiency but on the regulatory environment in which integrated firms operate, a consideration most existing research overlooks.

One limitation bears noting: the NAIC data do not allow us to separately identify transfer pricing inflation from genuine utilization increases at owned facilities. Resolving this distinction would require data linking insurer payments to owned providers with independent measures of care delivered—a direction we leave for future work. Taken together, these results illustrate a broader principle: when regulatory rules are imperfect, organizational structure determines which firms can satisfy them on paper while preserving profits. Vertical integration in health insurance is a case study in this dynamic, one likely to grow in relevance as integration expands across health care markets.

References

Abraham, Jean M., Pinar Karaca-Mandic, and Kosali Simon. 2014. "How Has the Affordable Care Act's Medical Loss Ratio Regulation Affected Insurer Behavior?" *Medical Care* 52 (4): 370–377.

Angeles, January, and Michael Bailit. 2025. "How Insurers That Own Providers Can Game the Medical Loss Ratio Rules." *Health Affairs Forefront*, September 29. https://www.healthaffairs.org/content/forefront/insurers-own-providers-can-game-medical-loss-ratio-rules.

Argüello, Andrés, and Natasha Murphy. 2025. "Medical Loss Ratio Reform Can Help Curb Corporate Power and Lower Health Care Costs." Center for American Progress, October 6. https://www.americanprogress.org/article/medical-loss-ratio-reform-can-help-curb-corporate-power-and-lower-health-care-cost

Born, Patricia H., E. Tice Sirmans, and Petra Steinorth. 2023. "Health Insurers'

Use of Quality Improvement Expenses to Achieve a Minimum Medical Loss Ratio Requirement." *Journal of Risk and Insurance* 90 (1): 123–154.

Callaghan, S. Reed, Elizabeth Plummer, and William F. Wempe. 2020. "Health Insurers' Claims and Premiums Under the Affordable Care Act: Evidence on the Effects of Bright Line Regulations." *Journal of Risk and Insurance* 87 (1): 67–93.

Cicala, Steve, Ethan M. L. Lieber, and Victoria Marone. 2019. "Regulating Markups in US Health Insurance." *American Economic Journal: Applied Economics* 11 (4): 71–104.

Dafny, Leemore, Kate Ho, and Robin S. Lee. 2019. "The Price Effects of Cross-Market Mergers: Theory and Evidence from the Hospital Industry." *RAND Journal of Economics* 50 (2): 286–325.

Dafny, Leemore. 2005. "How Do Hospitals Respond to Price Changes?" *American Economic Review* 95 (5): 1525–1547.

Eastman, Evan M., David L. Eckles, and Andrew Van Buskirk. 2021. "Accounting-Based Regulation: Evidence from Health Insurers and the Affordable Care Act." *The Accounting Review* 96 (2): 231–259.

Gaynor, Martin, Kate Ho, and Robert J. Town. 2015. "The Industrial Organization of Health-Care Markets." *Journal of Economic Literature* 53 (2): 235–284.

Gupta, Atul, Ambar La Forgia, and Adam Sacarny. Forthcoming. "Turbocharging Profits? Contract Gaming and Revenue Allocation in Healthcare." *American Economic Journal: Applied Economics*.

