Vertical Integration and Regulatory Compliance in U.S. Health Insurance Markets

Rui Li *

Abstract

I analyze how vertical integration between insurers and providers affects the compliance with Medical Loss Ratio regulation and the consumer welfare in the U.S. health insurance market, using a rich dataset from 2014 to 2024. Reduced-form evidence suggests that vertical integration raises the medical loss ratios of their subsidiary insurers, primarily through higher claims expenditures rather than changes in premiums or quality-improvement expenditures. Motivated by these reduced-form results, I propose and estimate a model of insurer–provider bargaining and find that removing vertical integration leads to lower negotiated prices and premiums, improving consumer welfare. In contrast, stricter regulations raise both prices and total spending. Therefore, vertical integration can weaken the effectiveness of Medical Loss Ratio regulation and reduce its intended consumer welfare.

Keywords: Vertical integration, Medical Loss Ratio, Bargaining

JEL Codes: I11, L10, L51.

*Department of Economics, University at Albany, State University of New York, Albany, NY 12222. Email: kli7@albany.edu. **Acknowledgements:** The authors received no financial or other research support for this study and declare no financial relationships or other potential conflicts of interest.

1 Introduction

Vertical integration (VI) between the insurers and providers has become a significant concern for regulators in U.S. health insurance markets (FierceHealthcare, 2025). VI occurs when an insurer owns or controls multiple parts of the healthcare supply chain, such as physician practices, clinics, or hospitals. As of January 2024, more than three-fourths of all U.S. doctors are now employed by hospitals, health insurers, private equity or other corporate entities (HealthCareDive, 2024). Fein (2025) shows that many major insurers[1] are not stand-alone insurers but are embedded in vertically integrated networks, controlling multiple segments (insurance, drug management, provider delivery). For example, UnitedHealth Group, the parent company of the nation's largest private insurer, UnitedHealthcare (UHC), employs approximately 10 percent of the active U.S. physician workforce, delivering primary, specialty, urgent, and surgical care through Optum Health (Weber, 2024). Despite the prevalence of VI between the insurers and providers, there is limited evidence quantifying its impact.

The rapid expansion of insurers has raised concerns about competition and access. In 2011, the Medical Loss Ratio(MLR) regulation, introduced under the Affordable Care Act, required insurance companies to spend at least 80% or 85% of premium revenue on medical care or quality improvement.[2] Noncompliant insurers who do not meet these thresholds should issue rebates to policyholders each year. The purpose of the MLR regulation is to promote greater transparency and accountability and to ensure that policyholders receive value for their premium dollars.[3] This regulation has been binding for many insurers, resulting in a total of $11.8 billion in rebates issued to individuals and employers from 2012 through 2022 (Norris, 2024).

[1] such as UnitedHealth Group, CVS/Aetna, Cigna, Elevance, and Humana

[2] Individual and small-group insurers must spend at least 80% of premiums on medical expenses, while large-group insurers must spend at least 85%.

[3] Department of Health and Human Services, "Health Insurance Issuers Implementing Medical Loss Ratio (MLR) Requirements Under the Patient Protection and Affordable Care Act; Interim Final Rule," Federal Register, 45 C.F.R. Part 158, December 1, 2010, pp. 74864 -74934; https://www.federalregister.gov/articles/2010/12/01/2010-29596/health-insurance-issuers-implementing-medical-loss-ratio-mlr-requirements-under-the-patient.

However, insurers that are vertically integrated with healthcare providers may use integration to reduce the effective impact of the MLR regulation and improve their reported compliance. In several states in 2023, there was a significant increase in payments that were not tied to specific medical claims, particularly in Medicare Advantage (Angeles and Bailit, 2025). Frank and Milhaupt (2022) find that among large insurers operating across multiple markets, internal transactions with related businesses can account for roughly 20% to 71% of reported medical spending. Under a vertically integrated structure, payments to affiliated providers are counted as medical spending in the insurer's MLR calculation. This creates an incentive for insurers to shift spending toward their affiliated providers, potentially inflating reported medical loss ratios (MLRs) without actually improving care quality or delivering additional services.

In this paper, I examine the equilibrium effects of VI between insurers and providers under the MLR regulation in the U.S. private health insurance market. Using Form 10-K filings from the SEC's EDGAR database and the National Plan and Provider Enumeration System (NPPES), I construct ownership links between providers and the four largest parent insurance companies from 2014 to 2024. Combining these data with MLR reports from the Center for Consumer Information and Insurance Oversight (CCIIO) in the individual market, I present reduced-form evidence that VI between parent insurers and providers increases the reported MLRs of their subsidiary insurers. In particular, this increase is primarily associated with higher reported claims expenditures rather than changes in premiums or quality improvement expenditures. The results are robust to alternative measures of the scale and intensity of provider ownership and to an expanded control group. The results are consistent with internal transfers between subsidiary insurers and affiliated providers that raise reported medical spending without necessarily improving efficiency. Motivated by this evidence, I develop a structural model to examine how VI affects market outcomes, including negotiated service prices, premiums, and consumer welfare, under the MLR constraint.

To quantify the internal transfer effects of VI, I develop a structural model in which

insurers negotiate service prices with providers and set premiums under VI and the MLR regulation. On the demand side, I use a discrete choice model with random coefficients following Berry et al. (1993) to capture how consumers choose among insurance plans based on premiums and cost-sharing characteristics. On the supply side, insurers maximize profits subject to the MLR constraint and bargain with providers over service prices. VI between parent insurers and affiliated providers introduces an internalization parameter that measures the extent to which the insurer values the provider's profit in its objective function. Affiliated providers also account for their own profit and a symmetric internalization share of the insurer's profit. This framework connects insurer pricing, provider bargaining, and regulatory compliance within the health insurance market.

I estimate the demand model using individual plan-level data from Healthcare.gov's ACA Health Insurance Exchange Marketplace. After estimating the demand for health insurance plans, I estimate how VI and the MLR regulation affect the internalization parameter, insurer bargaining power, non-compliance costs, and fixed costs. The counterfactual analysis show that removing VI lowers both negotiated provider prices and premiums, thereby improving consumer welfare. In contrast, stricter MLR threshold from 80% to 90% increases prices and reduces welfare, as insurers expand claims to maintain compliance. These results suggest that VI helps insurers meet MLR requirements without improving actual efficiency. Instead, insurers rely on internal cost shifting and accounting adjustment that inflate reported medical spending. Overall, VI weakens the effectiveness of MLR regulation and reduces its intended consumer protection benefits.

My paper contributes to the literature on VI and healthcare pricing. Existing studies have mainly examined VI between hospitals and physicians (Baker et al., 2016; Capps et al., 2018; Koch et al., 2021; Encinosa and Dor, 2025) or between insurers and pharmacy benefit manager (Brot-Goldberg et al., 2022; Gray et al., 2023). While these studies provide valuable insights into how consolidation affects prices and market power, there is relatively little research on VI directly between insurers and healthcare providers. My paper fills this gap by

quantifying how insurer–provider integration affects market outcomes in healthcare markets in U.S. healthcare markets.

This paper also contributes to the literature on how the MLR regulation influences insurer pricing behavior. Previous studies have examined how binding MLR constraints shape insurers' strategic responses in pricing and cost management. Harrington (2013) discusses how MLR regulation can distort pricing incentives by encouraging higher upfront premiums, reducing product innovation, and increasing consolidation in health insurance markets. Abraham et al. (2014) provide early evidence that insurers with MLRs more than 10% below the threshold showed the largest subsequent increases in MLRs for compliance, mainly through higher medical claims and lower administrative expenses. McCue et al. (2013) and McCue and Hall (2015) have found that MLRs rose significantly in the individual market, especially among for-profit insurers. Cicala et al. (2019) shows that insurers subject to MLR regulation tend to raise medical claims costs rather than reduce premiums, highlighting how the rule alters incentives in the insurance market. While these studies shed light on how MLR regulation affects insurer behavior and pricing, few have explored how VI interacts with these regulatory constraints. My paper fills this gap by quantifying how insurer–provider integration influences pricing, compliance, and welfare under the MLR regulation.

Several recent studies are closely related to my work. Zhao (2021) develops a structural model of insurer–provider bargaining under the MLR constraint and shows that the MLR regulation lead to higher health service prices and consumer welfare loss. However, due to data limitations, this study does not provide reduced-form evidence on the effects of MLR regulation. Cuesta et al. (2024) examine VI between insurers and hospitals using administrative data from Chile and demonstrate that VI distorts plan networks, raises healthcare spending, and lowers welfare. Their analysis does not consider the role of regulation in shaping insurers' behavior under VI. More recently, Angeles and Bailit (2025) discuss how insurers that are vertically integrated with providers may inflate medical service prices to game MLR rules, raising policy concerns about the regulation's effectiveness. However,

they do not provide direct empirical evidence or quantify the welfare outcomes to support these concerns. My paper builds on these studies in two ways. First, I use a difference-in-differences(DID) approach to empirically examine how insurer–provider integration affects MLR compliance. Second, I develop a structural model that incorporates both VI and the MLR constraint. This framework allows me to measure how VI influences pricing, compliance, and welfare, providing new evidence on the interaction between market structure and regulatory incentives in the U.S. health insurance market.

The remainder of the paper is organized as follows. Section 2 provides institutional background on the MLR regulation and the recent trend of insurer-provider VI. Section 3 describes the data sources and summary statistics. Section 4 presents the reduced-form analysis for how VI affect the MLRs. Section 5 introduces the model. Section 6 discusses the structural analysis. Finally, Section 7 concludes.

2 Industry Background

2.1 The Medical Loss Ratio Regulation

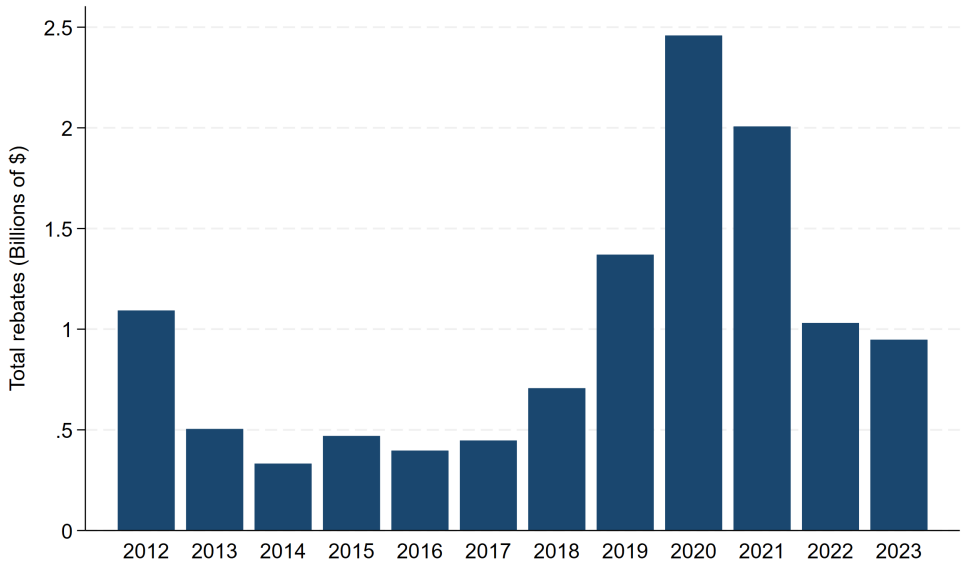
The MLR regulation was introduced under the Affordable Care Act (ACA) of 2010 and took effect in 2011 as a federal regulation. Under this rule, insurers are required to spend at least 80 percent of premium revenue on health care claims and quality improvement in the individual and small-group markets, and at least 85 percent in the large-group market. If an insurer's MLR falls below these thresholds, it must issue rebates to enrollees. Moreover, insurers that remain noncompliant for three consecutive years are prohibited from participating in the market. The goal of the regulation is to create stronger incentives for insurers to operate efficiently and improve the healthcare quality for the society. It limits the portion of premium dollars for insurers that may excessively spend on administration, marketing, and profits.

Under the ACA, the MLR regulation requires that all health insurers to publicly report

aggregated state-level MLR data by market segment(i.e., individual, small group, and large group) in each year. The MLR is calculated as the ratio of medical claims incurred plus spending on quality improvement activities (numerator) to premium revenue net of federal and state taxes, licensing, and regulatory fees (denominator). Premium revenue reflects the earnings collected from policyholders, while incurred claims represent insurer payments for medical care and prescription drugs. Quality improvement expenditures include investments intended to enhance care quality, improve patient safety, prevent hospital readmissions, and increase transparency. The tax and fee adjustments account for federal assessments, state and local taxes, and regulatory licensing fees. Insurers with fewer than 1,000 enrollees in a given market are exempt from the MLR regulation, while those with relatively small enrollment (1,000 to fewer than 75,000 life-years) are eligible for credibility adjustments that effectively increase their reported MLR.

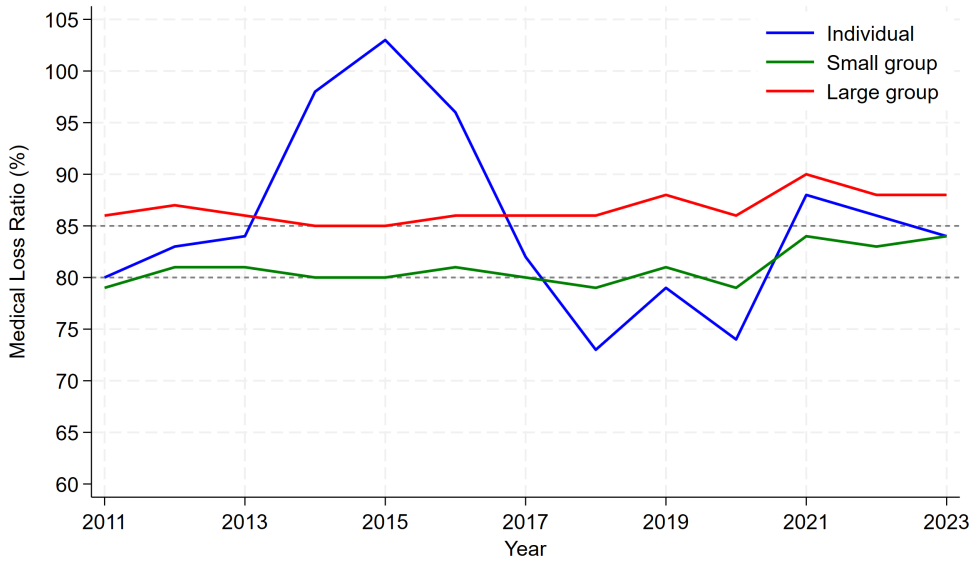
By the end of 2023, insurers had paid nearly $11.76 billion in rebates to consumers for failing to meet the MLR thresholds since 2012. Figure 1 displays the total rebates issued each year. The pattern shows that rebate payments increased sharply after 2018, peaking at nearly $2.5 billion in 2020. This sharp increase reflects the rising financial burden of MLR regulation on insurers and its influence on their market behavior. Figure 2 presents average MLRs across the individual, small group, and large group markets between 2011 and 2023. Large-group and small-group market have consistently maintained MLRs close to their respective regulatory thresholds. In contrast, the individual market exhibits substantial volatility, with MLRs exceeding 100 percent in 2015–2016 before declining sharply thereafter. This divergence reflects the greater pricing uncertainty in the individual market and highlights the MLR rule's central role in protecting consumers. This greater volatility and pricing uncertainty motivate the paper's focus on the individual market. These trends demonstrate that the MLR regulation not only influences insurers' pricing but also delivers significant consumer rebates, reinforcing its importance in today's health insurance markets.

Figure 1: Total Rebates by Year



Notes: Total rebates are calculated from CCIIO data, combining individual, small group, and large group markets. Rebates are expressed in billions of dollars.

