The No Surprises Act's Surprises: Arbitration, Provider Network, and Insurer Pricing*

Panle Jia Barwick† Anran Li‡ Tianli Xia§ Shizhe Yu¶

August 31, 2026

Abstract

The No Surprises Act (NSA) capped patients' liability for surprise bills and established federal final-offer arbitration for insurer–provider payment disputes. Linking arbitration records to the national longitudinal provider-insurer network panel and plan design data in the ACA marketplace, we find that the NSA reduced provider network participation in the specialties and states most affected by the reform, contrary to the standard prediction that removing balance billing makes network participation more attractive. The decline is consistent with arbitration strengthening providers' outside option: arbitration is high-volume and provider-favorable, with awards well above the statutory benchmark; and network participation declines after provider groups' first favorable arbitration outcomes. Suggestive evidence indicates that insurers more exposed to arbitration raise premiums. These results reveal an unintended consequence of the NSA: the law's direct consumer protections may be accompanied by narrower networks and higher premiums.

JEL Codes: I11, I13, I18, L14, D82
Keywords: No Surprises Act, arbitration, provider networks, health insurance, surprise billing

*We thank Florian Ederer for his insightful discussion, Jaden Chen, Jacob Dorn, Leemore Dafny, Joe Harrington, Fiona Scott Morton, Ellie Prager, and participants at seminars and conferences for their helpful comments.

†Barwick: Department of Economics, University of Wisconsin-Madison, NBER, and CEPR, pbarwick@wisc.edu.
‡Li: Department of Economics, University of Minnesota-Twin Cities and NBER, anranli@umn.edu.
§Xia: Simon Business School, University of Rochester, tianli.xia@rochester.edu.
¶Yu: Department of Economics, University of Wisconsin-Madison, syu334@wisc.edu.

1. Introduction

Prior to 2022, millions of Americans faced "surprise" medical bills when they unknowingly received care from out-of-network providers, particularly in emergency settings where patients have no meaningful choice of where to seek treatment. These bills frequently exceeded in-network rates by several multiples, generating financial hardship and contributing to the medical-debt crisis (Cooper et al., 2020a). The No Surprises Act (NSA), effective 2022, addressed this problem by capping patient cost-sharing at in-network levels for emergency and certain ancillary services. But the law did not set insurer–provider payments directly. Instead, it created federal independent dispute resolution (IDR), a final-offer arbitration process, to determine payments when insurers and providers fail to agree.

This design creates a potential tradeoff. By eliminating balance billing, the NSA makes out-of-network status less attractive: providers can no longer bill protected patients for the residual amount. At the same time, when arbitration awards favor providers, NSA not only transfers payments from insurers to providers but also improves providers' outside option in negotiations. Arbitration may therefore become more than a backstop for failed negotiations: providers may have weaker incentives to accept network contracts and bargain higher out-of-network payments, while insurers may pass the resulting costs on to consumers through higher premiums.[1] The predicted network response is therefore ambiguous. We ask how the NSA changed network participation, and whether the evidence favors the balance-billing-ban channel or the arbitration-option channel.

We study this question by linking federal IDR records to a national provider–insurer network panel and plan-design data from the ACA marketplace in 2017-2025. These data allow us to estimate the NSA's effect on provider networks and then examine whether observed arbitration and pricing patterns are consistent with an outside-option interpretation. Our analysis exploits three sources of variation: NSA-exposed specialties (emergency medicine, radiology, and anesthesiology) versus unexposed ones (primary care, general surgery, psychiatry, and other specialties); states without prior surprise-billing protections versus states with pre-existing laws similar to the NSA; and pre- versus post-NSA changes over time. We document four sets of findings.

Our first finding is that the NSA reduced provider network participation in the specialties and states most affected by the reform. Relative to states with comparable pre-existing protections and to specialties not covered by the law, in-network participation declines by 4.7 percentage points for emergency medicine, 3.4 percentage points for anesthesiology, and 1.8 percentage points for radiology. These estimates correspond to 13.5, 9.7, and 3.5 percent of their respective pre-NSA means. The pre-period coefficients are flat, and the decline emerges roughly a year after the NSA

[1] For contemporaneous media coverage, see Sarah Kliff and Margot Sanger-Katz, "A $440,000 Breast Reduction: How Doctors Cashed In on a Consumer Protection Law," *The New York Times*, April 22, 2026, https://www.nytimes.com/2026/04/22/us/politics/doctors-insurers-arbitration.html.

took effect, a lag consistent with the time needed for early disputes to resolve and for physicians and insurers to adjust. Placebo specialties, including psychiatry, primary care, and general surgery, do not show comparable responses in network participation. The estimates are also robust to alternative treatments of provider-directory gaps and large updates, and the exclusion of the largest IDR-filing staffing firms. This response runs counter to the conventional prediction that eliminating balance billing should make network participation more attractive.

We next present two pieces of complementary evidence consistent with arbitration strengthening providers' outside options. First, federal IDR rapidly became a high-volume and predominantly provider-favorable payment channel, exceeding 1.8 million cases cumulatively by 2025, five to ten times that regulators initially projected. Providers prevail in approximately 86 percent of resolved cases, and selected payments are frequently well above both insurer offers and the statutory Qualifying Payment Amount (QPA). After accounting for administrative and arbitration fees, providers' net gains from arbitration are positive for a large share of disputes and average $117,472 per participating provider-year. Although these statistics do not show that arbitration dominates every provider's pre-NSA outside option, they suggest that federal IDR can be financially meaningful for participating providers, giving at least some an incentive to treat IDR as a substitute for in-network contracting.

Second, network behavior responds when the arbitration option's value is realized. We use each provider group's first favorable IDR determination as an event and follow its network relationships with insurers. Following a first favorable outcome, NSA-exposed groups' in-network share falls by 6.8 percentage points (19.8 percent of the pre-event mean), and the fraction of groups entirely out-of-network for an insurer rises by 7.7 percentage points. The pre-event path is flat, and non-NSA-covered provider groups in the same ZIP codes exhibit no analogous change when assigned the same placebo event dates. This timing is consistent with providers learning about, or acting on, IDR as a valuable outside option. Nevertheless, because groups might select into arbitration, we read this timing as mechanism evidence consistent with providers learning about, or acting on, the value of IDR, rather than as the causal effect of a favorable determination on network exit.

Finally, we present suggestive evidence on insurer pricing. We construct an ex ante measure of arbitration exposure from insurers' pre-period contracting relationships with four physician-staffing organizations that subsequently became the largest federal IDR filers. Within the same state and year, issuers with greater predicted exposure experience larger post-NSA premium increases. The pooled estimates correspond to approximately $15 per month for a 27-year-old and $27 per month for a 50-year-old, or 3.6 to 3.8 percent of baseline premiums. Estimated changes in deductibles are positive but less precise. We interpret these results as suggestive evidence of an indirect pricing response, not as a causal estimate of arbitration-cost pass-through.

These results show that the NSA's direct and indirect effects can point in opposite directions.

The law protects patients from the visible shock of a surprise bill. But the arbitration-option mechanism changes provider–insurer bargaining, and those changes can reappear as narrower networks, higher premiums, and higher public subsidy outlays. We use a simple scale calculation to put these unintended consequences in perspective. Existing estimates imply direct balance-billing protection on the order of $55 per enrollee per year. Our premium estimates for high-exposure issuers are large relative to this benchmark. This comparison is not a net welfare calculation: our premium estimates are differential effects for more-exposed issuers, gross premiums do not map one-for-one into enrollee payments in the subsidized ACA Marketplace, and we do not monetize access, provider choice, or continuity-of-care costs from narrower networks. Still, the magnitudes indicate that the indirect cost channel is economically meaningful.