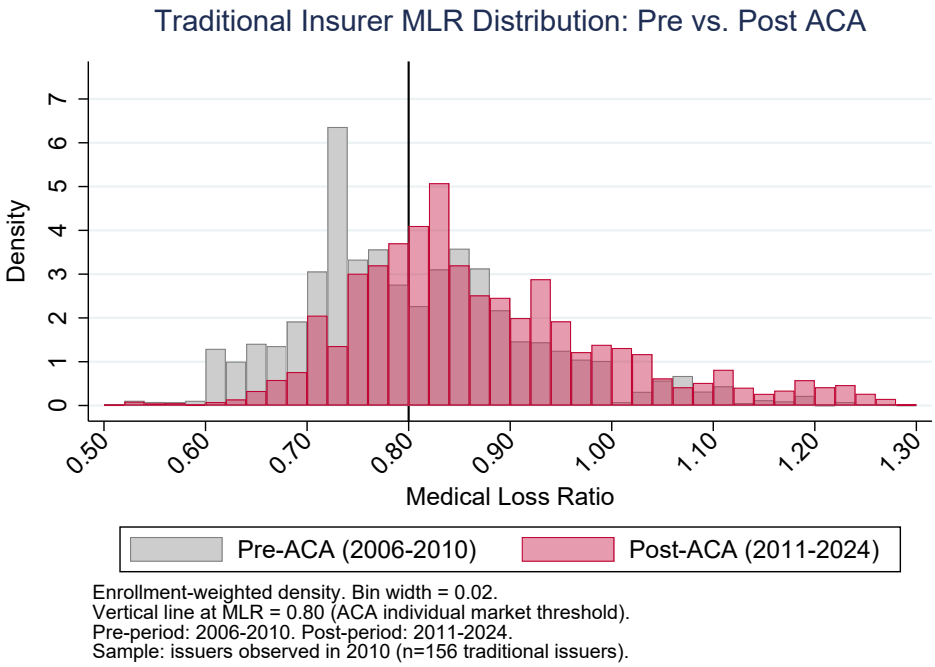
Karaca-Mandic, Pinar, Jean M. Abraham, and Kosali Simon. 2015. "Is the Medical

Loss Ratio a Good Target Measure for Regulation in the Individual Market for Health Insurance?" *Health Economics* 24 (1): 55–74.

Koch, Thomas G., Brett W. Wendling, and Nathan E. Wilson. 2021. "How Vertical Integration Affects the Quantity and Cost of Care for Medicare Beneficiaries." *Journal of Health Economics* 80: 102536.

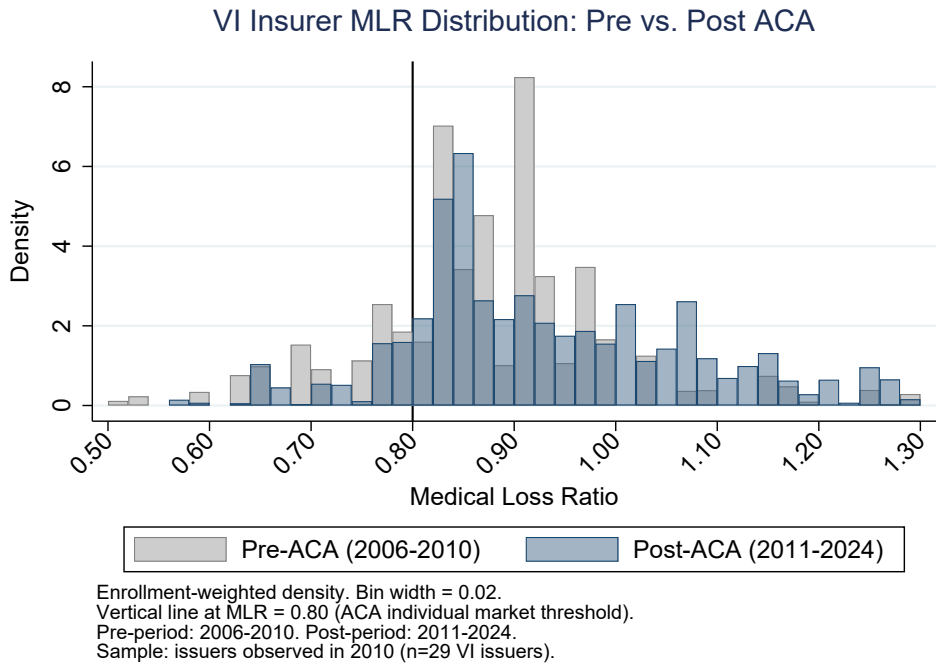
Ortaliza, Jared, and Cynthia Cox. 2024. "2024 Medical Loss Ratio Rebates." KFF, June 5. https://www.kff.org/private-insurance/medical-loss-ratio-rebates/.