Figure 2: Average MLR by Market, 2011–2023



Notes: Average MLR are calculated from CCIIO data.

2.2 The Vertical Relationship between the Insurers and Providers

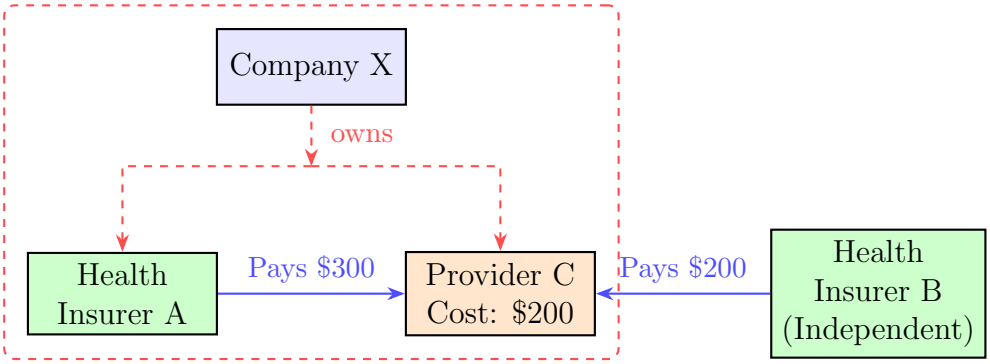
VI between health insurers and providers has accelerated in recent years because large parent companies expand ownership of physician groups, pharmacies, and pharmacy benefit managers (PBMs) alongside their insurance subsidiaries. A prominent example is UnitedHealth Group(UHG), the parent company of UnitedHealthcare, announced in December 2017 that its Optum would acquire DaVita Medical Group for approximately $4.9 billion (Mathews, 2017). UHG also owns OptumHealth, which employs or manages more than 90,000 physicians across the country by 2023, roughly 10% of the U.S. physician workforce (Weber, 2024). UHG has also expanded substantially, with approximately 2,700 subsidiaries in 2024 compared with about 985 in 2018.[4] Other major insurers have followed similar VI strategies. For example,Elevance Health (formerly Anthem, Inc.) owns Carelon, a health services provider that claims to serve one in three people in the US. CVS Health also intergrated with provider groups such as Oak Street Health and Signify Health. Such consolidation gives these firms substantial influence over the delivery, pricing, and reporting of healthcare services (Angeles and Bailit, 2025).

Figure 3 illustrates an example that how VI can affect reported medical spending under the MLR rule. Parent Company X owns both Health Insurer A and Provider C. It costs provider Y $200 to deliver a particular service. When a patient covered by independent Insurer B receives the service at Provider C, the payment is $200. For the same service delivered to a patient covered by Insurer A, the affiliated provider bills $300. Insurer A records the full $300 in its MLR numerator, although only $200 reflects actual care and the additional $100 is internal profit rather than additional care. This VI structure between the parent company and provider creates incentives for the subsidiary insurer to directly increase spending and raise reported MLRs without improving care. Consistent with these incentives, PhRMA (2024) reports that nearly two-thirds of OptumRx revenue comes from UnitedHealth Group affiliates, and Glickman (2025) documents that the share of Optum's

[4] Author's calculation based on UnitedHealth Group's 10-K filings.

revenue derived from UnitedHealthcare increased by nearly 50 percent. Internal transactions between CVS Pharmacy and other CVS Health businesses have totaled on the order of $50 billion (PhRMA, 2024).

Figure 3: Incentives for VI Insurers to Increase Spending to Affiliated Provider



Notes: This figure displays how VI affects insurers' spending. Independent Insurer B pays pays the full cost $200. However, the subsidiary Insurer A pays $300 for the same service ($200 for care and $100 internal profit). This VI structure between the parent company and provider creates incentives for the subsidiary insurer to directly increase spending and raise reported MLRs without improving care.

3 Data

3.1 Data Sources

I quantify insurer–provider vertical integration by focusing on four national parent companies: UnitedHealth Group, Inc.; The Cigna Group; Elevance Health, Inc. (formerly Anthem); and Humana Inc.[5] I collect information on each parent's licensed insurance subsidiaries from Form 10-K filings on the SEC's EDGAR website , which list legal entity names and states of registration.[6] Provider information comes from the National Plan and Provider Enumeration System (NPPES). I restrict attention to providers classified as organizations [7] because

[5]I exclude CVS Health Corporation because most provider organizations matched to CVS are pharmacies rather than care-delivery providers such as physician groups, clinics, or hospitals. Including CVS would therefore cause the vertical-integration measure to primarily capture insurer–pharmacy integration, which differs from the insurer–provider relationships examined in this paper.

[6]The subsidiary list in on "Exhibit 21: List of Subsidiaries"

[7]Entity Type Code = 2

VI typically occurs when insurers acquire or integrate with provider groups, hospitals, or pharmacies at the organizational level. I then match provider organization names to the five parent companies to determine ownership relationships. This procedure produces a state-level dataset indicating, for each year, whether a parent company owns providers in a given state.

I obtain MLR reports in the individual market from the Center for Consumer Information and Insurance Oversight (CCIIO). The CCIIO MLR Public Use Files provide detailed information by market segment (individual, small group, and large group), including MLRs, rebate amounts, enrollment, premium revenue, and related financial variables. If a parent insurer operates in multiple states, each licensed subsidiary insurers files a separate annual MLR report for each market segment.[8] Accordingly, the unit of observation in my analysis is the subsidiary insurer at the issuer–state–year level. I restrict the sample to issuers with at least 1,000 life-years [9], because issuers below this size are non-credible and not subject to rebate requirements.

For the structural model, I estimate the demand for insurance plans using plan-level data from Healthcare.gov's ACA Health Insurance Exchange Marketplace. The dataset reports premiums, deductibles, and other plan characteristics by insurer, state, and plan for individuals and families. The consumers can directly purchase insurance plans through the exchange. To measure plan-specific market shares, I merge Issuer Level Enrollment Data from Centers for Medicare & Medicaid Services (CMS) at the state–plan level with plan characteristics using the state specific issuer plan ID. Therefore I compute the average premium and other plan characteristrics at the state level. As a proxy for potential market size, I supplement these data with state-level counts of uninsured individuals from the U.S. Census Bureau.

[8]For example, UnitedHealthcare of Texas, Inc. and UnitedHealthcare of Florida, Inc. are distinct issuers. Each reports its own MLRs for Texas and Florida, respectively, and rebates owed in Texas cannot be offset by experience in Florida.

[9]Under the MLR regulation, 1,000 life-years is the credibility threshold. Issuers below this size are exempt from rebate requirements.

3.2 Summary Statistics

This paper focuses on the individual health insurance market from 2014 to 2024. The sample begins in 2014 because Marketplace coverage under the Affordable Care Act (ACA) took effect on January 1 of that year. Several major reforms also took effect that year, including guaranteed issue, modified community rating, and premium subsidies. Beginning the analysis in 2014 therefore provides a consistent regulatory setting and captures the individual market from the first year of Marketplace operation. The sample ends in 2024 because it is the most recent reporting year for which the Medical Loss Ratio Public Use Files are currently available. The resulting eleven-year period allows me to examine insurer–provider vertical integration and MLR outcomes throughout the development and maturation of the ACA individual market.

Table 1 reports summary statistics for the main analysis sample. After merging the MLR and VI data, the sample contains 19 subsidiary insurers operating in 41 insurer subsidiary–state markets, yielding 451 insurer subsidiary–state–year observations from 2014 to 2024. I restrict the sample to insurer subsidiary–state units that report MLR information in every year of the sample period. The mean MLR is 0.842, with a standard deviation of 0.076 and a range from 0.601 to 1.226. Approximately 35.9 percent of the observations are classified as vertically integrated. The mean claims share is 0.269, while the mean premium share is 1.009. Quality-improvement activities account for an average of 0.2 percent of the MLR denominator, and the mean gain share is 0.729. Covered life-years vary substantially across insurer subsidiary–state markets, with a mean of 76,163, a median of 31,908, and a range from 1,269 to 648,067. Panel B summarizes plan-level characteristics for the demand estimation. At the plan–state–year level, the mean annual premium is $6,657 and the average deductible is $4,764. Mean enrollment is 2,505 enrollees, average copayment is $27, and the average coinsurance rate is 16 percent. There is also substantial variation in plan characteristics across plans and over time.

Table 1: Descriptive statistics

	N	Mean	SD	Min	Median	Max
A. MLR reports and VI variables						
MLR	451	0.842	0.076	0.601	0.841	1.226
VI dummy	451	0.359	0.480	0.000	0.000	1.000
Claims share	451	0.269	0.145	0.000	0.246	0.990
Premium share	451	1.009	0.228	0.000	1.058	1.167
QIA share	451	0.002	0.002	0.000	0.002	0.012
Gain share	451	0.729	0.147	0.000	0.752	1.000
Life-years	451	76163	104226	1269	31908	648067
B. Plan variables						
N enrollees	20,170	2505.00	9577.46	11.00	490.00	421700.00
Premium ($)	20,170	6657.06	1670.98	2152.98	6505.71	19176.90
Deductible ($)	20,170	4763.81	2700.30	0.00	5250.00	9100.00
Copayment ($)	17,757	27.23	20.80	0.00	30.00	150.00
Coinsurance (percent)	6,973	16.14	19.95	0.00	0.00	60.00
Market share(percent)	20,170	0.48	1.31	0.00021	0.08	24.77

Notes: This table displays descriptive statistics for main variables. Panel A displays statistics after I merged MLR reports with VI. Panel B displays statistics for plan attributes in the demand estimation.

4 Reduced-Form Analysis

4.1 Effect of Vertical Integration on Medical Loss Ratios

I begin by examining whether insurer–provider vertical integration (VI) affects insurers' medical loss ratios (MLRs). The unit of observation is an insurer subsidiary–state–year in the ACA individual market. Although the MLR is reported at the subsidiary–state level, vertical integration is measured at the parent-insurer–state level because provider ownership decisions are generally made by the parent organization. I therefore assign each subsidiary the VI status of its parent insurer in the state and year in which the subsidiary operates.

The baseline sample includes subsidiaries affiliated with the major parent insurers for which I construct the vertical-integration measure. The treatment group consists of insurer–state units whose parent insurers transition from no identified provider ownership to vertical

integration during the sample period. The control group consists of insurer–state units whose parent insurers remain nonintegrated throughout the sample period. I restrict the analysis to a balanced panel covering 2014–2024 and apply the MLR credibility filter to reduce measurement error arising from particularly small enrollment bases.

For insurer subsidiary i operating in state s and year t, I estimate the following two-way fixed-effects specification:

$$y_{ist} = \beta \mathrm{VI}_{p(i)st} + \boldsymbol{X}'_{ist}\boldsymbol{\gamma} + \mu_{is} + \lambda_t + \varepsilon_{ist}, \tag{1}$$

where y_{ist} is the MLR of subsidiary i in state s and year t. The variable $\mathrm{VI}_{p(i)st}$ equals one if the parent insurer $p(i)$ of subsidiary i owns at least one identified care-delivery provider in state s and year t, and zero otherwise. The vector $\boldsymbol{X}_{ist}$ includes the logarithm of covered life-years, which controls for changes in the size of the insurer's enrolled population. Insurer subsidiary–state fixed effects, denoted by μ_{is}, absorb time-invariant differences across insurer–state units, including persistent differences in product composition, organizational structure, and local market position. Year fixed effects, denoted by λ_t, capture aggregate shocks and regulatory changes common to all insurers in a given year. The error term is ε_{ist}.

The coefficient of interest, β, measures the average within-insurer–state change in the MLR following vertical integration, relative to the contemporaneous change among insurer–state units that remain untreated. Because VI status varies at the parent-insurer–state level, the baseline specifications cluster standard errors at that level. I also report standard errors clustered at the company–state level as an alternative inference specification.

The causal interpretation of Equation (1) relies on the parallel-trends assumption: absent vertical integration, MLRs among treated and untreated insurer–state units would have evolved similarly. It also requires that the timing of integration is not driven by time-varying unobserved determinants of MLRs that are not absorbed by the included fixed effects and controls. I evaluate the plausibility of the parallel-trends assumption using event-study

specifications described below.

Table 2: Effect of Vertical Integration on Medical Loss Ratios

	(1)	(2)	(3)
Vertical integration	0.023*	0.020*	0.020*
	(0.012)	(0.012)	(0.011)
Log covered life-years		-0.012	-0.012*
		(0.007)	(0.007)
Insurer subsidiary × state FE	✓	✓	✓
Year FE	✓	✓	✓
Clustering level	Parent × state	Parent × state	Company × state
Observations	451	451	451
Within R^2	0.014	0.035	0.035

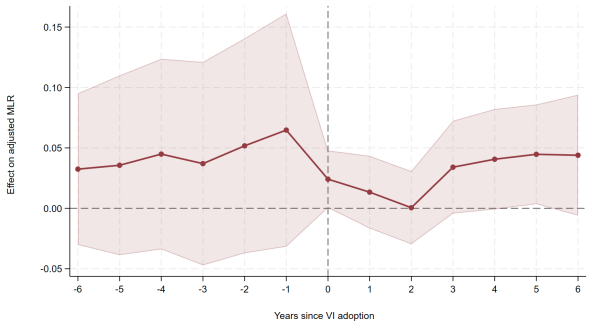
Notes: This table reports two-way fixed-effects estimates of the effect of insurer–provider vertical integration on adjusted medical loss ratios in the ACA individual market. The unit of observation is an insurer subsidiary–state–year. The estimation sample is a balanced panel covering 2014–2024 and is subject to the MLR credibility filter. All specifications include insurer subsidiary–state fixed effects and year fixed effects. Columns (2) and (3) additionally control for the logarithm of covered life-years. Standard errors are reported in parentheses. In Columns (1) and (2), standard errors are clustered at the parent-insurer–state level. In Column (3), standard errors are clustered at the company–state level. $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table 2 presents the baseline estimates. Column (1), which includes insurer subsidiary–state and year fixed effects but no additional covariates, indicates that vertical integration is associated with a 2.3 percentage points increase in the MLR. Column (2) controls for the logarithm of covered life-years and produces a slightly smaller estimate of 2.0 percentage points. Column (3) retains this control but clusters standard errors at the company–state level; the estimated effect remains 2.0 percentage points. The coefficients are statistically significant at the 10 percent level across all three specifications. The similarity of the estimates across specifications suggests that the result is not sensitive to controlling for enrollment or to the two reported clustering levels. Nevertheless, the relatively limited statistical precision warrants interpreting the baseline estimates together with the dynamic results and robustness analyses presented below.

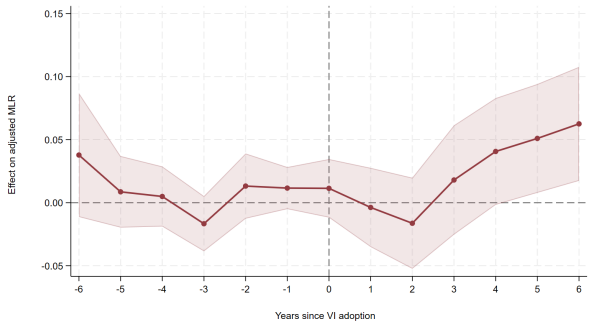
4.2 Dynamic Treatment Effects

A conventional two-way fixed-effects event study can produce contaminated comparisons when treatment is adopted at different times and treatment effects vary across cohorts or event time. I therefore estimate dynamic treatment effects using three complementary methods designed for staggered-adoption settings. The first is the imputation estimator of Borusyak et al. (2024), which estimates the untreated outcome model using untreated observations and imputes counterfactual outcomes for treated observations. The second is the group-time average treatment-effect estimator of Callaway and Sant'Anna (2021), which compares each treated cohort with appropriate untreated or not-yet-treated units. The third is a stacked difference-in-differences design that constructs cohort-specific samples and stacks them into a common event-study specification.