This paper proceeds as follows. Section 2 describes the institutional setting and data. Section 3 presents the paper's primary estimates of the NSA's effect on provider network participation. Section 4 assembles complementary evidence on the arbitration-option mechanism, first documenting federal IDR use and payment outcomes and then examining network participation around provider groups' first favorable determinations. Section 5 presents suggestive evidence on insurer pricing. Section 6 discusses implications.

Related Literature This paper contributes to the literature on surprise billing and its regulation. Prior work documents the prevalence and strategic use of surprise billing (Cooper and Scott Morton, 2016; Garmon and Chartock, 2017; Cooper et al., 2020a,b). Existing studies on surprise-billing regulation largely focus on state-level protections (Adler et al., 2019; Chartock et al., 2021; Duffy et al., 2022a; Garmon et al., 2024a,b). The NSA marks a sharp change in both scale and institutional design: it extended protections nationwide and created a federal arbitration system that exceeds the scope of state regulation. Recent work begins to study the federal IDR process. Fiedler and Adler (2024) and Duffy et al. (2024) document dispute outcomes; Chartock and Whaley (2025) show a standard final-offer arbitration model (Farber, 1980; Ashenfelter and Bloom, 1984) struggles to rationalize the observed dispute offers and determinations. Whereas these studies examine outcomes and behavior within disputes, we study the NSA's market consequences for provider network participation and, more suggestively, insurer premiums.

We also contribute to the literature on provider–insurer bargaining and network formation. Existing work primarily uses structural models to quantify how bargaining and other contracting incentives determine network breadth (Gowrisankaran et al., 2015; Ho and Lee, 2017; Shepard, 2022; Serna, 2024). Most closely related to our mechanism, Prager and Tilipman (2025) use counterfactual simulations to quantify how regulating out-of-network payments shifts disagreement payoffs and affects negotiated hospital prices and access. Our paper instead exploits quasi-random variation generated by the NSA, which shifts providers' outside options and is directly reflected in dispute volumes, win rates, and arbitration awards. This policy variation allows us to provide

model-free evidence that changes in providers' outside options causally affect network formation, without imposing the structure of a bargaining model. Our findings show that final-offer arbitration can do more than resolve payment disputes after negotiations fail: by changing providers' outside options, it can reshape provider–insurer contracting itself.

Lastly, the paper is related to the literature on insurance design and the pass-through of cost shocks (Polsky et al., 2016; Cicala et al., 2019; Decarolis, 2015; Kim and Li, 2026). We provide suggestive evidence that issuers with greater ex ante exposure to federal IDR raise gross premiums following the NSA. This result points to a channel through which the costs of a consumer-protection policy may reappear as diffuse premium increases.

2. Institutional Setting and Data

Balance Bills and State-level Regulations. Before the NSA, an out-of-network provider could "balance bill" a privately insured patient for the difference between the provider's billed charge and the insurer's payment. This practice was especially salient in settings where patients have little control over the treating physician: emergency departments and ancillary services such as anesthesiology, radiology, and pathology that are delivered by out-of-network providers at otherwise in-network facilities. Several states regulated surprise billing before 2022, but the scope of these rules varied considerably. Following Ruhter et al. (2023), we classify states as having comprehensive, partial, or no pre-NSA protections. Comprehensive-protection states had binding arbitration or payment standards covering both emergency and ancillary services, while no-prior-protection states lacked comparable rules. Partial-protection states had protections that covered only some surprise-billing situations. Figure A1 maps this classification.

2022 NSA. The NSA, effective 2022, caps patient cost-sharing for emergency and covered ancillary services at in-network levels and requires that cost-sharing count toward in-network deductibles and out-of-pocket limits, regardless of the provider's network status. The law protects patients from the out-of-network price, but does not directly set the payment from insurer to provider. Instead, if the parties fail to agree after a 30-day open-negotiation period, either party may initiate independent dispute resolution (IDR). Take-up of the new IDR process has been far larger than regulators anticipated; Section 4 documents dispute volumes and payment outcomes in detail.

IDR uses final-offer arbitration: the provider and insurer each submit one payment offer, and the arbitrator selects one in its entirety. The Qualifying Payment Amount (QPA), roughly the plan's median contracted rate for the service in the geographic area, serves as the statutory benchmark, though arbitrators may consider additional information.[2] Importantly, because the QPA is

[2]According to CMS, QPA is "the median of the contracted (in-network) rates recognized by the health plan or insurer on January 31, 2019, for the same or a similar item or service furnished by a provider in the same or similar

constructed from in-network contracted rates, it need not measure either pre-NSA out-of-network payments or providers' pre-NSA disagreement payoffs.

Although the NSA was implemented nationwide, its impact varies across states. In no-prior-protection states, the federal law newly replaced balance billing with a regulated insurer-provider payment-dispute process. In comprehensive-protection states, patient protections and payment rules were already in place, making the NSA a smaller incremental change. Our empirical design exploits these pre-existing differences, comparing no-prior-protection states, where the NSA most sharply changed providers' potential payoffs, with comprehensive-protection states, where the pre-existing regime already limited balance billing.

We distinguish the NSA's statutory effective date from the operational launch of federal arbitration. The Act's surprise-billing protections took effect in January 2022, whereas the Departments launched the federal IDR portal in April 2022 (Centers for Medicare & Medicaid Services, 2026; U.S. Department of Labor, 2023). 2022Q3 was the first full calendar quarter of federal IDR operations, and we therefore designate it as the first post-treatment period in our empirical analysis.[3]

Data. We link several data sources covering the ACA marketplace (both individual and small group) from 2017 through 2025. Table A1 reports summary statistics for each data source.

First, CMS IDR public-use files provide the universe of federal arbitration cases closed under the NSA. These records identify disputing parties, service codes, each side's offer relative to the QPA, the final determination, and the selected payment. We use these records to construct arbitration volume, provider win rates, payments relative to the QPA, and net arbitration gains. The details of data construction are in Appendix B. Section 4.1 reports key summary statistics.

Second, the Ideon provider-network data report whether a physician participates in an issuer's network in the ACA marketplace, annually before 2022, and quarterly thereafter. Linked to NPPES by National Provider Identifier (NPI), the data provide physician specialty, practice location, and organizational affiliation, which we use to construct issuer-state-specialty-quarter network participation rates. Provider networks are difficult to measure accurately. Existing studies typically rely on low-frequency, plan-level snapshots of network size or composition (e.g, Dafny et al. (2015), Ericson and Starc (2015), Polsky et al. (2016)). Claims-based measures capture realized patient-provider encounters rather than the set of providers available through insurer contracts. Our national, high-frequency panel of physician network participation instead lets us observe network entry and exit at the issuer-specialty level before and after the NSA and trace it to realized arbitration outcomes, a dynamic contracting margin that point-in-time measures cannot capture. We restrict the sample to individual physicians practicing in the insurer's state.

specialty and geographic region, adjusted annually for inflation."

[3]The case-level public-use files themselves begin one quarter-set later: CMS first published dispute-level records for 2023Q1, so the arbitration statistics in Section 4.1 and Appendix B cover 2023Q1–2025Q2, while the network analysis uses the full 2020–2025 panel with 2022Q3 as the treatment date.

We link physician-issuer-level IDR records to the Ideon data using NPIs and standardized issuer names. The federal IDR records do not identify the line of business associated with a dispute. Thus, the IDR data include disputes arising from the issuer's broader commercial business, including both ACA marketplace and employer-sponsored coverage, and cannot be linked to a specific ACA plan. Our linkage is therefore at the provider and issuer-state level, rather than the plan level. We use this linkage to study whether arbitration exposure in broader commercial markets is reflected in ACA marketplace networks and plan design.