Figure 1: MLR Distribution, Pre- and Post-ACA — Traditional Insurers



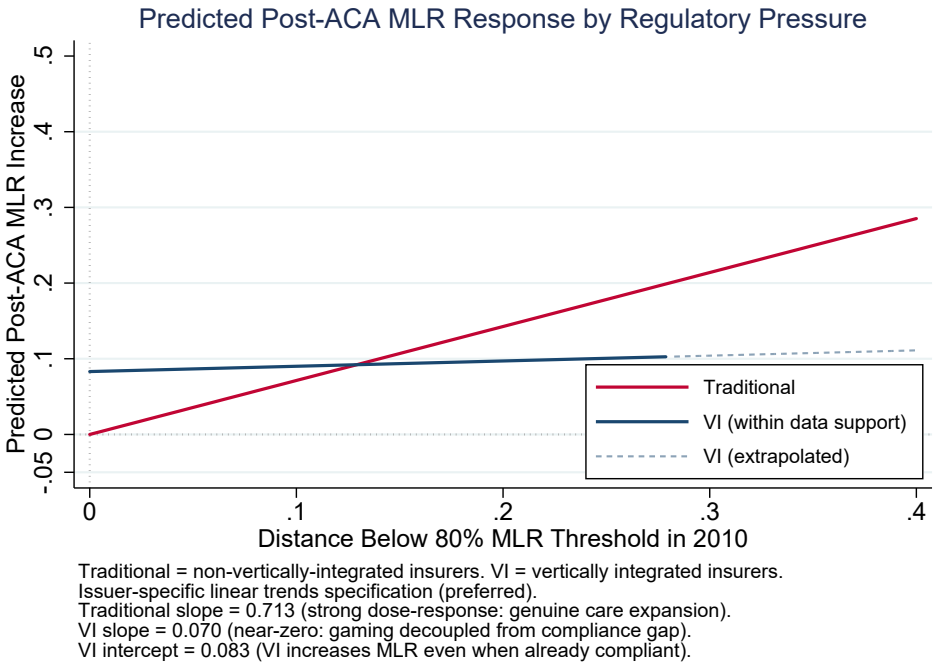
Notes: Distribution of medical loss ratios for traditional (non-integrated) insurers, individual market, 2001–2024. Pre-ACA period: 2001–2010. Post-ACA period: 2011–2024. The vertical dashed line marks the 80 percent regulatory threshold.

Figure 2: MLR Distribution, Pre- and Post-ACA — Vertically Integrated Insurers



Notes: Distribution of medical loss ratios for vertically integrated insurers, individual market, 2001–2024. Pre-ACA period: 2001–2010. Post-ACA period: 2011–2024. The vertical dashed line marks the 80 percent regulatory threshold.

Figure 3: Predicted MLR Response by Vertical Integration Status and Pre-Regulation Compliance Gap



Notes: Predicted post-ACA MLR increase as a function of pre-regulation distance from the 80 percent threshold, separately for traditional and VI issuers, based on Column 3 coefficients from Table 1. Lines cross at approximately 0.13 distance units.

Table 1: MLR Gaming by Vertical Integration Status

	(1) Baseline DiD	(2) Two-Way FE	**(3) Preferred**	(4) +HHI	(5) Excl. 2020
Distance × Post	0.759***	0.791***	0.713***	0.712***	0.717***
(trad. baseline; replicates Cicala)	(0.159)	(0.169)	(0.160)	(0.159)	(0.160)
Distance × Post × VI		0.315	−0.643***	−0.648***	−0.617***
(triple-diff: VI gaming)		(0.229)	(0.175)	(0.177)	(0.182)
Distance × VI		−0.925***	−0.990***	−0.994***	−0.957***
(pre-period VI selection)		(0.231)	(0.140)	(0.140)	(0.141)
Post × VI		0.064**	0.083***	0.083***	0.084***
(VI baseline shift)		(0.032)	(0.023)	(0.023)	(0.023)
State market HHI				0.011	
				(0.034)	
Log enrollment	−0.000	−0.001	0.016*	0.016*	0.017*
	(0.009)	(0.008)	(0.008)	(0.008)	(0.009)
Issuer-State FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Issuer Linear Trends	No	No	Yes	Yes	Yes
Sample	All	All	All	All, HHI ctrl	Excl. 2020
N	3,380	3,380	3,380	3,380	3,261
R^2	0.460	0.464	0.605	0.605	0.611