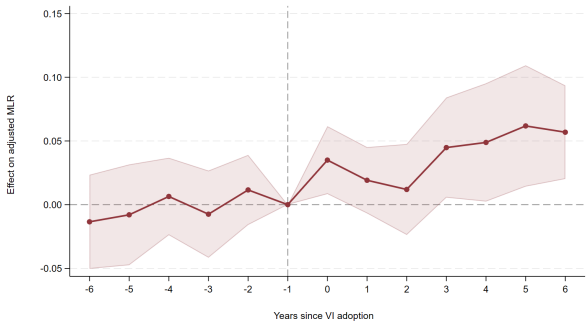
Figure 4: Dynamic Effects of Vertical Integration on Adjusted Medical Loss Ratios



(a) Borusyak–Jaravel–Spiess



(b) Callaway–Sant'Anna



(c) Stacked difference-in-differences

Notes: This figure reports event-study estimates of the effect of insurer–provider vertical integration on adjusted medical loss ratios in the ACA individual market. Panels (a), (b), and (c) present estimates from the imputation estimator of Borusyak et al. (2024), the group-time average treatment-effect estimator of Callaway and Sant'Anna (2021), and a stacked difference-in-differences estimator, respectively. The pre-treatment coefficients in Panels (a) and (b) are placebo estimates and therefore are not mechanically normalized to zero. Shaded areas represent 95% confidence intervals. All specifications control for the logarithm of covered life-years. Standard errors are clustered at the parent-insurer–state level.

Figure 4 presents the dynamic estimates. The pre-treatment coefficients are imprecisely estimated and generally statistically indistinguishable from zero, providing no clear evidence of systematic differential trends before vertical integration. This evidence supports, but does not establish, the parallel-trends assumption. The post-treatment estimates are generally positive, although their magnitude and precision vary across estimators and event years. The estimated effects are relatively modest during the first two years following integration and become larger approximately three to six years after adoption. This delayed pattern suggests that vertical integration may take time to affect insurer spending, provider contracting, or

the internal allocation of healthcare expenditures. Taken together, the three estimators indicate that the positive relationship between vertical integration and MLRs is not driven solely by the conventional two-way fixed-effects specification.

4.3 Robustness Checks

4.3.1 Alternative Measures of Vertical Integration

The baseline analysis measures vertical integration with an indicator for whether a parent insurer owns at least one identified provider in a given state and year. Although this measure captures the extensive margin of integration, it does not reflect potentially important differences in the scale of provider ownership. I therefore replace the binary indicator with four continuous measures: the logarithm of the number of owned providers, the number of distinct owned entities, the number of owned providers per 1,000 covered life-years, and the logarithm of this provider-intensity measure.

Table 3: Alternative Measures of Vertical Integration

	(1)	(2)	(3)	(4)
Log providers	0.010*			
	(0.006)			
Owned entities		0.002**		
		(0.001)		
Provider intensity			0.005***	
			(0.002)	
Log provider intensity				0.031*
				(0.016)
Controls	Yes	Yes	Yes	Yes
Insurer subsidiary × state FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	451	451	451	451
R^2	0.541	0.550	0.574	0.556

Notes: The dependent variable is the medical loss ratio. Each column replaces the binary vertical-integration indicator with an alternative continuous measure of integration intensity. Log providers is the logarithm of one plus the number of owned care-delivery providers identified through NPI registrations. Owned entities is the number of distinct owned entities in a parent-insurer–state–year. Provider intensity is the number of owned providers per 1,000 covered life-years. Log provider intensity is the logarithm of one plus provider intensity. All specifications control for the logarithm of covered life-years and include insurer subsidiary–state and year fixed effects. The estimation sample is the balanced individual-market panel and is subject to the MLR credibility filter. Standard errors are clustered at the parent-insurer–state level and reported in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table 3 reports the results. The estimated coefficients are positive across all four specifications and statistically significant at conventional levels. The number of distinct owned entities is positively associated with MLRs at the 5 percent level, while provider intensity is significant at the 1 percent level. The logarithmic measures are significant at the 10 percent level. Because these measures have different units and transformations, their coefficient magnitudes should not be compared directly across columns. The consistency of the coefficient signs nevertheless indicates that the baseline result is not driven solely by the binary definition of vertical integration. Instead, the results suggest that greater integration scale and intensity are also associated with higher MLRs.

4.3.2 Expanded Control Group

The primary analysis is restricted to the major parent insurers for which I systematically construct provider-ownership histories. As a second robustness check, I expand the sample to include other insurers operating in the ACA individual market. I assign these additional insurers to the control group because no vertical integration is identified for them in my ownership data. This exercise substantially broadens the comparison group and examines whether the baseline finding is driven by restricting the analysis to the major insurers included in the VI data. This specification requires a stronger classification assumption than the baseline analysis. Specifically, it assumes that the additional insurers classified as untreated do not own care-delivery providers during the relevant sample period. If some of these insurers are in fact vertically integrated, they are misclassified as untreated. Such misclassification would generally be expected to attenuate the estimated treatment effect if their integration also raises MLRs.

Table 4: Robustness: Expanded Sample Including Other Insurers

	(1) No controls	(2) With controls	(3) Alternative clustering
Vertical integration	0.067^{***}	0.057^{***}	0.054^{***}
	(0.011)	(0.011)	(0.011)
Log life-years		-0.012^{*}	-0.012^{*}
		(0.007)	(0.007)
Controls	No	Yes	Yes
Panel fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Clustering level	Parent insurer–state	Parent insurer–state	Company–state
Observations	2,145	2,145	2,394
Within R^2	0.015	0.024	0.019

Notes: The dependent variable is the medical loss ratio. This robustness analysis expands the estimation sample to include insurers outside the major parent insurers used to construct the vertical-integration measure. These additional insurers are assumed not to be vertically integrated and are assigned to the control group. Column (1) includes panel and year fixed effects without additional covariates. Column (2) additionally controls for the logarithm of covered life-years. Column (3) uses the same controls but clusters standard errors at the company–state level. The sample is a balanced panel covering 11 years and is subject to the credibility filter. Standard errors are reported in parentheses and clustered at the level indicated in the table. $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

Table 4 shows that the relationship between vertical integration and MLRs remains positive and statistically significant after expanding the control group. Column (1) indicates that vertical integration is associated with a 6.7 percentage point increase in the MLR when no additional covariates are included. After controlling for the logarithm of covered life-years in Column (2), the estimate decreases to 5.7 percentage points but remains statistically significant at the 1 percent level. Column (3) reports an estimate of 5.4 percentage points when standard errors are clustered at the company–state level. Overall, the expanded-sample estimates reinforce the conclusion that vertical integration is associated with higher MLRs and indicate that the baseline result is not driven by restricting the comparison group to the major insurers covered by the primary VI data.

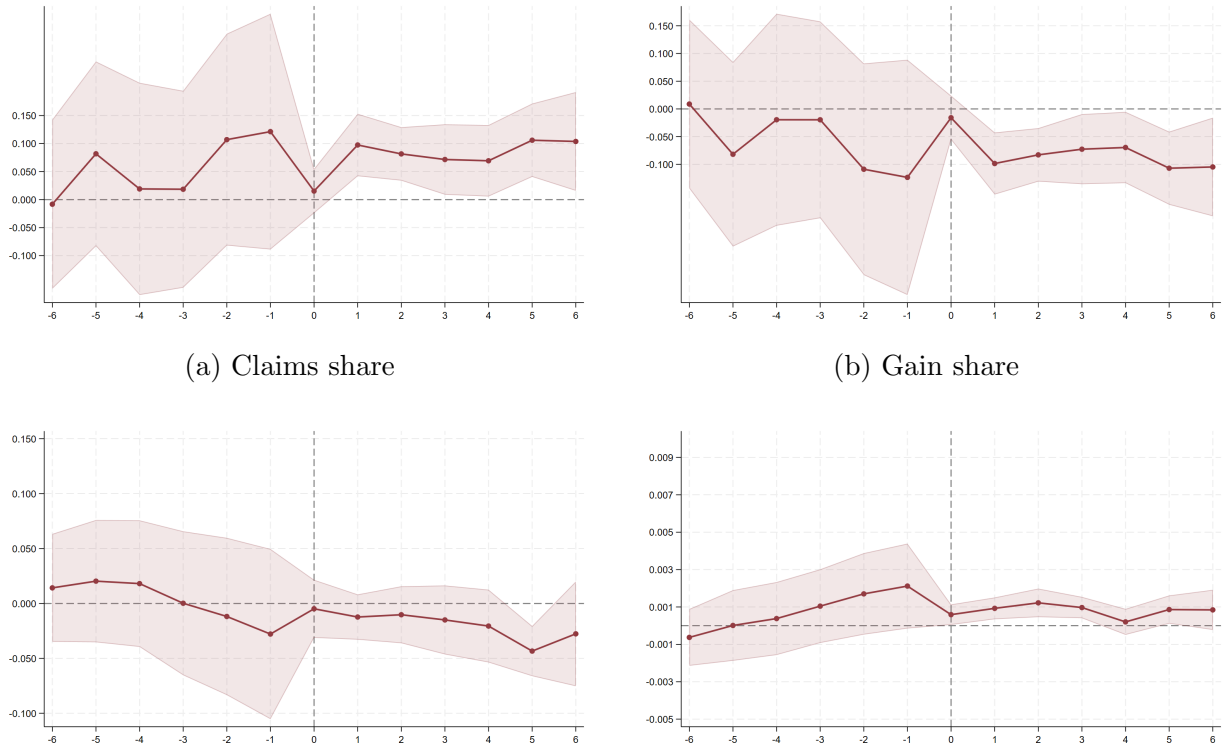
4.4 Mechanisms

I next examine which components of the MLR account for the increase following vertical integration. The MLR depends primarily on incurred claims and expenditures on quality-improvement activities relative to the applicable premium base. Vertical integration could increase the MLR by raising claims expenditures paid to affiliated or unaffiliated providers, increasing quality-improvement expenditures, changing premium revenues, or reducing the residual share retained by the insurer.

I consider four outcomes: the claims share, premium share, quality improvement activity share, and gain share. The claims share is total incurred claims divided by the MLR denominator. The premium share is gross earned premiums, including applicable risk-adjustment components, divided by the MLR denominator. Because the MLR denominator is net of certain taxes and regulatory fees, the premium share may be slightly greater than one. The quality improvement activity share is allowable quality-improvement expenditures divided by the MLR denominator. Finally, the gain share is defined as the premium share minus the claims share and the quality improvement activity share. A higher gain share therefore represents a larger residual retained by the insurer before accounting for other expenses not

included in this decomposition. For each outcome, I estimate dynamic treatment effects using the imputation estimator of Borusyak et al. (2024). The specifications include the logarithm of covered life-years and cluster standard errors at the parent-insurer–state level.

Figure 5: Dynamic Effects of Vertical Integration on Medical Loss Ratio Components



(a) Claims share

(b) Gain share

(c) Premium share

(d) Quality improvement activity share

Notes: This figure reports event-study estimates of the effects of insurer–provider vertical integration on components of the medical loss ratio using the imputation estimator of Borusyak et al. (2024). Panels (a)–(d) report effects on the claims share, gain share, premium share, and quality-improvement activity (QIA) share, respectively. Pre-treatment coefficients represent placebo estimates. Shaded areas represent 95% confidence intervals. All specifications control for the logarithm of covered life-years. Standard errors are clustered at the parent-insurer–state level.

Figure 5 indicates that the increase in MLRs is primarily associated with an increase in the claims share. The positive claims-share estimates become more pronounced several years after integration, broadly matching the delayed increase in the MLR. The gain-share estimates generally move in the opposite direction, suggesting that the increase in claims expenditures is accompanied by a reduction in the residual share retained by insurers. By comparison, the premium-share and quality-improvement-activity-share estimates provide

less evidence of persistent changes following integration. These patterns suggest that the increase in MLRs operates primarily through higher reported claims expenditures rather than through a sustained increase in premiums or quality improvement expenditures.

This decomposition is descriptive rather than a definitive test of the underlying mechanism. In particular, an increase in claims expenditures could reflect higher utilization, higher negotiated prices, changes in the composition of care, payments to affiliated providers, or a combination of these channels. The reduced-form evidence therefore identifies the accounting component through which the MLR changes but does not, by itself, distinguish among these economic mechanisms. In the next section, I develop a structural model to formally quantify these mechanisms and estimate the underlying effects of VI on insurer behavior.

5 The Model

The reduced-form analysis shows that VI between parent insurers and providers affects MLR and related mechanisms for subsidiary insurers, particularly through changes in claims costs. However, quantifying the relative importance of these mechanisms and determining their equilibrium implications remains challenging. For example,an increase in claims may occur if the MLR constraint induces insurers to reallocate internal transfers in a way that inflates reported medical spending and reduces rebate liability. These mechanisms cannot be separately identified with reduced-form evidence alone.

To address these limitations, I develop a structural model of the health insurance market that explicitly characterizes the interactions among subsidiary insurers, providers, and consumers. The model integrates consumer demand for insurance plans with insurers' premium-setting behavior under MLR regulation, explicitly distinguishing cases with and without VI. This structural framework enables me to quantify the impact of VI on consumer demand, insurer pricing, and to evaluate welfare outcomes under alternative regulatory scenarios.

The key agents in the model are (i) subsidiary insurers, who offer insurance plans and

set premiums; (ii) consumers, whose choices determine demand for these plans; and (iii) providers, who negotiate service prices with insurers. There are two stages in this model. In the first stage, insurers and providers bargain over reimbursement rates, and insurers simultaneously set premiums in anticipation of consumer demand and MLR regulation. In the second stage, consumers choose among available insurance plans based on premiums, cost-sharing provisions, and other plan characteristics. I describe the model by backward induction.

5.1 Demand for Insurance Plans

To allow for heterogeneity in consumer preferences, I model the demand for insurance plans using a discrete choice random utility model. This approach is based on Berry et al. (1993) and has been widely used in demand estimation (Grennan, 2013).

In the second stage, consumers purchase the insurance plans. Each consumer i chooses among plans j in state s and year t to maximize the following indirect utility:

$$u_{ijst} = \delta_{jst} + \mu_{ijst} + \varepsilon_{ijst}, \tag{2}$$

where δ_{jst} is the mean utility of plan j in market st, μ_{ijst} represents individual-specific deviations from the mean utility, and ε_{ijst} is the iid idiosyncratic error term distributed Type I Extreme Value distribution. The random component μ_{ijst} arises from interactions between individual characteristics and plan attributes, capturing taste heterogeneity. This utility function can be considered that how the consumer incorporates her own preferences and welfare into her plan decision.

The mean utility of plan j in state s and year t is given by:

$$\delta_{jst} = \sum_{k=1}^{K} \bar{\beta}_k x_{kjst} + \bar{\alpha}\phi_{jst} + \eta_{st} + \xi_{jst} \tag{3}$$

where x_{kjst} is the k-th observed characteristic of plan j in market st (e.g., deductible, copay-

ment, coinsurance) and ϕ_{jst} is the premium of plan j. The parameter $\bar{\beta}_k$ captures average consumer preferences for characteristic k, and $\bar{\alpha}$ represents the average marginal utility of premium for consumers. η_{st} is a state–year fixed effect that controls for market-level shocks and regulatory changes common to all plans in the same market. The term ξ_{jst} represents the unobserved plan characteristics, which is observed by consumers and insurers but not by the econometrician. It captures factors such as brand reputation, provider network quality, and service attributes that are not included in the observable plan characteristics.