Finally, the ACA marketplace plan public-use files provide premiums, deductibles, and maximum out-of-pocket limits at the plan–rating-area–year level, for both plans on the individual and small group markets. These linkages allow us to study two margins of response: provider network participation at the issuer–state–specialty–quarter level and insurer plan design at the plan–rating-area–year level.

3. The Effect of the NSA on Provider Network Participation

The NSA changed providers' out-of-network payoffs in two opposing ways: it eliminated balance billing of protected patients, which should make out-of-network status less attractive; it also created federal IDR, an enforceable insurer payment process that often yields payments above issuer offers and QPA, which can make out-of-network status more attractive. The net network response therefore reveals which force dominates. If the NSA primarily operated as a balance-billing ban, network participation should rise. In contrast, if provider-favorable arbitration makes IDR a credible alternative to in-network contracting, network participation should fall. We refer to the latter as the *arbitration-option* mechanism. It predicts the largest network responses in NSA-covered specialties, in states where the NSA most sharply changed out-of-network payment regimes, and after providers have had time to observe IDR outcomes.

3.1. Empirical Strategy

We test this prediction using three dimensions of quasi-experimental variation. First, exposure varies across specialties. Emergency medicine, radiology, and anesthesiology are directly affected by the NSA's balance-billing protections and IDR process, whereas many other physician services are not. Second, exposure varies across states. The federal law was a larger institutional shock in states without prior surprise-billing protections than in states with comprehensive pre-existing protections. Third, we observe provider networks before and after the NSA took effect.

Our preferred design combines these three dimensions in a triple-difference framework. We pool all remaining specialties into a single control group, "Other", and compare each of the three

exposed specialties (emergency medicine, radiology, and anesthesiology) against it.[4] This removes any change in the no-prior-protection-versus-comprehensive participation gap common across specialties, leaving the change specific to that specialty in no-prior-protection states after the NSA.

Primary care, general surgery, and psychiatry are held out as explicit placebos, excluded from "Other".[5] Each is estimated against the same pooled control group and should exhibit no differential post-NSA decline if our estimates isolate exposure to the NSA. The placebo set targets two complementary threats. Primary care and general surgery predominantly provide scheduled, patient-selected services and had relatively low pre-NSA out-of-network (OON) use. Their null coefficients test whether no-prior-protection states experienced broad network changes coincident with the NSA. Psychiatry provides a more demanding placebo: their office visits had substantial pre-NSA OON use but fall outside the NSA's balance-billing protections.[6] A null coefficient for psychiatry therefore helps distinguish an NSA-specific response from a broader post-2022 change among specialties with high OON reliance or distinct pre-existing contracting environments.

We stack each focal specialty with the control specialties and estimate a triple-difference:

$$y_{isjt} = \beta_0 \left(\text{Post}_t \times \text{NP}_s \times \text{Focal}_j\right) + \alpha_{isj} + \alpha_{ist} + \alpha_{jt} + \varepsilon_{isjt}, \tag{1}$$

where NP_s indicates a no-prior-protection state, Post_t indicates quarters after the NSA took effect, Focal_j equals one for the focal specialty and zero for the pooled control group.

The outcome y_{isjt} is the in-network share for issuer i, state s, specialty j, and quarter t: the number of specialty-j physicians in (i, s) who are in-network in quarter t, divided by the total number of specialty-j physicians ever in-network in (i, s) over the panel.[7] The denominator is fixed across quarters, so movements in y_{isjt} reflect physicians entering or leaving the network. This outcome measures physician network breadth from the patient's perspective: each NPI represents one physician listed as in-network. Thus, the exit of a 100-physician group removes more physicians

[4] Although pathology is an ancillary specialty affected by the NSA, we exclude it because of our sample restriction requiring each insurer network to include at least 25 physicians in a given specialty. Most pathology networks do not meet this threshold because they are too small. Consequently, there is insufficient variation to identify the event-study estimates, and the triple-difference specification cannot be estimated for pathology. We therefore exclude pathology from the participation sample and report results for the remaining specialties.

[5] We also exclude specialty surgery and hospitalists from "Other", which might belong to the NSA's auxiliary services.

[6] Mark and Parish (2024) reports 15.3% of psychiatrist office visits were out of network pre-NSA, compared with 2.2% for primary care and 3.4% for general surgery, while Cooper et al. (2020a) report a 22% OON rate for emergency care.

[7] We restrict the sample to issuer-state-specialty cells with at least 25 in-network specialty-j physicians throughout the sample period. The threshold of 25 is the 10th percentile of the number of emergency-medicine physicians ever in-network in an issuer's network, and it drops thin cells whose in-network shares are mechanically volatile. Our results are quantitatively similar under alternative sample filters, such as screening on the total number of physicians in the network rather than the own-specialty count, or retaining all cells and adding an additive indicator for cells with fewer than 25 own-specialty (respectively, total-network) in-network physicians, and under alternative state-pooled denominators, i.e., the total number of specialty-j physicians ever in-network in state s over the panel.

from patients' listed choice set, and appropriately counts more than the exit of a solo practitioner. This measure does not assume that physicians negotiate independently; contracting may occur at the provider-group level.

The specification includes issuer-state-specialty fixed effects α_{isj}, issuer-state-quarter fixed effects α_{ist}, and specialty-quarter fixed effects α_{jt}. Issuer-state-specialty fixed effects absorb all persistent differences in the participation of a given specialty in a local issuer network. Issuer-state-quarter fixed effects absorb any contemporaneous shock to an issuer's network in a state that is common across specialties, including local market conditions, plan changes, and issuer-specific directory updates. Specialty-quarter fixed effects absorb national changes specific to each specialty, such as workforce trends, pandemic recovery, and common reporting changes.

Our parameter of interest, β_0, is the post-minus-pre change in the focal specialty's no-prior-protection-versus-comprehensive participation gap, net of the same change for the control group. Network exit through the arbitration option implies $\beta_0 < 0$: participation should fall most for an NSA-exposed specialty, in the states where the NSA bound hardest, relative to a specialty with no access to IDR. The identifying assumption is that, absent the NSA, the focal-minus-*Other* participation gap would have evolved similarly in no-prior-protection and comprehensive-protection states.

We estimate equation (1) and every specification in the paper with the Borusyak et al. (2024) imputation estimator.

3.2. Primary Estimates and Placebos

Figure 1 presents the event-study analog of equation (1) by specialty. We highlight three findings.

First, the exposed specialties show flat pre-trends and then turn negative after the NSA. Table A2 reports corresponding average treatment effects, which are economically large: pooling all post-NSA quarters, NSA reduces in-network provider share by -4.7 percentage points for emergency medicine, -3.4 percentage points for anesthesiology, and -1.8 percentage points for radiology, equal to 13.5, 9.7, and 3.5 percent of their pre-NSA means.[8] The negative estimates indicate that the balance-billing channel did not dominate the network response. Their concentration in covered specialties and states without prior protections is consistent with an arbitration-outside-option interpretation, which Section 4 evaluates using complementary evidence.

Second, the placebo specialties do not exhibit the same pattern. General surgery, primary care, and psychiatry, which are not directly exposed to the NSA's IDR process, have coefficients that remain close to zero throughout the post period. This contrast is central to the identification. It

[8]Two features of radiology are consistent with a smaller and slower response: a substantial share of radiology is scheduled outpatient imaging for which patients select in-network providers, so a smaller fraction of radiologists' revenue arises in NSA-covered situations, diluting NSA exposure. Remote interpretation and national practice organizations make radiology's participation decisions less tied to the practice-state regime that assigns treatment.

indicates that the estimates are not picking up a general change in networks in no-prior-protection states, but rather a response concentrated among specialties with access to IDR. The psychiatry null is informative: despite pre-NSA OON reliance comparable to the exposed specialties, its participation shows no post-NSA decline, consistent with the response reflecting the NSA's coverage of a specialty's out-of-network care rather than high baseline out-of-network exposure itself.