Notes: Clustered SEs at issuer level in parentheses. Treatment intensity = $\max(0, 0.80 - MLR_{2010})$; zero if already compliant. Cols. 1–2: two-way fixed effects only. Cols. 3–5: issuer-specific linear trends. Col. 3 is the preferred specification; Cols. 4–5 are robustness checks. Issuer linear trends absorb differential pre-existing MLR trajectories between VI and traditional issuers; continuous treatment pre-trend test $p = 0.158$. *dist_X_post* = traditional insurer dose-response (replicates Cicala, Lieber, and Marone 2019). *dist_X_post_X_vi* = VI triple-diff (key coefficient). *dist_X_vi* = VI pre-period selection (strategic pre-ACA suppression). *post_X_vi* = VI baseline post-ACA shift independent of distance. Implied VI slope (Col. 3): $0.713 + (-0.643) = 0.070$ (90% smaller than traditional issuers).
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2: Mechanism Tests — Utilization and Price Responses

	(1) Amb. Enc./LY	(2) Admits/LY	(3) Claims/Util	(4) Claims PMPM	(5) Premiums PMPM
Distance × Post	2.570	0.017	3,555.051**	284.203**	139.240
(traditional baseline)	(3.677)	(0.020)	(1,598.093)	(110.688)	(123.307)
Distance × Post × VI	−3.636	−0.034	−2,819.991	−475.373*	−285.160
(triple-diff: VI gaming)	(4.467)	(0.032)	(2,258.870)	(284.121)	(322.785)
Post × VI	0.491	0.002	−32.968	45.031	34.919
(VI baseline shift)	(0.744)	(0.004)	(434.864)	(50.968)	(61.812)
Distance × VI	−16.243***	−0.203***	1,622.666*	−472.433***	−362.729***
(pre-period VI selection)	(2.981)	(0.020)	(902.380)	(129.055)	(132.686)
Issuer-State FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Issuer Linear Trends	Yes	Yes	Yes	Yes	Yes
N	2,888	2,863	3,232	3,380	3,380
R^2	0.772	0.712	0.700	0.834	0.829

Notes: All regressions use the preferred issuer-specific linear trends specification. Clustered SEs at issuer level in parentheses. Appendix Table A2 replicates under panel + year FE; mechanism results are sensitive to specification as expected given differential pre-trends. Outcome units: encounters and admits per life-year; claims and premiums in dollars per member month; claims/util in dollars per utilization unit (inpatient days + total encounters).

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

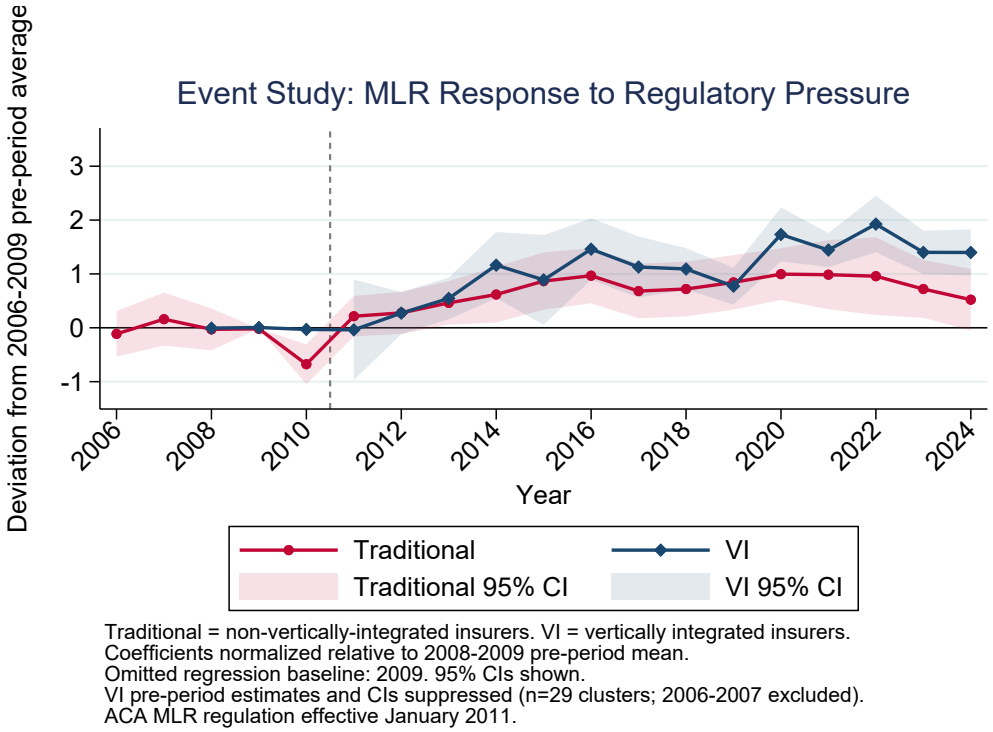
Preliminary Draft: Please Do Not Circulate