The individual deviations from the mean utility is:

$$\mu_{ijst} = \sum_{k=1}^{K} x_{kjst} \left(\sigma_k^o d_{ikst} + \sigma_k^u v_{ik}\right) + \phi_{jst} \left(\sigma_\alpha^o d_{ist} + \sigma_\alpha^u v_i\right) \tag{4}$$

where d_{ikst} are the observed age demographics and v_{ik} are the unobserved individual characteristics. σ_k^o and σ_α^o are the random coefficients for observed demographics and premiums. σ_k^u and σ_α^u are the random coefficients for unobserved heterogeneity.

The total demand for plan j in market st is obtained by multiplying the predicted market share s_{jst} by the total market size M_{st}. The market share represents the expected probability that a consumer chooses plan j, integrating over the distribution of both observable (d_i) and unobservable (v_i) individual heterogeneity. The integral averages the individual choice probabilities across the population, assuming independent distributions $F(d_i)$ and $F(v_i)$. Formally, the total demand is given by:

$$D_{jst} = M_{st} \cdot s_{jst} = M_{st} \cdot \int \frac{\exp\left(\delta_{jst} + \mu_{ijst}\right)}{1 + \sum_{j'=1} \exp\left(\delta_{j'st} + \mu_{ij'st}\right)} \, dF(d_i) \, dF(v_i) \tag{5}$$

5.2 Supply Side

On the supply side, I define a *market* as a state-year combination. Each subsidiary insurer m maximizes profits by treating each state-year pair (s, t) as a independent market. To simplify notation, I omit the subscripts s and t in the theoretical exposition that follows. Due to

data limitations, I couldn't observe the network for the insurers. To simplify the analysis, I model a representative insurer–provider pair, which captures the key strategic interactions relevant for VI.

5.2.1 No Vertical integration

When there is no VI, insurer m sets premiums to maximize total profits, taking the negotiated service price and the MLR regulation as given. The expected profit for insurer m is:

$$\pi_m = \sum_j D_j(\phi_j - p_j\gamma_j\chi - C_F) \tag{6}$$

where D_j is the total demand for plan j, and ϕ_j is the observed premium. The term p_j denotes negotiated service price, γ_j is the coverage rate for those services, and χ is the fixed fraction of enrollees expected to utilize providers' care. $p_j\gamma_j$ is just the insurer's payment to provider for medical care. I assume that χ is constant across insurers and plans, reflecting average population-level utilization patterns. This simplification allows the model to focus on variation in insurer incentives and coverage design rather than on heterogeneity in utilization behavior. The product $p_j\gamma_j\chi$ captures the expected per-enrollee claims cost for plan j, while C_F represents the baseline administrative cost.

For the provider, profit maximization depends on the marginal cost of delivering care. The provider's profit function is given by

$$\pi_h = \sum_j D_j(p_j - c_h)\chi \tag{7}$$

where c_h is the provider's marginal cost per patient, assumed constant across insurance plans. Only a fixed fraction χ of these enrollees are expected to utilize provider services, implying that the provider's realized demand for care is $D_j\chi$.

5.2.2 Insurer's Behavior under Vertical Integration

If the subsidiary insurer and the provider are owned by the same parent insurer, the insurer partially internalizes the provider's profit when setting premiums. I model the insurer's supply-side decision as choosing premiums ϕ_j to maximize an internalized profit objective that reflects VI with the affiliated provider, given the negotiated service price p_j. Given the definitions of insurer and provider profits from the non-VI case, I define the general payoff function of the insurer subject to the MLR constraint:

$$\begin{aligned} \max_{\{\phi_j\}_j} \quad & \tilde{\pi}_m = \pi_m + \lambda \pi_h \\ \text{s.t.} \quad & \sum_j p_j \gamma_j \chi D_j \geq \bar{R} \sum_j \phi_j D_j \end{aligned} \tag{8}$$

where $\lambda \in [0, 1]$ is a parameter capturing the degree to which the insurer internalizes the provider's profit. In the absence of VI ($\lambda = 0$), the insurer maximizes only its own profit π_m. If vertically integrated ($\lambda > 0$), the insurer's objective incorporates a share λ of the provider's downstream profit π_h. When $\lambda = 1$, the insurer fully internalizes the provider's profit and maximizes the joint profit of both divisions. Intermediate values of λ represent partial internalization, indicating that the vertically integrated firm places an unbalanced weight on upstream and downstream profits. A lower value of λ implies that the insurer gives greater priority to downstream (insurance) profits, whereas a higher λ reflects stronger concern for upstream (provider) profits. This imbalance highlights that even within a vertically integrated parent company, the subsidiary insurer and provider divisions may not behave as a single, unified profit-maximizing entity.

This formulation accounts for internal imbalances that may result from uneven regulatory oversight across insurance and provider divisions, tax-related asymmetries, or limited profit internalization, as emphasized by Crawford et al. (2018) and Cuesta et al. (2024). Moreover, insurers and providers within the same parent firm often operate under separate budgets, performance metrics, and managerial incentives, which limits coordination across divisions.

Accounting rules, reimbursement systems, and tax treatment may further restrict profit consolidation, and partial ownership or joint ventures can limit the insurer's financial claim on provider profits. Therefore, λ summarizes these frictions by indicating how much weight the insurer places on upstream provider profits relative to downstream insurance profits, reflecting the degree of internalization within the vertically integrated firm.

Now, the insurer's problem can be:

$$f^I = \sum_j D_j(\phi_j - p_j\gamma_j\chi - C_F) + \mu\left[\sum_j p_j\gamma_j\chi D_j - \bar{R}\sum_j \phi_j D_j\right] + \lambda\sum_j D_j(p_j - c_h)\chi$$

Then, I rewrite it as:

$$f^I = \underbrace{\sum_j D_j\left[(1-\mu\bar{R})\phi_j - (1-\mu)p_j\gamma_j\chi - C_F\right]}_{\substack{\text{Insurer profit without} \\ \text{internalization } (L^I)}} + \lambda\sum_j D_j(p_j - c_h)\chi \tag{9}$$

where μ is the Lagrangian multiplier, which means the extra cost for MLR regulation. Here, L^I denotes the insurer's effective profit component under MLR regulation in the absence of VI. It summarizes the insurer's revenues, expected claims costs, and administrative expenses after accounting for the shadow price μ associated with the MLR constraint. Therefore, L^I captures the insurer's MLR-adjusted operating profit, while the additional term $\lambda\sum_j D_j(p_j - c_h)\chi$ reflects the internalized provider profit under VI.

The first order condition of f^I with respect to ϕ_j given optimal p_j:

$$\begin{aligned}\frac{\partial f^I}{\partial \phi_j} =& (1-\mu\bar{R})\left[D_j + \sum_k \phi_k \frac{\partial D_k}{\partial \phi_j}\right] - (1-\mu)p_j\chi\sum_k \gamma_k\frac{\partial D_k}{\partial \phi_j} - C_F\sum_k \frac{\partial D_k}{\partial \phi_j} \\ &+ \lambda\chi(p_j - c_h)\sum_k \frac{\partial D_k}{\partial \phi_j} = 0\end{aligned}$$

Then, I can get:

$$(1-\mu\bar{R})D_j + \sum_k \left[(1-\mu\bar{R})\phi_k - (1-\mu)\gamma_k\chi p_j - C_F + \lambda\chi(p_j - c_h)\right]\frac{\partial D_k}{\partial \phi_j} = 0$$

$$\Rightarrow \sum_k \left[(1-\mu\bar{R})\phi_k - (1-\mu)\gamma_k\chi p_j - C_F\right]\cdot\frac{\partial D_k}{\partial \phi_j} + \sum_k \lambda\chi(p_j - c_h)\cdot\frac{\partial D_k}{\partial \phi_j} = -(1-\mu\bar{R})D_j$$

I rewrite as the matrix form for total derivative of f^I with respect to $\boldsymbol{\phi}$:

$$\begin{aligned}(1-\mu\bar{R})\,\boldsymbol{D}(\boldsymbol{\phi}) \;+\; (1-\mu\bar{R})\,\mathbf{J}^\top\boldsymbol{\phi} - (1-\mu)\chi\,(\mathbf{J}^\top\boldsymbol{\gamma})\odot\boldsymbol{p} \\ - C_F\,\mathbf{J}^\top\mathbf{1_J} + \lambda\chi\,(\mathbf{J}^\top\mathbf{1_J})\odot(\boldsymbol{p} - c_h\mathbf{1_J}) = \mathbf{0_J}.\end{aligned}$$

where $\boldsymbol{D}$ denotes the vector of plan-level demands, $\boldsymbol{\phi}$ the vector of premiums , $\boldsymbol{\gamma}$ the vector of coverage rates, and $\boldsymbol{p}$ the vector of negotiated service price . All vectors are of dimension $J \times 1$ unless otherwise stated. $\odot$ indicates element-wise (Hadamard) multiplication. The Jacobian matrix $\mathbf{J} \in \mathbb{R}^{J\times J}$ represents the sensitivity of demand to premiums. μ, $\bar{R}$, χ, C_F, and c_h are scalar parameters corresponding to the MLR shadow value, MLR threshold, utilization rate, administrative cost, and provider marginal cost, respectively.

Then, I rearrange for $\mathbf{J}^\top\boldsymbol{\phi}$:

$$\begin{aligned}(1-\mu\bar{R})\,\mathbf{J}^\top\boldsymbol{\phi} = -(1-\mu\bar{R})\,\boldsymbol{D}(\boldsymbol{\phi}) + (1-\mu)\chi\,(\mathbf{J}^\top\boldsymbol{\gamma})\odot\boldsymbol{p} \\ + C_F\,\mathbf{J}^\top\mathbf{1_J} - \lambda\chi\,(\mathbf{J}^\top\mathbf{1_J})\odot(\boldsymbol{p} - c_h\mathbf{1_J}).\end{aligned}$$

Finally, solving for $\boldsymbol{\phi}$:

$$\boldsymbol{\phi} = -(\mathbf{J}^\top)^{-1}\boldsymbol{D} + \frac{1}{1-\mu\bar{R}}\,(\mathbf{J}^\top)^{-1}\left[(1-\mu)\chi\,(\mathbf{J}^\top\boldsymbol{\gamma})\odot\boldsymbol{p} + C_F\,\mathbf{J}\mathbf{1}_J - \lambda\chi\,(\mathbf{J}^\top\mathbf{1_J})\odot(\boldsymbol{p} - c_h\mathbf{1_J})\right]$$

If $p_j = p$, then

$$\boldsymbol{\phi} = -(\mathbf{J}^\top)^{-1}\boldsymbol{D} + \frac{(1-\mu)}{1-\mu\bar{R}}\,\chi p\boldsymbol{\gamma} + \frac{C_F\mathbf{1}_J}{1-\mu\bar{R}} - \frac{\lambda\chi}{1-\mu\bar{R}}\,(p - c_h)\mathbf{1}_J \tag{10}$$

Equation (10) indicates that equilibrium premiums depend on both the weighted medical cost component and the elasticity-adjusted markup implied by differentiated-product demand. The first term, $-(\mathbf{J}^\top)^{-1}\boldsymbol{D}$, represents the standard markup condition implied by differentiated-product demand, where $(\mathbf{J}^\top)^{-1}$ captures the inverse of the demand elasticity matrix. The second term, $\frac{(1-\mu)}{1-\mu\bar{R}}\,\chi p\boldsymbol{\gamma}$, reflects the expected claims cost adjusted for the MLR regulation, where $(1-\mu)/(1-\mu\bar{R})$ scales the share of medical spending effectively retained by the insurer after accounting for potential rebate obligations. The third term, $\frac{C_F\mathbf{1}_J}{1-\mu\bar{R}}$, shows per-enrollee administrative and fixed operating costs, which are amplified under stricter MLR constraints. The final term, $-\frac{\lambda\chi\,(p-c_h)}{1-\mu\bar{R}}\mathbf{1}_J$, captures the effect of VI, where $\lambda \in [0,1]$ denotes the degree of profit internalization between the insurer and its affiliated provider. In the absence of VI ($\lambda = 0$), equation (10) simplifies to a standard MLR-constrained pricing rule in which premiums equal medical costs plus a markup scaled by demand elasticity.

5.2.3 Provider's Behavior Under Vertical Integration

Under VI, both the insurer and the provider internalize a portion of each other's profits because they operate within the same parent organization. The key idea is that integration not only reshapes internal accounting flows but also fundamentally modifies the objective functions of the affiliated divisions. Rather than maximizing stand-alone profits, each division behaves as if it cares about the consolidated profit of the parent company. This assumption of *symmetric internalization* is standard in the vertical relations literature (Crawford et al., 2018), and is commonly interpreted as partial profit pooling within the firm. It captures that the parent company coordinates pricing and transfer decisions across divisions to internalize complementarities, while still allowing for incomplete alignment of incentives due to regulatory, contractual, or managerial frictions.

Formally, the vertically integrated provider maximizes a weighted sum of its own profit and the internalized share of the insurer's profit, taking into account the MLR regulation

faced by the insurer. The provider's objective function is given by

$$f^H = \sum_j D_j(p_j - c_h)\chi + \lambda\left[\sum_j D_j(\phi_j - p_j\gamma_j\chi - C_F) + \mu\Big(\sum_j p_j\gamma_j\chi D_j - \bar{R}\sum_j \phi_j D_j\Big)\right] \tag{11}$$

The objective in equation (11) captures the cooperative nature of pricing under VI: the provider's negotiated price p_j now balances the provider's direct revenue from patient services with its indirect effects on insurer premiums, MLR compliance, and overall joint profitability. This internalization mechanism ensures that the provider considers the downstream implications of its pricing decisions on insurer outcomes, thereby connecting the supply-side bargaining problem to the regulatory environment.

5.3 Prices Negotiations

In the first stage, insurers and health care providers negotiate the service price p_j, which determines the payment per unit of care delivered to insured enrollees. I assume that the prices p_j are determined simultaneously with insurer premium setting through bilateral negotiations between one insurer-provider pair. In practice, insurers typically negotiate service prices for a portfolio of plans offered within the same market simultaneously. Therefore, I assume that all plans offered by the same insurer in a given market share the same negotiated service price, that is, $p_j = p$ for all $j \in J_m$. Each insurer-provider pair maximizes a bilateral Nash product, taking the prices negotiated by all other pairs as given.

A key difference between this framework and standard bargaining models is the presence of the MLR constraint, which restricts the insurer's feasible profit. Hence, the insurer's objective entering the Nash product corresponds to its constrained profit f^I, as defined in equation (9). Similarly, the provider's profit is given by f^H, as defined in equation (11). The negotiated service price p is then determined as the outcome of the bilateral Nash bargaining

problem:

$$p = \arg\max_{p} \left(f^I - f_0^I\right)^{\tau} \left(f^H - f_0^H\right)^{1-\tau} \tag{12}$$

where $\tau \in [0, 1]$ denotes the insurer's bargaining power, f_0^I and f_0^H represent the respective disagreement payoffs in the event of bargaining failure, and (f^I, f^H) are the profits realized if agreement is reached. Consistent with the literature, I normalize the outside options to zero ($f_0^I = f_0^H = 0$), implying that the insurer exits the market and the provider treats patients covered by other insurers if no agreement is reached.

This Nash product formulation reflects the joint surplus-maximizing nature of price negotiation between insurers and providers, weighted by their relative bargaining power. The MLR constraint modifies the insurer's marginal benefit from higher prices by limiting the fraction of premium revenue that can be retained as profit. In addition, VI introduces a further channel through the internalization parameter λ, which causes both sides to partially internalize the other's payoff. These two forces, MLR regulation and internalization, shape how bargaining outcomes translate into service prices and insurer premiums.