Third, the timing is consistent with gradual adjustment in provider–insurer contracting. In the first 0–3 post-NSA quarters, the exposed-specialty coefficients are close to zero. The declines appear later: four to seven quarters after implementation, emergency medicine and anesthesiology fall by 6.3 and 5.5 percentage points, and the effect persists afterward. The delayed response is also compatible with providers learning about early IDR outcomes before adjusting network contracts. Timing alone, however, does not distinguish this explanation from other sources of delayed contracting adjustment. Section 4 examines additional evidence on the arbitration mechanism.

Figure A2 reports robust patterns when we instead conduct the simpler difference-in-differences for each specialty in isolation.

3.3. Ruling Out Alternative Explanations

We examine four alternative explanations for the estimated network decline: transient directory noise, physicians' workforce attrition, NSA's own provider-directory provisions, and changes concentrated among named IDR participants or large staffing organizations. We address each in turn.

First, our estimates are not driven by provider-directory churn. Table A3 reports two checks: gap-bridged participation measures that treat absences shorter than two or four consecutive quarters as continued in-network status when the physician later returns, and samples that exclude issuer–state-quarter cells with large directory changes (changes above 25 percent of the total NPI count or in the upper tail of the change distribution). The estimates are stable throughout: the post-NSA decline reflects persistent exit rather than transient reporting gaps or bulk mechanical updates.

Second, the estimates are not driven by physician workforce attrition. Physician disappearance could reflect retirement, relocation, or post-pandemic contraction in the workforce, rather than exit from a particular insurer's network. We therefore recompute the outcome under two stricter exit definitions, counting as exits only disappearances by physicians who remain active in NPPES and, more stringently, only those who remain listed in at least one other insurer's directory in the same state. Table A3 shows the results remain quantitatively similar. The latter restriction focuses on physicians who remain observable in the same local market, making general workforce attrition an unlikely explanation for the estimated decline.

Third, the estimates are unlikely to be driven by insurers removing stale listings under the NSA's directory-verification requirements. Several facts weigh against this hypothesis. First, the requirements applied nationwide and across specialties, whereas the estimated decline is concen-

trated in NSA-exposed specialties in no-prior-protection states and is absent among the placebo specialties. Second, the institutional setting did not generate a synchronized cleanup: pending rulemaking, the Departments directed plans to apply a "good faith, reasonable interpretation" of the provisions, and the enforcement backstop protected patients' cost sharing rather than mandating delisting (U.S. Departments of Labor, Health and Human Services, 2021). Finally, subsequent IDR activity indicates that at least some directory exits involve active physicians rather than stale listings. Among NPIs observed in the IDR records, physicians exiting an issuer's Ideon network average 6.94 disputes. This evidence links directory disappearance to subsequent out-of-network activity for a subset of physicians, although it does not validate every exit.

Finally, the estimates are not driven by providers observed in arbitration or by large physician staffing organizations. Table A4 addresses both possibilities. First, NPIs that ever appear as disputing providers account for only 2.6% of exit events in exposed specialties and 0.5% of all exit events in the ACA marketplace. Re-estimating equation (1) after dropping all such IDR-participating NPIs leaves the estimates close to baseline.[9] Second, because IDR filing is concentrated among a small number of national physician-staffing organizations (Fiedler and Adler, 2024), two of which failed in 2023[10], the estimated decline could reflect the behavior of a handful of firms rather than a market-wide response. We therefore exclude NPIs affiliated with these organizations[11] and re-estimate equation (1). The estimates are quantitatively similar. These exercises show that the decline in patient-facing network breadth extends beyond the NPIs in IDR records and beyond the largest staffing firms. They do not imply that every exiting NPI represents a separate contracting decision. Physicians whose NPIs never appear in IDR may belong to a group that negotiates or files disputes collectively. We therefore interpret this evidence as showing that the observed loss of physician availability is broader than the set of named arbitration participants.

Taken together, our findings point to an economically meaningful network-participation response to the NSA: the declines are concentrated in covered specialties and in states where the federal reform represented the largest institutional change, emerge with a contracting lag, and are absent in placebo specialties.

[9]The complementary sample restricted to disputing NPIs is too thin to estimate: almost no issuer-state-specialty cell retains 25 disputing physicians.

[10]Envision Healthcare filed for Chapter 11 in May 2023, and American Physician Partners wound down its operations in July 2023; both cited the NSA's implementation among the causes.

[11]Fiedler and Adler (2024) identifies TeamHealth, SCP Health, Envision Healthcare, and Radiology Partners as top IDR filers; we additionally include American Physician Partners. Affiliation is assigned using the provider email domain reported in the IDR records, supplemented by normalized provider-group names.

4. Evidence Consistent with the Arbitration-Option Mechanism

Section 3 shows that the NSA reduced provider network participation where the federal law represented the largest change. The triple-difference design identifies this net policy effect but does not isolate the role of arbitration. We now assess whether the arbitration-option mechanism is consistent with the observed network response. We first ask whether federal IDR became sufficiently prevalent and financially meaningful to affect provider–insurer bargaining. We then examine whether network participation among selected IDR-participating provider groups changes around their first favorable determinations. These analyses provide complementary mechanism evidence rather than separate causal estimates of arbitration's effect.

4.1. IDR is a Widely-Used, Provider-Favorable Payment Channel

For arbitration to affect network contracting, it must be used frequently enough and generate sufficiently meaningful payments to influence providers' disagreement payoffs. Figure 2 documents the scale and payment outcomes of federal IDR.

Panel (a) shows that quarterly IDR volume rises steadily and then accelerates from late 2023, reaching more than 1.8 million disputes cumulatively by 2025Q2. This scale implies that IDR has become a routine payment channel for out-of-network services rather than a rarely used fallback. Panel (b) shows that arbitration outcomes are strongly tilted toward providers: among non-default determinations, providers win about 85% of cases, and the tilt is stable across quarters.

Panels (c)–(e) assess whether these provider-favorable outcomes are economically meaningful. Accessing IDR is costly: both parties pay an administrative fee, and the losing party pays the arbitration fee.[12] We calculate the net payments the provider received from the dispute as the arbitrated amount, net of expected administrative fees and, if the provider loses, the arbitration fee. We define the providers gain from IDR as the net payment received, less the issuer's offer, so a positive value means the provider receives more from pursuing IDR than from accepting the issuer's pre-arbitration offer.

Panel (c) plots the quarterly distribution of this net gain per dispute. Provider gains are small or negative early in implementation, when filing costs were high, and the IDR system was still developing. After the administrative fee reduction in late 2023, however, the distribution shifts upward: the median gain becomes positive, and the upper tail grows substantially. By 2025Q2, the 75th percentile reaches $2,574, and the 1-percentile-winsorized mean reaches $3,373, implying that a substantial set of disputes generates economically meaningful payments to providers, even after accounting for the direct costs of accessing arbitration.

[12]The administrative fees were $50 per party in 2022, increased to $350 for disputes initiated in early 2023, reverted to $50 in August 2023, and were set at $115 in January 2024. The arbitration fee paid by the losing party is also nontrivial: $200–$884 for a single determination and $268–$1,173 for a batched determination.