Table 3: Mechanism Analysis — MLR Numerator Decomposition by VI Status

	(1) MLR	(2) Claims Share	(3) QI Share	(4) Admin Share
Distance × Post	1.154***	1.149***	0.004**	−0.045
(traditional baseline)	(0.072)	(0.071)	(0.002)	(0.052)
Distance × VI	0.989***	1.000***	−0.010	0.135
(pre-period selection)	(0.044)	(0.053)	(0.014)	(0.144)
Post × VI	0.031*	0.033*	−0.003	0.019
(VI baseline shift)	(0.018)	(0.019)	(0.003)	(0.032)
Distance × Post × VI	−1.052***	−1.061***	0.008	−0.014
(triple-diff)	(0.091)	(0.095)	(0.014)	(0.147)
Observations	3,762	3,762	3,762	3,762

Notes: Unit of observation is issuer-state-year, 2010–2024. Sample restricted to issuers appearing in main NAIC panel (N = 3,762 after dropping singletons). Claims share = incurred claims / adjusted premiums (SHCE line 5.0/1.8). QI share = QI expenses / adjusted premiums (SHCE lines 6.3 or 6.6/1.8). Admin share = (claims adj. exp. + G&A) / adjusted premiums (SHCE lines 8.3+10.5/1.8). $dist = 0.80 - MLR$ (distance to regulatory threshold), winsorized at 1/99. All specifications include issuer-state FE, year FE, and issuer linear trends. *flag_hmo* omitted due to collinearity with issuer-state fixed effects. Standard errors clustered by issuer.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure 4: Event Study — Year-by-Year MLR Coefficients by VI Status, 2001–2024



Notes: Year-by-year triple-difference coefficients from equation (2), estimated separately for traditional (non-integrated) and vertically integrated issuers. Coefficients normalized to the 2006–2009 pre-period mean. Omitted baseline year: 2009. Pre-period VI confidence intervals suppressed due to width (29 VI issuer clusters). Shaded bands indicate 95 percent confidence intervals.

Table 4: Robustness Checks — Two-Way Fixed Effects

	(1) Large	(2) Balanced	(3) No COVID	(4) No 2014–15	(5) 2-Way Clust.	(6) 85% Thresh.	(7) 3+ Pre-yrs
Distance × Post	0.514***	1.007***	0.783***	0.787***	0.791***		0.661***
(traditional baseline)	(0.168)	(0.223)	(0.170)	(0.170)	(0.167)		(0.178)
Distance × Post × VI	2.483***	2.309***	0.248	0.342	0.315*		2.589***
(triple-diff: VI gaming)	(0.422)	(0.686)	(0.221)	(0.222)	(0.185)		(0.679)
Post × VI	0.020	0.086**	0.068**	0.068**	0.064**	0.077**	0.063*
(VI baseline shift)	(0.022)	(0.038)	(0.032)	(0.032)	(0.025)	(0.034)	(0.036)
dist85 × Post						0.706***	
						(0.138)	
dist85 × Post × VI						0.217	
						(0.224)	
Sample	≥10K lives	Balanced	Excl. 2020	Excl. 2014–15	2-way cluster	Thresh. 85%	3+ pre-yrs
Issuer Trends	No	No	No	No	No	No	No
N	2,308	2,217	3,261	3,058	3,380	3,380	2,938
R^2	0.561	0.441	0.468	0.468	0.464	0.468	0.454