The F.O.C with respect to p for the logarithmic derivative of Nash product is:

$$\frac{\tau}{f^I}\left(\frac{\partial f^I}{\partial p} + \sum_j \frac{\partial f^I}{\partial \phi_j} \cdot \frac{\partial \phi_j}{\partial p}\right) + \frac{1-\tau}{f^H}\left\{\chi \sum_j \left[D_j + (p - c_h) \sum_j \frac{\partial D_j}{\partial \phi_k} \cdot \frac{\partial \phi_k}{\partial p}\right] + \lambda \cdot \frac{\partial L^I}{\partial p}\right\} = 0 \tag{13}$$

I first use the envelope theorem to get the optimal ϕ given optimal p^*:

$$f^I = \boldsymbol{D}^\top \cdot \left[-(\mathbf{J}^\top)^{-1}(1 - \mu\bar{R})\boldsymbol{D}\right] = -(1 - \mu\bar{R})\boldsymbol{D}^\top(\mathbf{J}^\top)^{-1}\boldsymbol{D} \tag{14}$$

Then, I substitute into equation (13):

$$
\begin{aligned}
&\frac{\tau}{f^I}\left(-\sum_j \gamma_j \chi D_j + \mu \sum_j \gamma_j \chi D_j + \lambda \chi \sum_j D_j\right) \\
&\quad + \frac{(1-\tau)}{f^H}\Bigg\{\chi \sum_j D_j + \lambda \sum_j \left[\left((1-\mu\bar{R})\frac{\partial \phi_j}{\partial p} - (1-\mu)\gamma_j \chi\right) D_j\right] \\
&\quad + \left[\sum_j (p - c_h)\chi \sum_k \frac{\partial D_j}{\partial \phi_k} \cdot \frac{\partial \phi_k}{\partial p} + \lambda \sum_j \left((1-\mu\bar{R})\phi_j - (1-\mu)\gamma_j \chi p - C_F\right) \sum_k \frac{\partial D_j}{\partial \phi_k} \cdot \frac{\partial \phi_k}{\partial p}\right]\Bigg\} = 0
\end{aligned}
$$

Finally, I solve the negotiated price:

$$
\begin{aligned}
p = c_h &+ \underbrace{\frac{\lambda(1-\mu\bar{R})\,\boldsymbol{D}^\top (\mathbf{J}^\top)^{-1}\boldsymbol{D}}{(1-\lambda^2)\,\chi\,\mathbf{1}^\top \boldsymbol{D}}}_{\text{first term}} \\
&- \underbrace{\frac{\boldsymbol{D}^\top\left[\chi \mathbf{1} + \lambda(1-\mu\bar{R})\,\frac{\partial \boldsymbol{\phi}}{\partial p} - \lambda(1-\mu)\chi\boldsymbol{\gamma}\right]}{(1-\lambda^2)\chi\,\mathbf{1}^\top \boldsymbol{D}}}_{\text{second term}} \\
&\times \underbrace{\left[\frac{\mathbf{1}^\top \mathbf{J}\,\frac{\partial \boldsymbol{\phi}}{\partial p}}{\mathbf{1}^\top \boldsymbol{D}} + \frac{\tau\chi}{(1-\tau)(1-\mu\bar{R})}\,\frac{\boldsymbol{D}^\top\big[(1-\mu)\boldsymbol{\gamma} - \lambda\big]}{\boldsymbol{D}^\top (\mathbf{J}^\top)^{-1}\boldsymbol{D}}\right]^{-1}}_{\text{third term}}
\end{aligned}
\tag{15}
$$

Equation (15) shows that the negotiated price equals the provider's marginal cost c_h plus a markup. This markup can be decomposed into three components, each capturing a distinct economic mechanism in the insurer–provider negotiation. The *first term* reflects the internalization effect of VI: as λ increases, the insurer and provider partially internalize each other's profits, which raises the negotiated price above cost. The *second term* represents the regulatory adjustment, illustrating how the MLR constraint and the insurer's premium response to price changes limit the marginal benefit of higher prices. The *third term* captures the inverse effective elasticity, scaling the overall price response by how sensitive demand and premiums are to changes in the negotiated price. Together, these components demonstrate how VI, regulation, and demand elasticity jointly determine the equilibrium service price.

To obtain $\frac{\partial \boldsymbol{\phi}}{\partial p}$, I apply the implicit function theorem to equation (10). For each plan $j \in$

$\{1, 2, \ldots, J\}$, differentiating the insurer's premium condition with respect to the negotiated provider price p yields:

$$\left\{ \left[(\mathbf{J}^\top)^{-1} \boldsymbol{D} \right]^\top \boldsymbol{H}_j - \mathbf{J}^\top - \mathbf{J} \right\} \boldsymbol{\phi}'(p) \;=\; \mathbf{J}^\top \left[-\frac{1-\mu}{1-\mu\bar{R}} \boldsymbol{\gamma} + \frac{\lambda\chi}{1-\mu\bar{R}} \mathbf{1} \right] \tag{16}$$

where $\boldsymbol{H}_j$ denotes the Hessian of the demand system with respect to premiums for plan j, and $\boldsymbol{\phi}'(p) \equiv \frac{\partial \boldsymbol{\phi}}{\partial p}$ is the vector of premium responses to changes in the negotiated service price. The left-hand side collects the Jacobian and curvature effects of the demand system, while the right-hand side captures the direct influence of medical cost adjustments and profit internalization through $\boldsymbol{\gamma}$ and λ. This expression links the premium adjustment to both the regulatory parameters μ and the degree of VI λ. The detailed derivation is provided in Appendix B.

6 Structural Analysis

6.1 Demand Estimation for Insurance Plans

I estimate consumer demand for health insurance plans following a random-coefficients discrete choice model provided by Berry et al. (1993). I recover the predicted market shares based on equation (5). In practice, this integral is approximated by simulation using N_s random draws of (d_i, v_i):

$$\hat{s}_{jst} = \frac{1}{N_s} \sum_{i=1}^{N_s} \frac{\exp(\delta_{jst} + \mu_{ijst})}{1 + \sum_{j' \in \mathcal{J}_{st}} \exp(\delta_{j'st} + \mu_{ij'st})}.$$

Then, I calculate the mean utilities of product j, δ_{jst} such that predicted market shares match the observed market shares using the contraction mapping. The unobserved plan charateristics ξ_{jst} is recovered as the model-implied residual. Then, the moment conditions

of the Generalized Method of Moments (GMM) method to identify the demand model is:

$$\mathbb{E}[z'_{jst}\xi_{jst}] = \mathbb{E}[z'_{jst}(\delta_{jst} - x_{jst}\bar{\beta} - \bar{\alpha}\phi_{jst})] = 0, \tag{17}$$

The identifying assumption is that unobserved plan charateritics ξ_{jst} is uncorrelated with the instruments z_{jst}. I use two sets of instrument variables: (i) average plan characteristics of rival plans within the same market (ii) average premiums and plan characteristics of plans offered by the same insurer in other markets. The identification of the demand parameters arises from exogenous variation within state-year in plan characteristics, based on different premiums and other characteristics across different insurers.

Table 5 reports the results from the random coefficients demand estimation. The estimated coefficient on the premium variable and deductible are negative and statistically significant, indicating that higher premiums and deductible reduce plan demand. The copayment coefficient is also negative and significant, implying that consumers are sensitive to the amount of out-of-pocket for each service directly. In contrast, the coinsurance rate is insignificant, indicating limited responsiveness to marginal cost-sharing rates. This finding is consistent with the Loewenstein et al. (2013) that consumers prefer insurance plan they are easily to understand. The significant standard deviations for premium, deductible, and copayment indicate substantial heterogeneity in consumers' price sensitivity and risk preferences. These results suggest meaningful variation in demand behavior across individuals, with important implications for insurer pricing strategies and regulatory compliance in vertically integrated markets. The estimated parameters further allow the computation of the Jacobian and Hessian matrices of demand, which serve as key inputs for the supply-side bargaining estimation in Section 6.2.

Table 5: Results of Random-Coefficients Demand Estimation

Variable	Means	Standard Deviations (σ)
Premium (per \$1k)	-0.3712^{**}	0.007^{***}
	(0.145)	(0.0002)
Deductible (per \$1k)	-0.0194^{***}	0.014^{***}
	(0.005)	(0.002)
Copayment (\$)	-0.003^{***}	0.012^{*}
	(0.0005)	(0.007)
Coinsurance (%)	0.001	0.002
	(0.122)	(2.602)
Fixed Effects	Yes	
Markets	215	
Observations	27,862	

Notes: This table reports estimates from the random-coefficients demand model using the efficient GMM estimator. Standard errors are reported in parentheses. The fixed effects include metal category FE, plan type FE, state-year FE. The *premium* and *deductible* variables are measured per \$1,000, the *copayment* variable is measured in dollars per visit, and *coinsurance* is expressed in percentage points. The mean coefficients capture average consumer preferences, while the standard deviations reflect heterogeneity across consumers. *Significant at the 10% level, **significant at the 5% level, ***significant at the 1% level.

6.2 Price Negotiation Estimation

The identification for price negotiation estimation based on the marginal cost c_h. The provider's marginal cost c_h is determined endogenously through the bargaining equilibrium between the insurer and the provider. Due to data limitations, I do not observe the negotiated prices between insurers and providers. As a result, the utilization parameters χ and γ are not directly identified. I normalize $\chi = 1$, which implies that a fixed fraction of enrollees are expected to use provider services.

In practice, I approximate γ by the complement of the plan's coinsurance rate, $\gamma = 1 - \text{Coinsurance}$. This assumption treats the insurer's coverage rate as the fraction of allowed costs paid by the plan after cost sharing, consistent with standard actuarial definitions. Becasuse this paper don't focus on the plan design. This simple calibration of coverage rate

offers a way to incorporate plan-level benefit generosity into the bargaining model without introducing additional unobserved parameters.

Following the model developed in section 5, I eliminate p and get a function of c_h:

$$c_h = f\left(\lambda, \tau, \mu, C_F \,\middle|\, \boldsymbol{\phi}, \boldsymbol{\gamma}, \bar{R}, \chi, \mathbf{D}, \mathbf{J}, \mathbf{H}\right) \tag{18}$$

where $f(\cdot)$ represents the equilibrium cost function implied by the joint insurer–provider problem. Equation (18) summarizes how the equilibrium provider's cost c_h depends on the structural primitives $(\lambda, \tau, \mu, C_F)$ and equilibrium observables $(\boldsymbol{\phi}, \boldsymbol{\gamma}, \bar{R}, \chi, \mathbf{D}, \mathbf{J}, \mathbf{H})$.[10] The full expression of $f(\cdot)$ are provided in Appendix C.

To identified the parameters, I express the provider's marginal cost c_h as a function of obervables g_h:

$$c_h = \omega g_h + \varphi_h \tag{19}$$

where g_h represents provider characteristics that influence the cost of delivering care, φ_h is the unobservable component of provider costs. In principle, g_h could include detailed provider-level attributes such as input cost, market structure and provider size. However, I do not observe detailed providers characteristics due to data limitation. To address this limitation, I interpret g_h as the average cost environment faced by the providers in each state–year market. Specifically, I construct g_h using aggregate measures of provider input costs and market structure, including the average Medicare wage index across states from CMS[11], average operating expense per discharge from Healthcare Cost Report Information System (HCRIS). [12] These variables vary across states and years, capturing systematic differences in the cost of care delivery.

The econometric error for the GMM estimator is then defined as the difference between

[10] I compute demand,Jacobian and Hessian matrices of demand from demand estimation
[11] The data come from CMS Wage Index Public Use File.
[12] Hospital Provider Cost Report from CMS

the projected cost from Equation (19) and the cost implied by the equilibrium conditions from Equation (18). That is, I define the econometric error for a insurer-provider pair as:

$$\boldsymbol{\psi}_h = \omega g_h - f\left(\lambda, \tau, \mu, C_F \,\middle|\, \mathbf{J}, \boldsymbol{\phi}, \boldsymbol{\gamma}, \mathbf{D}, \chi, \bar{R}\right) \tag{20}$$

Then, the moment conditions for the supply side estimation is:

$$\mathbb{E}[\boldsymbol{z}'\boldsymbol{\psi}] = 0 \tag{21}$$

For the intruments in $\boldsymbol{z}$, I use the variables in g_h, lagged credibility adjustment factor, lagged average plan characteristics of rival plans within the same market, lagged average plan characteristics offered by the same insurer in other markets, lagged number of insurers in the market.

Table 6 presents separate parameter estimates for non-compliant and compliant insurers. The estimated internalization parameter (λ) is substantially larger for non-compliant insurers, indicating that firms facing a binding MLR constraint internalize a greater share of provider profits. This result suggests that VI can reallocate spending and inflate reported medical claims to meet the regulatory threshold. Consistent with this interpretation, the estimated MLR tightness parameter (μ) is higher among non-compliant insurers, reflecting a stronger binding constraint. In contrast, the bargaining power parameter (τ) is lower for non-compliant insurers, implying that when the MLR binds, insurers have weaker incentives to negotiate lower provider prices since higher provider payments facilitate compliance.Finally, the estimated fixed cost parameter (C_F) is larger for non-compliant insurers, consistent with higher administrative and compliance-related expenditures. All these results indicate that VI alters the insurer–provider bargaining equilibrium by intensifying internal transfers under MLR regulation.

Table 6: Effect of MLR Regulation and Bargaining

		Non-compliant	Compliant
Effect of VI	λ	0.325	0.124
		(0.001)	(0.010)
Effect of MLR Regulation	μ	0.178	0.124
		(0.005)	(0.021)
Nash Bargaining Parameters	τ	0.563	0.648
		(0.025)	(0.035)
Insurer Fixed Cost	C_F	1.325	0.895
		(0.235)	(0.005)

Notes: This table reports parameter estimates from the insurer–provider bargaining model, shown separately for non-compliant and compliant insurers. λ measures the effect of VI, μ captures the binding strength of the MLR constraint, τ denotes the insurer's bargaining weight, and C_F represents fixed costs. Standard errors are in parentheses.

6.3 Counterfactual Analysis

This section presents a set of counterfactual experiments to evaluate how alternative ownership structures and regulatory regimes affect equilibrium outcomes in the health insurance market. Using the estimated structural parameters from the demand and supply model, I simulate how insurers' pricing and welfare outcomes respond to changes in market structure and policy design. I consider consider four counterfactual scenarios: (1) removing VI, (2) higher VI (3) eliminating the MLR regulation (4) imposing a stricter MLR threshold. The detailed equilibrium price formulations for each scenario are provided in Appendix D.

In the first scenario, I remove all insurer–provider ownership links by setting the internalization parameter to $\lambda = 0$, allowing insurers and providers to operate as fully independent entities. The second scenario imposes higher degree of VI ($\lambda = 0.9$) where profit internalization is strong, providing an upper-range benchmark for the influence of integration on prices and welfare. The third scenario eliminates the MLR regulation by setting $\mu = 0$, enabling an evaluation of how the market would operate without regulatory constraints. Finally, the fourth scenario raises the MLR threshold from 80% to 90%, assessing how a stricter compli-

ance standard affects insurer incentives and the trade-off between consumer protection and market efficiency. All these scenarios provide a structured framework to identify the separate and joint effects of VI and MLR regulation on market performance.

Table 7 shows the counterfactual simulation results in different scenarios. Removing VI ($\lambda = 0$) eliminates internal transfer markups and foreclosure incentives, leading to lower negotiated provider prices and premiums, thereby improving consumer surplus. In contrast, higher integration ($\lambda = 0.9$) or stricter MLR enforcement ($\bar{R} = 0.9$) raises negotiated prices and premiums as insurers internalize provider margins and expand reported medical claims to satisfy the regulatory threshold. These higher costs are largely passed on to consumers, reducing consumer surplus. When the MLR requirement is removed ($\mu = 0$), insurers have weaker incentives to inflate claims for relax regulatory pressure, resulting in lower negotiated prices and premiums and higher welfare. These patterns suggest that VI allows insurers to meet compliance thresholds through accounting adjustments and internal pricing rather than through true efficiency gains.