Panel (d) compares issuer offers and prevailing offers as a share of the Qualifying Payment Amount (QPA). The QPA provides an observable benchmark for in-network payments because it is the plan's median contracted in-network rate for the same or similar service, specialty, and geographic area. It is not, however, a measure of providers' pre-NSA out-of-network outside option. CMS also describes QPA as a factor that certified IDR entities consider in payment determinations. Issuer offers cluster near QPA, while prevailing offers are shifted sharply to the right. The median prevailing offer is 374 percent of QPA, compared with 100 percent for issuer offers. Thus, although we do not observe each provider's own contracted rates, arbitrated payments often exceed the statutory median contracted-rate benchmark that anchors the NSA's payment framework.

Panel (e) summarizes the same incentives at three levels. At the IDR-case level, prevailing offers average 1,048 percent of QPA, and 75.9 percent of disputes generate positive provider net gains after fees. The mean and median net provider gains are $2,645 and $363 per dispute. These case-level gains accumulate into meaningful annual amounts for participating providers: at the provider-year level, net gains average $117,472, with the 75th and 90th percentiles reaching $22,315 and $231,970. Finally, the state-quarter rows show that IDR participation is not limited to a few isolated cases: the average state-quarter has 128 participating providers and 3,542 disputes, with substantially larger volumes in the upper tail.

These facts provide several benchmarks for interpreting IDR as a potential provider-favorable outside option. First, relative to accepting the issuer's offer, IDR generates positive net gains for a substantial share of disputes. Second, relative to QPA, the best observable benchmark for median in-network reimbursement, prevailing offers are often substantially higher. Third, relative to the pre-NSA balance-billing regime, IDR changes the nature of the out-of-network payments: before the NSA, an out-of-network provider's residual claim was against the patient; after the NSA, the provider can pursue payment from the insurer through IDR.

Finally, we compare IDR payments with those under the pre-NSA out-of-network regime. Using 2019 commercial claims for professional emergency services, Duffy et al. (2022b) estimate that mean out-of-network allowed amounts were 2.13 times the estimated QPA. By comparison, the median prevailing offer in our full IDR sample is 3.74 times QPA, and 92.5% of determinations exceed the pre-NSA benchmark. Thus, many selected IDR awards exceed even a benchmark that recognizes that pre-NSA out-of-network allowed amounts were already substantially above QPA.

This comparison is necessarily descriptive. We do not claim that IDR dominates every provider's previous outside option, as entry into IDR is selected, and we do not observe provider-specific pre-NSA contracted rates, balance-bill collections, or collection costs. Rather, statistics show that IDR is high-volume, provider-favorable, and financially attractive relative to accepting observed issuer offers. Awards frequently exceed the statutory in-network benchmark and, for many selected disputes, the external pre-NSA out-of-network benchmark. These patterns make arbitration

a plausible countervailing force to the balance-billing prohibition.

4.2. Network Exit Around Favorable Arbitration Outcomes

We now examine whether, among provider groups that participate in IDR, network participation changes around the first favorable determination.

We conduct an event study at the provider-group×issuer level. Let g index a provider group, i an issuer, s a state, and t a calendar quarter. A provider group may consist of either multiple providers or a solo practitioner. Provider group affiliation is observed only through the Provider/Facility Group Name field in the IDR data and is therefore available only for physicians who appear in at least one IDR dispute. As a result, our sample includes only physicians with at least one observed IDR dispute, rather than all physicians in provider groups. Define the event quarter e_g as the quarter of the group's first favorable IDR determination in the state, i.e., the arbitrator selects the provider's offer. The event is defined at the group level and is common to all of the group's issuer relationships, rather than specific to the issuer it disputes. This allows us to examine whether a favorable determination is followed by broader changes in the group's network position across all insurers, consistent with the group updating the value of IDR as an outside option. We estimate

$$y_{gist} = \sum_{k \neq -1} \beta_k \, \mathbb{1}\{t - e_g = k\} + \alpha_{gis} + \alpha_t + \varepsilon_{gist}, \tag{2}$$

where y_{gist} is either the in-network share of the group g's roster in issuer i's network or an indicator for complete exit (all physicians in the provider group g exit issuer i's network); α_{gis} and α_t are provider group–issuer and time fixed effects. Standard errors are clustered at provider group-issuer level. We estimate equation (2) using the Borusyak et al. (2024) imputation estimator.

The coefficients β_k trace the timing of network participation relative to the group's first favorable IDR outcome. We use the not-yet-treated groups as the comparison group for those who have already won in the same calendar quarter. The identifying assumption is that, absent a favorable determination, a winning group's network status would have continued in parallel with that of not-yet-treated groups, and a group does not adjust its network position before the determination is realized.

A limitation is that the timing of a group's first win may not be random. A group wins only after it elects to arbitrate and an arbitrator rules in its favor. Thus, groups that are already pursuing an out-of-network strategy may be more likely both to generate IDR disputes and to experience a favorable determination. We therefore interpret our estimates as evidence of a descriptive mechanism rather than as a standalone causal estimate that favorable determination itself causes exit.

We use two checks to discipline the mechanism interpretation. First, pre-event coefficients test directly for differential trends and pre-emptive withdrawal before the favorable determination.

Second, we run the same event study for non-NSA-affected provider groups (general surgery, primary care, psychiatry, and other specialists) located in the same ZIP code and assigned the same event quarter as the affected group. These placebo specialties face the same local market conditions and timing, but they do not obtain an IDR outside option from the treated group's favorable determination. A local shock, directory update, or broad issuer network change coincident with the event would therefore move the placebo specialties as well; the arbitration-option mechanism predicts a response only for the affected group.[13]

Figure 3 reports the estimates of the event study. We highlight three findings.

First, network participation among NSA-exposed provider groups declines after the first favorable arbitration outcome. The pre-event path is flat, while the in-network share falls after the event. The pooled estimates in Table A6 report a 6.8 percentage-point decline in the in-network share, equal to 19.8 percent of the pre-event mean, and complete exit rises by 7.7 percentage points. We interpret this timing as consistent with providers learning about, or acting on, the value of IDR as an outside option, so that network status changes after, not before, the first favorable determination. This complements the roughly one-year lag in Figure 1, the time required for initial IDR outcomes to become observable.

Second, the placebo groups' network participation does not move at the treated groups' event dates. Their coefficients are centered around zero throughout the event window, so the decline of NSA-exposed specialties is unlikely to be driven by local shocks common to all groups coinciding with the favorable arbitration outcome. Complete exit has no meaningful placebo counterpart because the data never show all placebo physicians in a ZIP code exiting an insurer.

Third, the network response scales with the size of the favorable event. The arbitration-option mechanism predicts not only that exit follows a favorable determination, but that exit should be stronger when the determination reveals a more valuable option. We measure the size of each group's first favorable event in four ways: the gain of the provider's offer relative to accepting the insurer's offer, and the provider's prevailing offer itself, each expressed as a percent of the QPA and in dollars. Table A7 splits groups at the median of each measure and estimates the responses separately. The post-event decline is uniformly larger for above-median events, and complete exit rises more. This dose-response pattern is consistent with groups acting on the magnitude of the revealed arbitration option rather than its mere occurrence.