Notes: All regressions include issuer-state FE and year FE. Clustered SEs at issuer level except Col. 5 (issuer × state). Cols. 1–5, 7: threshold = 80%. Col. 6: threshold = 85% (*dist85* variables shown). Large positive *dist_X_post_X_vi* in Cols. 1–2 and 7 reflects differential pre-existing trends among VI issuers in these subsamples; see Table 5.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Preliminary Draft: Please Do Not Circulate

Table 5: Robustness Checks — Preferred Specification (Issuer Trends)

	(1) Large	(2) Balanced	(3) No COVID	(4) Pre-2014	(5) 3+ Pre-yrs
Distance × Post	0.619***	0.914***	0.717***	0.713***	0.708***
(traditional baseline)	(0.150)	(0.241)	(0.160)	(0.160)	(0.198)
Distance × Post × VI	0.815	0.848	−0.617***	−0.643***	0.755
(triple-diff: VI gaming)	(0.613)	(0.592)	(0.182)	(0.175)	(0.551)
Post × VI	0.076**	0.100***	0.084***	0.083***	0.085***
(VI baseline shift)	(0.030)	(0.024)	(0.023)	(0.023)	(0.023)
Sample	≥10K lives	Balanced	Excl. 2020	Pre-2014	3+ pre-yrs
Issuer Trends	Yes	Yes	Yes	Yes	Yes
N	2,308	2,217	3,261	3,380	2,938
R^2	0.680	0.575	0.611	0.605	0.589

Notes: All regressions include issuer-state FE, year FE, and issuer-specific linear trends. Clustered SEs at issuer level. Benchmark (full sample, Col. 3 of Table 1): *dist_X_post* = 0.713***, *dist_X_post_X_vi* = −0.643***. Cols. 1–2: attenuation toward zero (vs. large positive in Table 4) confirms issuer trends absorb differential pre-existing trajectories in these subsamples. Cols. 3–5: key result stable and significant.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix

Appendix Table A1: Vertical Integration Classification Procedure

We identify vertically integrated issuers using a two-step procedure applied to NAIC company names. In the first step, we construct a standardized company name string and apply string-matching rules to flag issuers affiliated with well-documented provider-sponsored health plan systems, including large integrated delivery networks and hospital-system-affiliated plans operating in the individual market. Several rules include state restrictions to avoid false positives from common names. In the second step, we supplement the automated string-matching with a manually verified issuer crosswalk that codes additional issuers not captured by the automated rules, with manual classifications taking precedence where they overlap. We classify an issuer as VI if its organizational parent is a health care delivery system rather than a standalone insurance company.

Appendix Table A2: Mechanism Analysis — MLR Numerator Decomposition by VI Status (Two-Way Fixed Effects)

	(1) MLR	(2) Claims Share	(3) QI Share	(4) Admin Share
Distance × Post	1.209***	1.206***	0.002	−0.003
(traditional insurer baseline)	(0.075)	(0.075)	(0.002)	(0.075)
Distance × VI	0.989***	1.007***	−0.017***	−0.030
(pre-period selection)	(0.056)	(0.059)	(0.006)	(0.134)
Post × VI	0.022	0.027	−0.005***	0.016
(VI baseline shift)	(0.021)	(0.021)	(0.002)	(0.028)
Distance × Post × VI	−0.997***	−1.013***	0.017**	0.114
(triple-diff)	(0.095)	(0.097)	(0.007)	(0.151)
Observations	3,762	3,762	3,762	3,762

Notes: Unit of observation is issuer-state-year, 2010–2024. Sample restricted to issuers appearing in main NAIC panel (N = 3,762 after dropping singletons). Claims share = incurred claims / adjusted premiums (SHCE line 5.0/1.8). QI share = QI expenses / adjusted premiums (SHCE lines 6.3 or 6.6/1.8). Admin share = (claims adj. exp. + G&A) / adjusted premiums (SHCE lines 8.3+10.5/1.8). $dist = 0.80 - MLR$ (distance to regulatory threshold), winsorized at 1/99. All specifications include issuer-state FE and year FE. No issuer linear trends. *flag_hmo* omitted due to collinearity with issuer-state fixed effects. Standard errors clustered by issuer.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$