Table 7: Counterfactual Simulation of Insurer Market Outcomes

	Δ Premiums (%)	Δ Negotiated Price (%)	Δ Consumer Surplus (%)
Scenario 1: No VI ($\lambda = 0$)	-3.51	-6.83	+2.25
Scenario 2: Higher VI ($\lambda = 0.9$)	+0.22	+2.30	-1.92
Scenario 3: No MLR ($\mu = 0$)	-1.83	-3.41	+4.78
Scenario 4: Stricter MLR ($\bar{R} = 0.9$)	+1.98	+3.15	–1.53
Baseline average	\$6.65	\$4.25	\$30.28

Notes: This table reports simulated changes in insurer market outcomes under each counterfactual, expressed as percentage deviations from the baseline equilibrium. Premiums are averaged at the plan level. Baseline levels are expressed in millions of USD for consumer surplus and thousands for premium and negotiated service price.

7 Conclusion

This paper investigates how VI between insurers and providers affects regulatory compliance and market outcomes in the U.S. health insurance market. Using comprehensive data, I find that VI between parent insurers and providers increases the MLRs of their subsidiary insurers. In particular, this increase is primarily associated with higher reported claims expenditures rather than changes in premiums or quality improvement expenditures. Motivated by this evidence, I develop a structural model with an MLR constraint to quantify how VI influences premiums, negotiated service prices, and consumer welfare. This study bridges the literatures on VI, MLR regulation, and healthcare pricing within integrated insurance markets.

The structural results show that VI raises both negotiated service prices and premiums, reducing consumer welfare. A higher level of integration increases internal transfers, which inflate reported MLRs. Stricter MLR thresholds induce insurers to increase premiums and negotiated service rates to maintain compliance. These findings suggest that VI helps insurers meet MLR requirements not through real efficiency gains, but through accounting and bargaining advantages that shift costs within the firm, ultimately harming consumers.

These findings have important policy implications. The results suggest that VI allows insurers to appear more compliant with MLR regulation without necessarily improving efficiency or reducing costs. This challenges the effectiveness of current oversight in achieving true consumer protection. To address this issue, regulators should pay closer attention to internal transfers between affiliated insurers and providers. Moreover, simply tightening the MLR threshold may not lead to better outcomes, as it could encourage price increases rather than real efficiency gains. Effective regulation should recognize the growing trend of VI and ensure that reported compliance reflects true improvements in consumer welfare, not accounting adjustments within integrated firms.

References

Abraham, J. M., P. Karaca-Mandic, and K. Simon (2014). How has the affordable care act's medical loss ratio regulation affected insurer behavior? *Medical care 52*(4), 370–377.

Angeles, J. and M. Bailit (2025). How insurers that own providers can game the medical loss ratio rules. *Health Affairs Forefront*.

Baker, L. C., M. K. Bundorf, and D. P. Kessler (2016). The effect of hospital/physician integration on hospital choice. *Journal of Health Economics 50*, 1–8.

Berry, S. T., J. A. Levinsohn, and A. Pakes (1993). Automobile prices in market equilibrium: Part i and ii.

Borusyak, K., X. Jaravel, and J. Spiess (2024). Revisiting event-study designs: robust and efficient estimation. *Review of Economic Studies 91*(6), 3253–3285.

Brot-Goldberg, Z. C., C. Che, and B. R. Handel (2022). Pharmacy benefit managers and vertical relationships in drug supply: state of current research. Technical report, National Bureau of Economic Research.

Callaway, B. and P. H. Sant'Anna (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics 225*(2), 200–230. Themed Issue: Treatment Effect 1.

Capps, C., D. Dranove, and C. Ody (2018). The effect of hospital acquisitions of physician practices on prices and spending. *Journal of health economics 59*, 139–152.

Cicala, S., E. M. Lieber, and V. Marone (2019). Regulating markups in us health insurance. *American Economic Journal: Applied Economics 11*(4), 71–104.

Crawford, G. S., R. S. Lee, M. D. Whinston, and A. Yurukoglu (2018). The welfare effects of vertical integration in multichannel television markets. *Econometrica 86*(3), 891–954.

Cuesta, J. I., C. E. Noton, and B. Vatter (2024). Vertical integration and plan design in healthcare markets. Technical report, National Bureau of Economic Research.

Encinosa, W. and A. Dor (2025). Does physician-hospital vertical integration signal care-coordination? evidence from mover-stayer analysis of commercially insured enrollees. *Journal of health economics 101*, 102997.

Fein, A. J. (2025). Mapping the vertical integration of insurers, pbms, specialty pharmacies, and providers. Available at drugchannels.net.

FierceHealthcare (2025). The perks and pitfalls of payer-provider vertical consolidation. Available at fiercehealthcare.com.

Frank, R. G. and C. Milhaupt (2022). Profits, medical loss ratios, and the ownership structure of medicare advantage plans. *Brookings Institution*.

Glickman, S. (2025). Gaming the system: Medical loss ratios and how insurers manipulate them. Available at substack.com.

Gray, C., A. E. Alpert, and N. Sood (2023). Disadvantaging rivals: Vertical integration in the pharmaceutical market. Technical report, National Bureau of Economic Research.

Grennan, M. (2013). Price discrimination and bargaining: Empirical evidence from medical devices. *American Economic Review 103*(1), 145–177.

Harrington, S. E. (2013). Medical loss ratio regulation under the affordable care act. *INQUIRY: The Journal of Health Care Organization, Provision, and Financing 50*(1), 9–26.

HealthCareDive (2024). More u.s. doctors are now employed by hospitals, payers, or corporations, report finds. Available at healthcaredive.com.

Koch, T. G., B. W. Wendling, and N. E. Wilson (2021). The effects of physician and hospital integration on medicare beneficiaries' health outcomes. *Review of Economics and Statistics 103*(4), 725–739.

Loewenstein, G., J. Y. Friedman, B. McGill, S. Ahmad, S. Linck, S. Sinkula, J. Beshears, J. J. Choi, J. Kolstad, D. Laibson, et al. (2013). Consumers' misunderstanding of health insurance. *Journal of Health Economics 32*(5), 850–862.

Mathews, A. W. (2017, December). Unitedhealth to buy large doctor group for $4.9 billion. *Wall Street Journal*. Available at wsj.com.

McCue, M., M. Hall, and X. Liu (2013). Impact of medical loss regulation on the financial performance of health insurers. *Health Affairs 32*(9), 1546–1551.

McCue, M. J. and M. Hall (2015). Health insurers' financial performance and quality improvement expenditures in the affordable care act's second year. *Medical Care Research and Review 72*(1), 113–122.

Norris, L. (2024). Billions in ACA rebates show 80/20 rule's impact. Available at healthinsurance.org.

PhRMA (2024). Explainer: Vertical integration may allow insurers to skirt Medical Loss Ratio requirements. *PhRMA*. Available at phrma.org.

Weber, S. (2024). 10% of us physicians work for or under unitedhealth. is that a problem? *Dermatology News 55*(1), 16–18.

Zhao, X. (2021). Medical loss ratio regulation and insurer pricing. Technical report, Working paper.

Appendix

A Proof of the Equilibrium Negotiated Price

This section derives the equilibrium negotiated service price p under bilateral Nash bargaining between the insurer and provider, subject to the insurer's MLR constraint and potential VI.

The negotiated price p solves the Nash product:

$$p = \arg\max_{p} \left(f^I - f_0^I\right)^{\tau} \left(f^H - f_0^H\right)^{1-\tau} \tag{A.1}$$

where $\tau \in [0, 1]$ denotes the insurer's bargaining power, f_0^I and f_0^H represent the respective disagreement payoffs in the event of bargaining failure, and (f^I, f^H) are the profits realized if agreement is reached. Consistent with the literature, I normalize the outside options to zero ($f_0^I = f_0^H = 0$), implying that the insurer exits the market and the provider treats patients covered by other insurers if no agreement is reached.

The F.O.C with respect to p for the logarithmic derivative of Nash product is:

$$\frac{\tau}{f^I}\left(\frac{\partial f^I}{\partial p} + \sum_j \frac{\partial f^I}{\partial \phi_j} \cdot \frac{\partial \phi_j}{\partial p}\right) + \frac{1-\tau}{f^H}\left\{\chi \sum_j \left[D_j + (p - c_h) \sum_j \frac{\partial D_j}{\partial \phi_k} \cdot \frac{\partial \phi_k}{\partial p}\right] + \lambda \cdot \frac{\partial L^I}{\partial p}\right\} = 0 \tag{A.2}$$

A.1 Envelope Theorem

Given optimal p^*, the insurer can get optimal ϕ based on the objective function:

$$
\begin{aligned}
f^I &= \sum_j D_j(\phi_j - p\gamma_j\chi - C_F) + \mu\left[\sum_j p\gamma_j\chi D_j - \bar{R}\sum_j \phi_j D_j\right] + \lambda\sum_j D_j(p - c_h)\chi \\
&= \sum_j (1-\mu\bar{R})\phi_j D_j - \sum_j (1-\mu)\gamma_j\chi p D_j - \sum_j C_F D_j + \lambda\sum_j D_j(p - c_h)\chi \\
&= \sum_j D_j\left[(1-\mu\bar{R})\phi_j - (1-\mu)p\gamma_j\chi - C_F\right] + \lambda\sum_j D_j(p - c_h)\chi
\end{aligned}
$$

Take total derivative of f^I with respect to ϕ_j:

$$
\begin{aligned}
\frac{\partial f^I}{\partial \phi_j} =& (1-\mu\bar{R})\left[D_j + \sum_k \phi_k \frac{\partial D_k}{\partial \phi_j}\right] - (1-\mu)p\chi\sum_k \gamma_k \frac{\partial D_k}{\partial \phi_j} - C_F\sum_k \frac{\partial D_k}{\partial \phi_j} \\
&+ \lambda\chi(p - c_h)\sum_k \frac{\partial D_k}{\partial \phi_j} = 0
\end{aligned}
$$

Then, I can get:

$$
\begin{aligned}
&(1-\mu\bar{R})D_j + \sum_k \left[(1-\mu\bar{R})\phi_k - (1-\mu)\gamma_k\chi p - C_F + \lambda\chi(p - c_h)\right]\frac{\partial D_k}{\partial \phi_j} = 0 \\
&\Rightarrow \sum_k \left[(1-\mu\bar{R})\phi_k - (1-\mu)\gamma_k\chi p - C_F\right]\cdot\frac{\partial D_k}{\partial \phi_j} + \sum_k \lambda\chi(p - c_h)\cdot\frac{\partial D_k}{\partial \phi_j} = -(1-\mu\bar{R})D_j
\end{aligned}
$$

I rewrite as the matrix form for total derivative of f^I with respect to $\boldsymbol{\phi}$:

$$
\begin{aligned}
&(1-\mu\bar{R})\boldsymbol{D} + \mathbf{J}^\top\left[(1-\mu\bar{R})\boldsymbol{\phi} - (1-\mu)\chi\boldsymbol{\gamma}\odot\boldsymbol{p} - C_F + \lambda\chi(p - c_h)\right] = 0 \\
&\Rightarrow \mathbf{J}^\top\left[(1-\mu\bar{R})\boldsymbol{\phi} - (1-\mu)\chi\boldsymbol{\gamma}\odot\boldsymbol{p} - C_F + \lambda\chi(p - c_h)\right] = -(1-\mu\bar{R})\boldsymbol{D} \qquad \text{(A.3)}
\end{aligned}
$$

Then, I can rearrange the matrix form:

$$
(1-\mu\bar{R})\boldsymbol{\phi} - (1-\mu)\boldsymbol{\gamma}\odot\boldsymbol{p}\chi - C_F + \lambda\chi(\boldsymbol{p} - c_h) = -(\mathbf{J}^\top)^{-1}(1-\mu\bar{R})\boldsymbol{D} \qquad \text{(A.4)}
$$

$$
(1-\mu\bar{R})\boldsymbol{\phi} - (1-\mu)\boldsymbol{\gamma}\odot\boldsymbol{p}\chi - C_F = -(\mathbf{J}^\top)^{-1}(1-\mu\bar{R})\boldsymbol{D} - \lambda\chi(\boldsymbol{p} - c_h) \qquad \text{(A.5)}
$$

Next, I rewrite the f^I as matrix form:

$$f^I = \boldsymbol{D}^\top \left[(1-\mu\bar{R})\boldsymbol{\phi} - (1-\mu)\boldsymbol{\gamma} \odot \boldsymbol{p}\chi - C_F - \lambda\chi(\boldsymbol{p} - c_h)\right] \tag{A.6}$$

I then substitute equation (A.4) to get the f^I:

$$f^I = \boldsymbol{D}^\top \cdot \left[-(\mathbf{J}^\top)^{-1}(1-\mu\bar{R})\boldsymbol{D}\right] = -(1-\mu\bar{R})\boldsymbol{D}^\top(\mathbf{J}^\top)^{-1}\boldsymbol{D} \tag{A.7}$$

A.2 First-order conditions with Respect to p

From equation (A.2), I can get:

$$\begin{aligned}
&\frac{\tau}{f^I}\cdot\frac{\partial f^I}{\partial p} + \frac{1-\tau}{f^H}\cdot\Bigg\{\chi\sum_j\left(D_j + (p-c_h)\sum_k\frac{\partial D_j}{\partial\phi_k}\cdot\frac{\partial\phi_k}{\partial p}\right) \\
&\qquad + \lambda\sum_j\Bigg[\left((1-\mu\bar{R})\frac{\partial\phi_j}{\partial p} - (1-\mu)\gamma_j\chi\right)D_j \\
&\qquad + \left((1-\mu\bar{R})\phi_j - (1-\mu)\gamma_j\chi p - C_F\right)\sum_k\frac{\partial D_j}{\partial\phi_k}\cdot\frac{\partial\phi_k}{\partial p}\Bigg]\Bigg\} = 0
\end{aligned}$$

$$\begin{aligned}
\Rightarrow\quad &\frac{\tau}{f^I}\left(-\sum_j\gamma_j\chi D_j + \mu\sum_j\gamma_j\chi D_j + \lambda\chi\sum_j D_j\right) \\
&+ \frac{1-\tau}{f^H}\Bigg\{\chi\sum_j D_j + \lambda\sum_j\left[\left((1-\mu\bar{R})\frac{\partial\phi_j}{\partial p} - (1-\mu)\gamma_j\chi\right)D_j\right] \\
&\quad + \left[\sum_j(p-c_h)\chi\sum_k\frac{\partial D_j}{\partial\phi_k}\frac{\partial\phi_k}{\partial p} + \lambda\sum_j\left((1-\mu\bar{R})\phi_j - (1-\mu)\gamma_j\chi p - C_F\right)\sum_k\frac{\partial D_j}{\partial\phi_k}\frac{\partial\phi_k}{\partial p}\right]\Bigg\} = 0
\end{aligned}$$

$$\Rightarrow \quad \frac{\tau}{f^I}\left[\sum_j \chi\left(\lambda-(1-\mu)\gamma_j\right)D_j\right]+\frac{(1-\tau)}{f^H}\left[\sum_j\left(\chi+\lambda(1-\mu\bar{R})\frac{\partial\phi_j}{\partial p}-\lambda(1-\mu)\gamma_j\chi\right)D_j\right]$$
$$+\frac{(1-\tau)\sum_j\left[(p-c_h)\chi+\lambda\left((1-\mu\bar{R})\phi_j-(1-\mu)\gamma_j\chi p-C_F\right)\right]\left(\sum_k\frac{\partial D_j}{\partial\phi_k}\cdot\frac{\partial\phi_k}{\partial p}\right)}{\sum_j D_j(p-c_h)\chi+\lambda\sum_j\left((1-\mu\bar{R})\phi_j-(1-\mu)\gamma_j\chi p-C_F\right)D_j}=0$$

$$\Rightarrow \quad \frac{\tau}{f^I}\left[\sum_j \chi\left(\lambda-(1-\mu)\gamma_j\right)D_j\right]+\frac{(1-\tau)\left[\sum_j\left(\chi+\lambda(1-\mu\bar{R})\frac{\partial\phi_j}{\partial p}-\lambda(1-\mu)\gamma_j\chi\right)D_j\right]}{\sum_j D_j(p-c_h)\chi+\lambda\sum_j\left((1-\mu\bar{R})\phi_j-(1-\mu)\gamma_j\chi p-C_F\right)D_j}$$
$$+\frac{(1-\tau)\sum_j\sum_k\frac{\partial D_j}{\partial\phi_k}\cdot\frac{\partial\phi_k}{\partial p}}{\sum_j D_j}=0$$