Taken together, we find that among selected IDR-participating groups, network participation

[13] Another natural alternative is to examine the event at a group's first unfavorable determination: a group that loses in arbitration gains no new outside option and so should not withdraw on that account. Two features of the data make this loss-based placebo uninformative. First, determinations are overwhelmingly provider-favorable, so first-loss events are comparatively rare. Second, a group's first loss and first win often occur in the same or adjacent quarters (72% of times), and over 80% of groups in the loss sample also appear in the win sample. A first-loss event study would therefore trace nearly the same groups at nearly the same event time as the win-based estimates, making it mechanically similar rather than an independent test. The non-affected local placebo avoids both problems.

declines after the first favorable determination, while nearby non-covered specialties do not show comparable changes. The same-ZIP placebo helps rule out coincident local shocks but does not address selection into arbitration or the timing of favorable determinations. Nevertheless, because groups might select into arbitration, we interpret these results as complementary mechanism evidence, not as proof that receiving a favorable determination causes network exit.

5. Suggestive Evidence on Insurer Pricing

Sections 3 and 4 show that the NSA reduced provider network participation and document complementary patterns consistent with arbitration strengthening providers' outside options. These responses can raise insurers' costs in two ways. First, when a provider prevails in IDR, the insurer must pay the arbitrated amount, which often exceeds both the insurer's offer and the QPA benchmark. Second, favorable arbitration outcomes strengthen providers' disagreement payoffs in subsequent network negotiations, potentially raising expected reimbursement costs even for services that settle before arbitration. We therefore examine whether premium growth is greater among issuers with higher ex ante predicted arbitration exposure. Because this exposure is not randomly assigned and the pricing analysis cannot exploit the cross-specialty variation used previously, we interpret the results as suggestive rather than causal.

Because the unit of observation is the issuer–state–plan rather than the specialty, we cannot use the triple-difference design from Section 3. We instead compare post-NSA changes across issuer–states with higher and lower ex ante predicted arbitration exposure, while absorbing state-year shocks. For an issuer-state pair, we define exposure by the issuer's pre-period contracting with the four physician staffing firms that dominate federal IDR filings—TeamHealth, SCP Health, Envision Healthcare, and Radiology Partners (Fiedler and Adler, 2024)—which together account for 46.5 percent of federal dispute line items over 2023Q1-2025Q2 (Figure A3). Specifically, we measure the staffing-firm affiliation share: the share of the issuer's pre-period in-network physicians in NSA-exposed specialties who are affiliated with one of the four firms from the IDR records. A low affiliation share means the staffing firms' physicians are largely out of the issuer's network, so their NSA-covered encounters can be routed to IDR, where these firms file at scale and typically prevail above the QPA, creating bigger financial losses to these issuers. We therefore construct the low-affiliation-share issuers as the treatment group and classify an issuer-state as high-exposure when its affiliation share is at or below the within-state median across sample issuers. All plans offered by that issuer in that state inherit the same exposure classification.

Let k index a plan cell, i.e., an issuer by state by metal tier by rating area by market segment (individual or small group). let $s(k)$ denote its state, and let t index a plan year. We collapse plans

to the cell-year level and estimate

$$y_{kt} = \beta \left(\text{HighExp}_k \times \text{Post}_t \right) + \alpha_k + \alpha_{s(k)t} + \varepsilon_{kt}, \tag{3}$$

where HighExp_k equals one for cells belonging to below-median issuer-states, Post_t indicates plan years from 2022 onward, and α_k and $\alpha_{s(k)t}$ are plan-cell and state-year fixed effects. The plan-cell fixed effects absorb the issuer-state, rating area, metal tier, and market segment dimensions and all of their interactions; standard errors are clustered by issuer-state.[14] We take monthly premiums and annual in-network deductible as outcomes y_{kt}.[15] We estimate equation (3) with the Borusyak et al. (2024) imputation estimator. Because ACA premiums for plan year t are filed before the plan year begins, we distinguish the legal implementation of the NSA from the years in which insurers could plausibly incorporate realized IDR experience into pricing: our preferred post-learning indicator equals one for 2024 and later; we report 2022 and 2023 separately as transition years.

Our parameter of interest, β, is the differential post-NSA change in plan design for high-exposure relative to low-exposure issuers in the same state and year. If greater predicted arbitration exposure raises insurers' expected costs, premium growth should be greater among high-exposure issuers, i.e., $\beta > 0$.

We have two findings. First, Figure 4 shows that pre-period coefficients on premiums are flat, whereas post-period coefficients gradually rise after the policy takes effect. Post-period estimates are largest in the later plan years, when insurers could plausibly have observed IDR volumes and arbitration outcomes. Because 2022-2023 premiums were filed before insurers had substantial realized IDR experience, we interpret those early coefficients as possible expectation or transition effects after the announcement of NSA, and place more weight on the later post-learning years.

Second, gross premium growth is greater among high-exposure issuers by $15.2 per month for a 27-year-old and $26.8 per month for a 50-year-old, equal to 3.6 and 3.8 percent of baseline premiums in 2024-2025 (compared to the low-exposure issuers), as reported in Table A8. Deductible responses are positive on average but noisier. This pattern is consistent with a broad pricing response to arbitration exposure. Because the NSA insulates patients from out-of-network prices for covered services, insurers cannot recover higher reimbursement costs by raising out-of-network cost-sharing for those services. Instead, they adjust plan parameters that apply more broadly to enrollees.

There are a few caveats in interpreting these pricing estimates. First, these are differential

[14] We do not use the no-prior-protection-versus-comprehensive state contrast as the main pricing specification, because premium setting is strongly state-specific. Instead, state-year fixed effects absorb all shocks common to a state-year, including changes in regulation, risk pools, local premium trends, and market-wide cost shocks.

[15] We exclude out-of-network parameters, because the NSA caps patient liability for covered services at in-network levels and makes out-of-network design largely irrelevant to affected enrollees. We exclude in-network maximum out-of-pocket, because it is bound by the ACA's statutory limit on cost-sharing and has little room to move.

effects for high-exposure issuers relative to lower-exposure issuers in the same state and year, not the average effect of the NSA on all ACA marketplace plans. Second, we examine gross premiums that measure insurer pricing, but not the enrollees' out-of-pocket premiums in the subsidized ACA marketplace. These premium increases are borne partly by federal subsidy spending and partly by unsubsidized enrollees. These estimates therefore provide suggestive evidence of a differential pricing response among more-exposed issuers. They are not estimates of the NSA's average effect on premiums or of the causal effect of realized arbitration losses.

6. Discussion and Conclusion

Our primary finding is that the NSA reduced provider network participation in covered specialties and in states where the federal reform represented the largest institutional change. This result runs counter to the standard prediction that eliminating balance billing should make network participation more attractive. Complementary evidence points to federal arbitration as a plausible explanation: IDR became a high-volume, provider-favorable payment channel; the network contraction extends beyond NPIs named in arbitration records; and network participation among selected IDR-participating groups declines after their first favorable determinations. Because arbitration participation and event timing are selected, these patterns do not establish that favorable arbitration outcomes cause network exit. Finally, we find suggestive evidence that premium growth is greater among insurers with higher ex ante arbitration exposure.

Out-of-network emergency bills averaged roughly $623 above in-network amounts and affected about 22 percent of emergency visits at in-network hospitals (Cooper et al., 2020a). Combined with an emergency-department utilization rate of 41.9 visits per 100 people per year, this implies direct savings of approximately $57 per enrollee per year ($= 0.419 \times 0.22 \times 623$). At the same time, our results suggest that insurers pass these additional costs on to consumers through higher premiums. These costs include both direct losses incurred in IDR and higher negotiated in-network rates resulting from providers' improved outside options. For high-exposure issuers, the implied increase in annual gross premiums is $183 for a 27-year-old and $322 for a 50-year-old, which is substantially larger than the $57 annual savings benchmark.