The matrix form is:

$$\frac{\tau\chi\boldsymbol{D}^\top[\lambda-(1-\mu)\boldsymbol{\gamma}]}{-(1-\mu\bar{R})\boldsymbol{D}^\top(\mathbf{J}^\top)^{-1}\boldsymbol{D}}+\frac{(1-\tau)\boldsymbol{D}^\top\left[\chi+\lambda(1-\mu\bar{R})\frac{\partial\boldsymbol{\phi}}{\partial p}-\lambda(1-\mu)\chi\boldsymbol{\gamma}\right]}{(p-c_h)\chi\mathbf{1}^\top\boldsymbol{D}+\lambda\boldsymbol{D}^\top\left[(1-\mu\bar{R})\boldsymbol{\phi}-(1-\mu)\chi\boldsymbol{\gamma}p-C_F\right]}+\frac{(1-\tau)\mathbf{1}^\top\mathbf{J}\frac{\partial\boldsymbol{\phi}}{\partial p}}{\mathbf{1}^\top\boldsymbol{D}}=0$$

Then, I substitute equation (A.5):

$$\frac{\tau\chi\boldsymbol{D}^\top[(1-\mu)\boldsymbol{\gamma}-\lambda]}{(1-\mu\bar{R})\boldsymbol{D}^\top(\mathbf{J}^\top)^{-1}\boldsymbol{D}}+\frac{(1-\tau)\boldsymbol{D}^\top\left[\chi+\lambda(1-\mu\bar{R})\frac{\partial\boldsymbol{\phi}}{\partial p}-\lambda(1-\mu)\chi\boldsymbol{\gamma}\right]}{(p-c_h)\chi\mathbf{1}^\top\boldsymbol{D}+\lambda\boldsymbol{D}^\top\left[-(\mathbf{J}^\top)^{-1}(1-\mu\bar{R})\boldsymbol{D}-\lambda\chi(\boldsymbol{p}-c_h)\right]}+\frac{(1-\tau)\mathbf{1}^\top\mathbf{J}\frac{\partial\boldsymbol{\phi}}{\partial p}}{\mathbf{1}^\top\boldsymbol{D}}=0$$

$$\Rightarrow \quad \frac{\tau\chi\boldsymbol{D}^\top[(1-\mu)\boldsymbol{\gamma}-\lambda]}{(1-\mu\bar{R})\boldsymbol{D}^\top(\mathbf{J}^\top)^{-1}\boldsymbol{D}}+\frac{(1-\tau)\boldsymbol{D}^\top\left[\chi+\lambda(1-\mu\bar{R})\frac{\partial\boldsymbol{\phi}}{\partial p}-\lambda(1-\mu)\chi\boldsymbol{\gamma}\right]}{(1-\lambda^2)\chi(p-c_h)\mathbf{1}^\top\boldsymbol{D}-\lambda(1-\mu\bar{R})\boldsymbol{D}^\top(\mathbf{J}^\top)^{-1}\boldsymbol{D}}+\frac{(1-\tau)\mathbf{1}^\top\mathbf{J}\frac{\partial\boldsymbol{\phi}}{\partial p}}{\mathbf{1}^\top\boldsymbol{D}}=0$$

Finally, I get:

$$\frac{(1-\tau)\boldsymbol{D}^\top\left[\chi+\lambda(1-\mu\bar{R})\frac{\partial\boldsymbol{\phi}}{\partial p}-\lambda\chi(1-\mu)\boldsymbol{\gamma}\right]}{(1-\lambda^2)\chi(p-c_h)\mathbf{1}^\top\boldsymbol{D}-\lambda(1-\mu\bar{R})\boldsymbol{D}^\top(\mathbf{J}^\top)^{-1}\boldsymbol{D}}=-\left[\frac{\tau\chi\boldsymbol{D}^\top[(1-\mu)\boldsymbol{\gamma}-\lambda]}{(1-\mu\bar{R})\boldsymbol{D}^\top(\mathbf{J}^\top)^{-1}\boldsymbol{D}}+\frac{(1-\tau)\mathbf{1}^\top\mathbf{J}\frac{\partial\boldsymbol{\phi}}{\partial p}}{\mathbf{1}^\top\boldsymbol{D}}\right] \tag{A.8}$$

I denote:

$$\begin{aligned}
A&=(1-\tau)\boldsymbol{D}^\top\left[\chi+\lambda(1-\mu\bar{R})\frac{\partial\boldsymbol{\phi}}{\partial p}-\lambda(1-\mu)\chi\boldsymbol{\gamma}\right]\\
B&=\lambda(1-\mu\bar{R})\boldsymbol{D}^\top(\mathbf{J}^\top)^{-1}\boldsymbol{D}\\
C&=(1-\lambda^2)\chi\mathbf{1}^\top\boldsymbol{D}
\end{aligned}$$

Define the RHS as:

$$R=-\left(\frac{\tau\chi\boldsymbol{D}^\top[(1-\mu)\boldsymbol{\gamma}-\lambda]}{(1-\mu\bar{R})\boldsymbol{D}^\top(\mathbf{J}^\top)^{-1}\boldsymbol{D}}+\frac{(1-\tau)\mathbf{1}^\top\mathbf{J}\frac{\partial\boldsymbol{\phi}}{\partial p}}{\mathbf{1}^\top\boldsymbol{D}}\right)$$

Now multiply both sides of the equation by the denominator:

$$A=[C(p-c_h)-B]\cdot R$$

Now solve for p:

$$p=c_h+\frac{1}{C}\left(\frac{A}{R}+B\right)$$

Then, I can get:

$$p = c_h + \frac{1}{(1-\lambda^2)\chi \mathbf{1}^\top \boldsymbol{D}} \left[\begin{array}{l} \dfrac{(1-\tau)\boldsymbol{D}^\top \left[\chi + \lambda(1-\mu\bar{R})\frac{\partial \boldsymbol{\phi}}{\partial p} - \lambda(1-\mu)\chi\boldsymbol{\gamma}\right]}{(1-\mu\bar{R})\boldsymbol{D}^\top(\mathbf{J}^\top)^{-1}\boldsymbol{D}} + \lambda(1-\mu\bar{R})\boldsymbol{D}^\top(\mathbf{J}^\top)^{-1}\boldsymbol{D} \\ - \left(\dfrac{\tau\chi \boldsymbol{D}^\top \left[(1-\mu)\boldsymbol{\gamma} - \lambda\right]}{(1-\mu\bar{R})\boldsymbol{D}^\top(\mathbf{J}^\top)^{-1}\boldsymbol{D}} + \dfrac{(1-\tau)\mathbf{1}^\top \mathbf{J}\dfrac{\partial \boldsymbol{\phi}}{\partial p}}{\mathbf{1}^\top \boldsymbol{D}} \right) \end{array} \right]$$

Finally, I get the equilibrium negotiated price:

$$p = c_h + \underbrace{\frac{\lambda(1-\mu\bar{R})\,\boldsymbol{D}^\top(\mathbf{J}^\top)^{-1}\boldsymbol{D}}{(1-\lambda^2)\,\chi\,\mathbf{1}^\top\boldsymbol{D}}}_{\text{first term}}$$

$$- \underbrace{\frac{\boldsymbol{D}^\top\left[\chi\mathbf{1} + \lambda(1-\mu\bar{R})\,\frac{\partial\boldsymbol{\phi}}{\partial p} - \lambda(1-\mu)\chi\boldsymbol{\gamma}\right]}{(1-\lambda^2)\chi\,\mathbf{1}^\top\boldsymbol{D}}}_{\text{second term}}$$

$$\times \underbrace{\left[\frac{\mathbf{1}^\top\mathbf{J}\,\frac{\partial\boldsymbol{\phi}}{\partial p}}{\mathbf{1}^\top\boldsymbol{D}} + \frac{\tau\chi}{(1-\tau)(1-\mu\bar{R})}\frac{\boldsymbol{D}^\top\left[(1-\mu)\boldsymbol{\gamma} - \lambda\right]}{\boldsymbol{D}^\top(\mathbf{J}^\top)^{-1}\boldsymbol{D}}\right]^{-1}}_{\text{third term}}$$

B Calculation of ϕ'

Given optimal p^*, the insurer can get optimal ϕ based on the objective function:

$$\begin{aligned} f^I &= \sum_j D_j(\phi_j - p\gamma_j\chi - C_F) + \mu\left[\sum_j p\gamma_j\chi D_j - \bar{R}\sum_j \phi_j D_j\right] + \lambda\sum_j D_j(p-c_h)\chi \\ &= \sum_j (1-\mu\bar{R})\phi_j D_j - \sum_j (1-\mu)\gamma_j\chi p D_j - \sum_j C_F D_j + \lambda\sum_j D_j(p-c_h)\chi \\ &= \sum_j D_j\left[(1-\mu\bar{R})\phi_j - (1-\mu)p\gamma_j\chi - C_F\right] + \lambda\sum_j D_j(p-c_h)\chi \end{aligned}$$

Take total derivative of f^I with respect to ϕ_j:

$$\begin{aligned} \frac{\partial f^I}{\partial \phi_j} =& (1-\mu\bar{R})\left[D_j + \sum_k \phi_k \frac{\partial D_k}{\partial \phi_j}\right] - (1-\mu)p\chi\sum_k \gamma_k \frac{\partial D_k}{\partial \phi_j} - C_F\sum_k \frac{\partial D_k}{\partial \phi_j} \\ &+ \lambda\chi(p-c_h)\sum_k \frac{\partial D_k}{\partial \phi_j} = 0 \end{aligned}$$

Then, I can get:

$$\begin{aligned} &(1-\mu\bar{R})D_j + \sum_k\left[(1-\mu\bar{R})\phi_k - (1-\mu)\gamma_k\chi p - C_F + \lambda\chi(p-c_h)\right]\frac{\partial D_k}{\partial \phi_j} = 0 \\ &\Rightarrow \sum_k\left[(1-\mu\bar{R})\phi_k - (1-\mu)\gamma_k\chi p - C_F\right]\cdot\frac{\partial D_k}{\partial \phi_j} + \sum_k \lambda\chi(p-c_h)\cdot\frac{\partial D_k}{\partial \phi_j} = -(1-\mu\bar{R})D_j \end{aligned}$$

Then, for each $j \in \{1, ..., J\}$

$$\sum_{k=1}^J \phi_k \cdot \frac{\partial D_k}{\partial \phi_j} + D_j = \frac{1-\mu}{1-\mu\bar{R}}\chi p\sum_{k=1}^J \gamma_k \cdot \frac{\partial D_k}{\partial \phi_j} + \frac{C_F}{1-\mu\bar{R}}\sum_{k=1}^J \frac{\partial D_k}{\partial \phi_j} - \frac{\lambda\chi}{1-\mu\bar{R}}(p-c_h)\sum_{k=1}^J \frac{\partial D_k}{\partial \phi_j}$$

I take derivative w.r.t. p:

$$\sum_{k=1}^{J}\frac{d\phi_k}{dp}\frac{\partial D_k}{\partial \phi_j} + \sum_{k=1}^{J}\phi_k\sum_{\ell=1}^{J}\frac{\partial^2 D_k}{\partial\phi_j\,\partial\phi_\ell}\frac{d\phi_\ell}{dp} + \sum_{k=1}^{J}\frac{\partial D_j}{\partial\phi_k}\frac{d\phi_k}{dp} = \frac{1-\mu}{1-\mu\bar{R}}\chi\sum_{k=1}^{J}\gamma_k\frac{\partial D_k}{\partial\phi j}$$
$$+\frac{1-\mu}{1-\mu\bar{R}}\chi p\sum_{k=1}^{J}\gamma_k\sum_{\ell=1}^{J}\frac{\partial^2 D_k}{\partial\phi_j\,\partial\phi_\ell}\frac{d\phi_\ell}{dp} + \frac{C_F}{1-\mu\bar{R}}\sum_{k=1}^{J}\sum_{\ell=1}^{J}\frac{\partial^2 D_k}{\partial\phi_j\,\partial\phi_\ell}\frac{d\phi_\ell}{dp}$$
$$-\frac{\lambda\chi}{1-\mu\bar{R}}\sum_{k=1}^{J}\frac{\partial D_k}{\partial\phi_j} - \frac{\lambda\chi}{1-\mu\bar{R}}(p-c_h)\sum_{k=1}^{J}\sum_{\ell=1}^{J}\frac{\partial^2 D_k}{\partial\phi_j\,\partial\phi_\ell}\frac{d\phi_\ell}{dp}$$

Then, I can get:

$$\sum_{k=1}^{J}\frac{d\phi_k}{dp}\frac{\partial D_k}{\partial \phi_j} + \sum_{k=1}^{J}\phi_k\sum_{\ell=1}^{J}\frac{\partial^2 D_k}{\partial\phi_j\,\partial\phi_\ell}\frac{d\phi_\ell}{dp} + \sum_{k=1}^{J}\frac{\partial D_j}{\partial\phi_k}\frac{d\phi_k}{dp}$$
$$=\frac{1-\mu}{1-\mu\bar{R}}\chi\left[\sum_{k=1}^{J}\gamma_k\frac{\partial D_k}{\partial\phi_j} + p\sum_{k=1}^{J}\gamma_k\sum_{\ell=1}^{J}\frac{\partial^2 D_k}{\partial\phi_j\,\partial\phi_\ell}\frac{d\phi_\ell}{dp}\right]$$
$$-\frac{\lambda\chi}{1-\mu\bar{R}}\sum_{k=1}^{J}\frac{\partial D_k}{\partial\phi_j} + \left[\frac{C_F}{1-\mu\bar{R}} - \frac{\lambda\chi}{1-\mu\bar{R}}(p-c_h)\right]\sum_{k=1}^{J}\sum_{\ell=1}^{J}\frac{\partial^2 D_k}{\partial\phi_j\,\partial\phi_\ell}\frac{d\phi_\ell}{dp}$$

Then, for plan j, the matrix form is:

$$\mathbf{J}_j^\top\boldsymbol{\phi}'(p) + \boldsymbol{\phi}^T\boldsymbol{H}_j\boldsymbol{\phi}'(p) + \mathbf{J}_j\boldsymbol{\phi}'(p) = \frac{1-\mu}{1-\mu\bar{R}}\mathbf{J}_j^\top\boldsymbol{\gamma} + \frac{1-\mu}{1-\mu\bar{R}}\chi p\left(\boldsymbol{\gamma}^T\boldsymbol{H}_j\boldsymbol{\phi}'(p)\right) + \frac{C_F}{1-\mu\bar{R}}\mathbf{1}^T\boldsymbol{H}_j\boldsymbol{\phi}'(p)$$
$$-\frac{\lambda\chi}{1-\mu\bar{R}}\mathbf{J}_j^\top\mathbf{1} - \frac{\lambda\chi}{1-\mu\bar{R}}(p-c_h)\mathbf{1}^T\boldsymbol{H}_j\boldsymbol{\phi}'(p)$$

where $\mathbf{J}_j^\top$ is the j^{th} row of $\mathbf{J}^\top$,$\boldsymbol{H}_j$ is the Hessian matrix of plan j's demand function