This comparison is informative about scale, but is not a net welfare estimate. The premium estimates are differential effects for high-exposure issuers, and gross premiums do not map one-for-one into enrollee payments in the subsidized ACA marketplace. Meanwhile, we do not monetize direct access, choice, or continuity-of-care costs from network narrowing. Thus, this calculation reveals that the indirect costs channel caused by NSA is economically large enough to matter for policy evaluation. A complete welfare analysis requires modeling the feedback from arbitration outcomes to subsequent in-network rate negotiations and the resulting equilibrium in network

breadth and premiums, which we examine in our companion structural work (Barwick et al., 2026).

These findings speak to ongoing debates about the NSA's implementation. Several design features of IDR could mitigate the outside-option incentive: higher filing fees, a stronger role for the QPA benchmark in arbitrator decisions, or measures targeting providers who file large dispute volumes while remaining out-of-network. More broadly, our analysis illustrates a general principle: a dispute-resolution mechanism intended as a backstop can become a primary channel of market activity when it offers participants more favorable terms than bilateral negotiation. Anticipating these responses is essential to designing policies that achieve their goals without costly unintended consequences.

References

Adler, Loren, Matthew Fiedler, Paul B. Ginsburg, Mark A. Hall, Erin Trish, Christen Linke Young, and Erin L. Duffy, "State Approaches to Mitigating Surprise Out-of-Network Billing," Technical Report, USC–Brookings Schaeffer Initiative for Health Policy February 2019.

Ashenfelter, Orley and David E. Bloom, "Models of Arbitrator Behavior: Theory and Evidence," *American Economic Review*, 1984, *74* (1), 111–124.

Barwick, Panle Jia, Anran Li, Tianli Xia, and Shizhe Yu, "The Price of Arbitration: How Federal Surprise Billing Regulation Reshapes Provider Networks and Healthcare Costs," 2026. Work in progress.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess, "Revisiting Event-Study Designs: Robust and Efficient Estimation," *Review of Economic Studies*, 2024, *91* (6), 3253–3285.

Centers for Medicare & Medicaid Services, "Independent Dispute Resolution Reports," https://www.cms.gov/initiatives/no-surprise-billing/overview/policies-resources/independent-dispute-resolution-reports August 2026. Last modified August 21, 2026; accessed August 25, 2026.

Chartock, Benjamin and Christopher Whaley, "Arbitration and Negotiated Prices: Evidence from Insurer–Doctor Disputes," 2025. Working paper, October 31, 2025.

_ , Loren Adler, Bich Ly, Erin Duffy, and Erin Trish, "How Have State Surprise Billing Protections Affected Out-of-Network Billing?," *Health Affairs*, 2021, *40* (7), 1137–1144.

Cicala, Steve, Ethan MJ Lieber, and Victoria Marone, "Regulating markups in US health insurance," *American Economic Journal: Applied Economics*, 2019, *11* (4), 71–104.

Cooper, Zack and Fiona Scott Morton, "Out-of-Network Emergency-Physician Bills — An Unwelcome Surprise," *New England Journal of Medicine*, 2016, *375* (20), 1915–1918.

_ , Fiona Scott Morton, and Nathan Shekita, "Surprise! Out-of-Network Billing for Emergency Care in the United States," *Journal of Political Economy*, 2020, *128* (9), 3626–3677.

_ , Hao Nguyen, Nathan Shekita, and Fiona Scott Morton, "Out-of-Network Billing and Negotiated Payments for Hospital-Based Physicians," *Health Affairs*, 2020, *39* (1), 24–32.

Dafny, Leemore, Igal Hendel, and Nathan Wilson, "Narrow Networks on the Health Insurance Exchanges: What Do They Look Like and How Do They Affect Pricing? A Case Study of Texas," *American Economic Review P&P*, 2015, *105* (5), 110–114.

Decarolis, Francesco, "Medicare part d: Are insurers gaming the low income subsidy design?," *American Economic Review*, 2015, *105* (4), 1547–1580.

Duffy, Erin L., Christopher Garmon, Loren Adler, Adam Biener, and Erin Trish, "No Surprises Act Independent Dispute Resolution Outcomes for Emergency Services," *Health Affairs Scholar*, 2024, *2* (11), qxae132.

_ , **Loren Adler, Benjamin L. Chartock, and Erin Trish**, "Dispute Resolution Outcomes for Surprise Bills in Texas," *JAMA*, 2022, *327* (23), 2350–2351.

Duffy, Erin Lindsey, Adam Biener, Christopher Garmon, and Erin E Trish, "Comparison of estimated no surprises act qualifying payment amounts and payments to in-network and out-of-network emergency medicine professionals," in "JAMA Health Forum," Vol. 3 American Medical Association 2022, pp. e223085–e223085.

Ericson, Keith Marzilli and Amanda Starc, "Measuring Consumer Valuation of Limited Provider Networks," *American Economic Review P&P*, 2015, *105* (5), 115–119.

Farber, Henry S., "An Analysis of Final-Offer Arbitration," *Journal of Conflict Resolution*, 1980, *24* (4), 683–705.

Fiedler, Matthew and Loren Adler, "A First Look at Outcomes under the No Surprises Act Arbitration Process," Technical Report, Brookings Institution, Center on Health Policy 2024.

Garmon, Christopher and Benjamin Chartock, "One in Five Inpatient Emergency Department Cases May Lead to Surprise Bills," *Health Affairs*, 2017, *36* (1), 177–181.

_ , **Yiting Li, Sheldon M. Retchin, and Wendy Yi Xu**, "The Impact of Surprise Billing Laws on Emergency Services," *Health Economics*, 2024, *33* (11), 2450–2462.

_ , _ , _ , **and** _ , "The Impact of Surprise Billing Laws on Hospital-Based Physician Prices and Network Participation," *Health Economics, Policy and Law*, 2024, pp. 1–39. Published online December 11, 2024.

Gowrisankaran, Gautam, Aviv Nevo, and Robert Town, "Mergers When Prices Are Negotiated: Evidence from the Hospital Industry," *American Economic Review*, 2015, *105* (1), 172–203.

Ho, Katherine and Robin S. Lee, "Insurer Competition in Health Care Markets," *Econometrica*, 2017, *85* (2), 379–417.

Kim, Paul H S and Anran Li, "Insurer Risk and Public Risk-Sharing: Quantifying the Value of Reinsurance," *Review of Economic Studies*, June 2026, p. rdag063.

Mark, Tami L. and William J. Parish, "Behavioral Health Parity—Pervasive Disparities in Access to In-Network Care Continue," Technical Report, RTI International April 2024.

Polsky, Daniel, Zuleyha Cidav, and Ashley Swanson, "Marketplace Plans With Narrow Physician Networks Feature Lower Monthly Premiums Than Plans With Larger Networks," *Health Affairs*, 2016, *35* (10), 1842–1848.

Prager, Elena and Nicholas Tilipman, "Regulating Out-of-Network Hospital Payments: Disagreement Payoffs, Negotiated Prices, and Access," Working Paper 33727, National Bureau of Economic Research 2025.

Ruhter, Joel, Steven Sheingold et al., "Evaluation of the Impact of the No Surprises Act on Health Care Market Outcomes: Baseline Trends and Framework for Analysis," Technical Report, Office of the Assistant Secretary for Planning and Evaluation (ASPE), U.S. Department of Health and Human Services July 2023. No Surprises Act Reports to Congress Series.