I can rearrange the matrix form:

$$\left[\frac{1-\mu}{1-\mu\bar{R}}\chi p\boldsymbol{\gamma}^T + \frac{C_F}{1-\mu\bar{R}}\mathbf{1}^T - \frac{\lambda\chi}{1-\mu\bar{R}}(p-c_h)\mathbf{1}^T - \boldsymbol{\phi}^T\right]\boldsymbol{H}_j\boldsymbol{\phi}'(p)$$
$$= \mathbf{J}_j^\top\boldsymbol{\phi}'(p) + \mathbf{J}_j\boldsymbol{\phi}'(p) - \frac{1-\mu}{1-\mu\bar{R}}\mathbf{J}_j^\top\boldsymbol{\gamma} + \frac{\lambda\chi}{1-\mu\bar{R}}\mathbf{J}_j^\top\mathbf{1} \qquad \text{(B.1)}$$

Based on equation (A.3), I can get:

$$\frac{1-\mu}{1-\mu\bar{R}}\chi p\boldsymbol{\gamma}^T + \frac{C_F}{1-\mu\bar{R}}\mathbf{1}^T - \frac{\lambda\chi}{1-\mu\bar{R}}(p-c_h)\mathbf{1}^T - \boldsymbol{\phi}^T = \left[(\mathbf{J}^\top)^{-1}\boldsymbol{D}\right]^T$$

Then, I can get for each plan j:

$$\begin{aligned}
&\mathbf{J}_j^\top\boldsymbol{\phi}'(p) + \mathbf{J}_j\boldsymbol{\phi}'(p) - \frac{1-\mu}{1-\mu\bar{R}}\mathbf{J}_j^\top\boldsymbol{\gamma} + \frac{\lambda\chi}{1-\mu\bar{R}}\mathbf{J}_j^\top\mathbf{1} = \left[(\mathbf{J}^\top)^{-1}\boldsymbol{D}\right]^T\boldsymbol{H}_j\boldsymbol{\phi}'(p)\\
\Rightarrow\quad &\left[(\mathbf{J}^\top)^{-1}\boldsymbol{D}\right]^T\boldsymbol{H}_j\boldsymbol{\phi}'(p) - \mathbf{J}_j^\top\boldsymbol{\phi}'(p) - \mathbf{J}_j\boldsymbol{\phi}'(p) = -\frac{1-\mu}{1-\mu\bar{R}}\mathbf{J}_j^\top\boldsymbol{\gamma} + \frac{\lambda\chi}{1-\mu\bar{R}}\mathbf{J}_j^\top\mathbf{1}\\
\Rightarrow\quad &\left[\left[(\mathbf{J}^\top)^{-1}\boldsymbol{D}\right]^T\boldsymbol{H}_j - \mathbf{J}_j^\top - \mathbf{J}_j\right]\boldsymbol{\phi}'(p) = \mathbf{J}_j^\top\left[-\frac{1-\mu}{1-\mu\bar{R}}\boldsymbol{\gamma} + \frac{\lambda\chi}{1-\mu\bar{R}}\mathbf{1}\right]
\end{aligned}$$

Finally, I get for $j \in 1, 2..., J$:

$$\left[\left[(\mathbf{J}^\top)^{-1}\boldsymbol{D}\right]^T\mathbf{H}_j - \mathbf{J}^\top - \mathbf{J}\right]\boldsymbol{\phi}'(p) = \mathbf{J}^\top\left[-\frac{1-\mu}{1-\mu\bar{R}}\boldsymbol{\gamma} + \frac{\lambda\chi}{1-\mu\bar{R}}\mathbf{1}\right]$$

Finally, I solve for the derivative of $\boldsymbol{\phi}'$

$$\boldsymbol{\phi}'(p) = \mathbf{A}^{-1}\mathbf{J}^\top\mathbf{B} \tag{B.2}$$

where

$$\mathbf{A} = \left[(\mathbf{J}^\top)^{-1}\boldsymbol{D}\right]^\top\mathbf{H}_j - \mathbf{J}^\top - \mathbf{J}, \qquad \mathbf{B} = \left[-\frac{1-\mu}{1-\mu\bar{R}}\boldsymbol{\gamma} + \frac{\lambda\chi}{1-\mu\bar{R}}\mathbf{1}\right]$$

C Solve for c_h

Firstly, I substitute equation (B.2) to equation From equation (A.8) to get:

$$\frac{(1-\tau)\boldsymbol{D}^\top\Big[\chi+\lambda(1-\mu\bar{R})\mathbf{A}^{-1}\mathbf{J}^\top\mathbf{B}-\lambda\chi(1-\mu)\boldsymbol{\gamma}\Big]}{(1-\lambda^2)\chi(p-c_h)\mathbf{1}^\top\boldsymbol{D}-\lambda(1-\mu\bar{R})\boldsymbol{D}^\top(\mathbf{J}^\top)^{-1}\boldsymbol{D}}= -\left[\frac{\tau\chi\boldsymbol{D}^\top\big((1-\mu)\boldsymbol{\gamma}-\lambda\mathbf{1}\big)}{(1-\mu\bar{R})\boldsymbol{D}^\top(\mathbf{J}^\top)^{-1}\boldsymbol{D}}+\frac{(1-\tau)}{\mathbf{1}^\top\boldsymbol{D}}\mathbf{1}^\top\mathbf{J}\mathbf{A}^{-1}\mathbf{J}^\top\mathbf{B}\right] \tag{C.1}$$

Then, I will eliminate p using the premium equation from insurer's problem. I rewrite the premium equation (A.3) as:

$$\left[\frac{(1-\mu)}{1-\mu\bar{R}}\chi\boldsymbol{\gamma}-\frac{\lambda\chi\mathbf{1}}{1-\mu\bar{R}}\right]p=\boldsymbol{\phi}+(\mathbf{J}^\top)^{-1}\boldsymbol{D}-\frac{C_F\mathbf{1}}{1-\mu\bar{R}}-\frac{\lambda\chi}{1-\mu\bar{R}}c_h$$
$$p-c_h=S^{-1}\left[\boldsymbol{\phi}+(\mathbf{J}^\top)^{-1}\boldsymbol{D}-\frac{C_F\mathbf{1}}{1-\mu\bar{R}}-\frac{\lambda\chi}{1-\mu\bar{R}}c_h\right]-c_h$$

Rearranging, I obtain:

$$p-c_h=S^{-1}V-\kappa c_h \tag{C.2}$$

where

$$S=\frac{(1-\mu)}{1-\mu\bar{R}}\chi\boldsymbol{\gamma}-\frac{\lambda\chi\mathbf{1}}{1-\mu\bar{R}},\qquad V=\boldsymbol{\phi}+(\mathbf{J}^\top)^{-1}\boldsymbol{D}-\frac{C_F\mathbf{1}}{1-\mu\bar{R}},\qquad \kappa=1+\frac{S^{-1}\chi\lambda}{1-\mu\bar{R}}$$

Then, I substitute equation (C.2) into equation (C.1):

$$\frac{(1-\tau)\Big[\chi\,\mathbf{1}^\top\boldsymbol{D}-\lambda\chi(1-\mu)\,\boldsymbol{D}^\top\boldsymbol{\gamma}+\lambda(1-\mu\bar{R})\,\boldsymbol{D}^\top\mathbf{A}^{-1}\mathbf{J}^\top\mathbf{B}\Big]}{(1-\lambda^2)\chi\,(S^{-1}V-\kappa c_h)\,\mathbf{1}^\top\boldsymbol{D}-\lambda(1-\mu\bar{R})\,\boldsymbol{D}^\top(\mathbf{J}^\top)^{-1}\boldsymbol{D}}= -\left[\frac{\tau\chi\,\boldsymbol{D}^\top\big((1-\mu)\boldsymbol{\gamma}-\lambda\mathbf{1}\big)}{(1-\mu\bar{R})\,\boldsymbol{D}^\top(\mathbf{J}^\top)^{-1}\boldsymbol{D}}+\frac{(1-\tau)}{\mathbf{1}^\top\boldsymbol{D}}\mathbf{1}^\top\mathbf{J}\mathbf{A}^{-1}\mathbf{J}^\top\mathbf{B}\right]. \tag{C.3}$$

Now I Define:

$$\Theta_D \equiv (1-\tau)\Big[\chi\,\mathbf{1}^\top \boldsymbol{D} - \lambda\chi(1-\mu)\,\boldsymbol{D}^\top\boldsymbol{\gamma} + \lambda(1-\mu\bar{R})\,\boldsymbol{D}^\top\mathbf{A}^{-1}\mathbf{J}^\top\mathbf{B}\Big],$$

$$\Psi_D \equiv \frac{\tau\chi\,\boldsymbol{D}^\top\big((1-\mu)\boldsymbol{\gamma} - \lambda\mathbf{1}\big)}{(1-\mu\bar{R})\,\boldsymbol{D}^\top(\mathbf{J}^\top)^{-1}\boldsymbol{D}} + \frac{(1-\tau)}{\mathbf{1}^\top\boldsymbol{D}}\mathbf{1}^\top\mathbf{J}\mathbf{A}^{-1}\mathbf{J}^\top\mathbf{B}, \qquad \Omega_D = \boldsymbol{D}^\top(\mathbf{J}^\top)^{-1}\boldsymbol{D}.$$

Finally, I solve for c_h:

$$c_h = \frac{S^{-1}V - Q}{\kappa} \tag{C.4}$$

where

$$Q = \frac{\lambda(1-\mu\bar{R})\,\Omega_D - \Theta_D/\Psi_D}{(1-\lambda^2)\,\chi\,\mathbf{1}^\top\mathbf{D}}$$

Equation (C.4) provides the closed-form expression for the provider's effective cost c_h as a function of model primitives $(\lambda, \tau, \mu, C_F, \chi, \gamma)$ and equilibrium observables $(\mathbf{J}, \mathbf{H}, \boldsymbol{\phi}, \mathbf{D}, \bar{R})$.

D Counterfactual Prices

The general equilibrium price expression is given by:

$$p = c_h + \underbrace{\frac{\lambda(1-\mu\bar{R})\,\boldsymbol{D}^\top(\mathbf{J}^\top)^{-1}\boldsymbol{D}}{(1-\lambda^2)\,\chi\,\mathbf{1}^\top\boldsymbol{D}}}_{\text{first term}}$$
$$\underbrace{-\frac{\boldsymbol{D}^\top\Big[\chi\mathbf{1}+\lambda(1-\mu\bar{R})\,\frac{\partial\boldsymbol{\phi}}{\partial p}-\lambda(1-\mu)\chi\boldsymbol{\gamma}\Big]}{(1-\lambda^2)\chi\,\mathbf{1}^\top\boldsymbol{D}}}_{\text{second term}}$$
$$\times\underbrace{\left[\frac{\mathbf{1}^\top\mathbf{J}\,\frac{\partial\boldsymbol{\phi}}{\partial p}}{\mathbf{1}^\top\boldsymbol{D}}+\frac{\tau\chi}{(1-\tau)(1-\mu\bar{R})}\frac{\boldsymbol{D}^\top\big[(1-\mu)\boldsymbol{\gamma}-\lambda\big]}{\boldsymbol{D}^\top(\mathbf{J}^\top)^{-1}\boldsymbol{D}}\right]^{-1}}_{\text{third term}}$$

D.1 No VI: $\lambda = 0$

$$p^{\text{NoVI}} = c_h - \frac{\boldsymbol{D}^\top}{\mathbf{1}^\top\boldsymbol{D}}\left[\frac{\mathbf{1}^\top\mathbf{J}\frac{\partial\phi}{\partial p}}{\mathbf{1}^\top\boldsymbol{D}}+\frac{\tau\chi}{(1-\tau)(1-\mu\bar{R})}\frac{\boldsymbol{D}^\top\big((1-\mu)\boldsymbol{\gamma}\big)}{\boldsymbol{D}^\top(\mathbf{J}^\top)^{-1}\boldsymbol{D}}\right]^{-1}.$$

D.2 Higher VI: $\lambda = 0.9$

$$p^{VI=0.9} = c_h + \frac{0.9\,(1-\mu\bar{R})\,\mathbf{D}^\top(\mathbf{J}^\top)^{-1}\mathbf{D}}{0.19\,\chi\,\mathbf{1}^\top\mathbf{D}} - \frac{\mathbf{D}^\top\Big[\chi\mathbf{1}+0.9(1-\mu\bar{R})\,\frac{\partial\boldsymbol{\phi}}{\partial p}-0.9(1-\mu)\chi\,\boldsymbol{\gamma}\Big]}{0.19\,\chi\,\mathbf{1}^\top\mathbf{D}}$$
$$\times\left\{\frac{\mathbf{1}^\top\mathbf{J}\,\frac{\partial\boldsymbol{\phi}}{\partial p}}{\mathbf{1}^\top\mathbf{D}}+\frac{\tau\chi}{(1-\tau)(1-\mu\bar{R})}\,\frac{\mathbf{D}^\top\big[(1-\mu)\boldsymbol{\gamma}-0.9\big]}{\mathbf{D}^\top(\mathbf{J}^\top)^{-1}\mathbf{D}}\right\}^{-1}$$

D.3 No MLR Regulation: $\mu = 0$

$$p^{\mu=0} = c_h + \frac{\lambda\,\boldsymbol{D}^\top(\mathbf{J}^\top)^{-1}\boldsymbol{D}}{(1-\lambda^2)\,\chi\,\mathbf{1}^\top\boldsymbol{D}} - \frac{\boldsymbol{D}^\top\Big[\chi+\lambda\frac{\partial\phi}{\partial p}-\lambda\chi\boldsymbol{\gamma}\Big]}{(1-\lambda^2)\,\chi\,\mathbf{1}^\top\boldsymbol{D}}$$
$$\times\left[\frac{\mathbf{1}^\top\mathbf{J}\frac{\partial\phi}{\partial p}}{\mathbf{1}^\top\boldsymbol{D}}+\frac{\tau\chi}{(1-\tau)}\frac{\boldsymbol{D}^\top\big(\boldsymbol{\gamma}-\lambda\big)}{\boldsymbol{D}^\top(\mathbf{J}^\top)^{-1}\boldsymbol{D}}\right]^{-1}.$$

D.4 Stricter MLR Threshold: $\bar{R} = 0.9$

$$p^{\bar{R}=0.9} = c_h + \frac{\lambda(1-0.9\,\mu)\,\boldsymbol{D}^\top(\mathbf{J}^\top)^{-1}\boldsymbol{D}}{(1-\lambda^2)\,\chi\,\mathbf{1}^\top\boldsymbol{D}} - \frac{\boldsymbol{D}^\top\Big[\chi\Big(1+\lambda(1-0.9\,\mu)\frac{\partial\phi}{\partial p}\Big)-\lambda(1-\mu)\chi\boldsymbol{\gamma}\Big]}{(1-\lambda^2)\,\chi\,\mathbf{1}^\top\boldsymbol{D}}$$
$$\times\left[\frac{\mathbf{1}^\top\mathbf{J}\frac{\partial\phi}{\partial p}}{\mathbf{1}^\top\boldsymbol{D}} + \frac{\tau\chi}{(1-\tau)(1-0.9\,\mu)}\frac{\boldsymbol{D}^\top\big((1-\mu)\boldsymbol{\gamma}-\lambda\big)}{\boldsymbol{D}^\top(\mathbf{J}^\top)^{-1}\boldsymbol{D}}\right]^{-1}$$