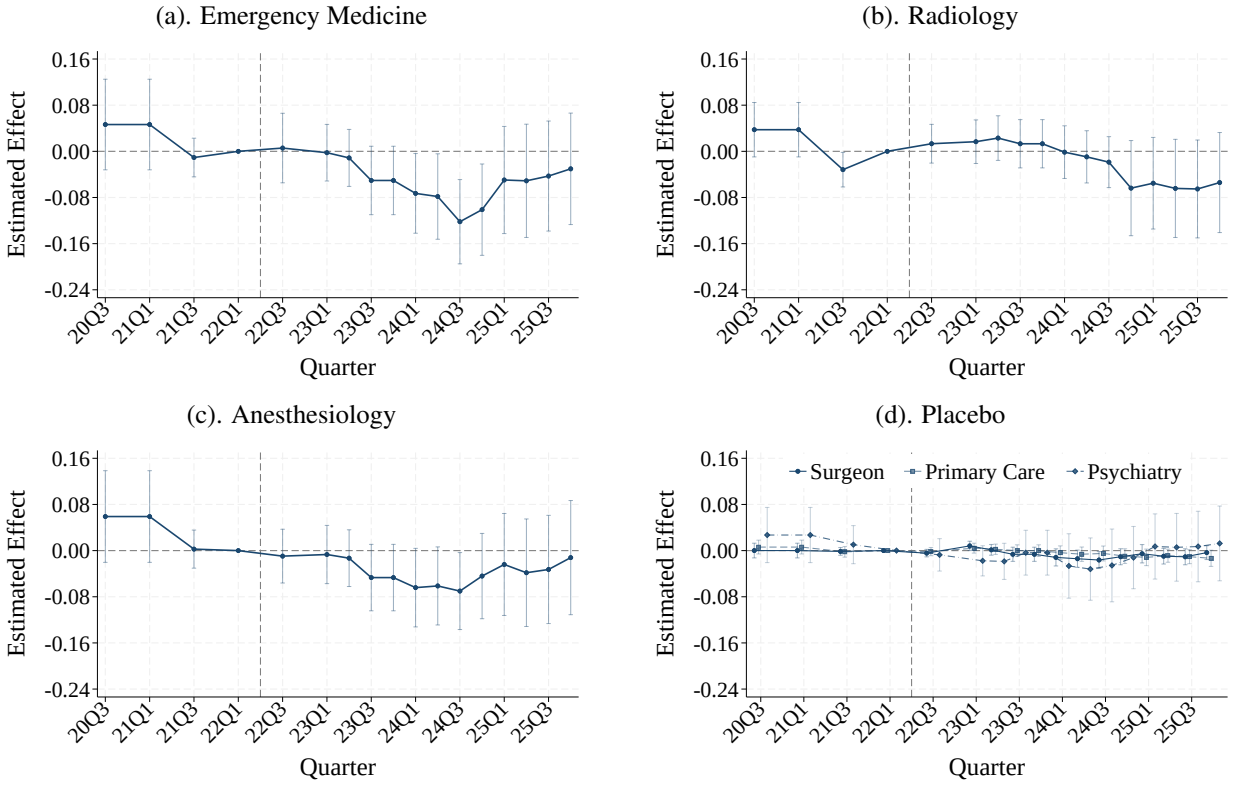
Serna, Natalia, "Determinants of Healthcare Provider Networks: Risk Selection vs. Administrative Costs," Technical Report 2024.

Shepard, Mark, "Hospital network competition and adverse selection: evidence from the Massachusetts health insurance exchange," *American Economic Review*, 2022, *112* (2), 578–615.

U.S. Department of Labor, "Federal Independent Dispute Resolution Process—Status Update," https://www.dol.gov/agencies/ebsa/laws-and-regulations/laws/no-surprises-act/federal-idr-process-status-update April 2023. Published April 27, 2023; accessed August 25, 2026.

U.S. Departments of Labor, Health and Human Services, "FAQs About Affordable Care Act and Consolidated Appropriations Act, 2021 Implementation Part 49," August 2021. Question 8.

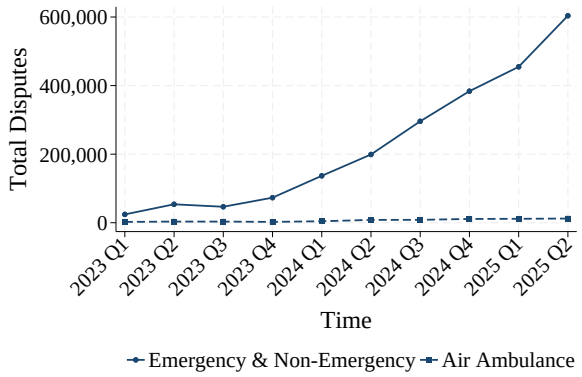
Figure 1. Network Participation Falls for NSA-Exposed Specialties



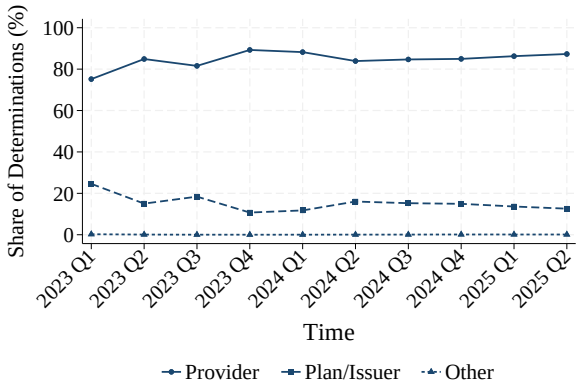
Notes: Each panel plots the event-study analog of the triple difference in equation (1), replacing $Post_t$ with event-time indicators. The focal specialty is indicated in each panel and is compared with the pooled "Other" specialty group; general surgery, primary care, and psychiatry are placebo focal specialties. The sample is the insurer-state-specialty panel in 2020-2025, restricted to issuer-state-specialty cells with at least 25 in-network own-specialty physicians in every quarter and to physicians practicing in the plan's state. The outcome variable is the share of a specialty's ever-in-network physicians that remain in the insurer's network in the specific quarter. Fixed effects are issuer-state-specialty, issuer-state-quarter, and specialty-quarter, as in equation (1). Cells are weighted by the number of ever-in-network physicians. Estimates use the imputation estimator of Borusyak et al. (2024). Standard errors are clustered by issuer–state; bars are 95% confidence intervals.

Figure 2. Federal IDR Use, Payment Outcomes, and Estimated Arbitration Gains

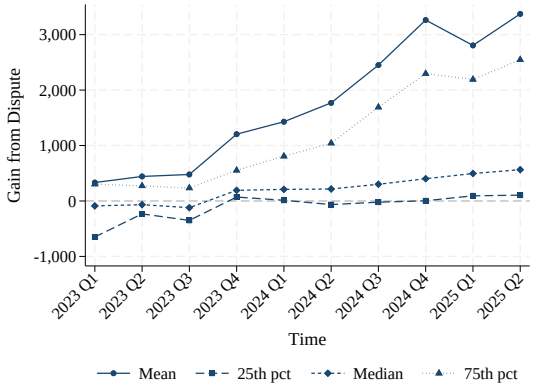
(a). Quarterly dispute volume



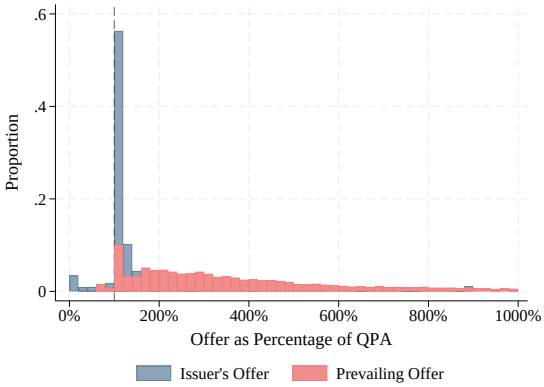
(b). Provider win rate over time



(c). Distribution of Provider Net Gains over time



(d). Issuer and prevailing offers, as a proportion of QPA

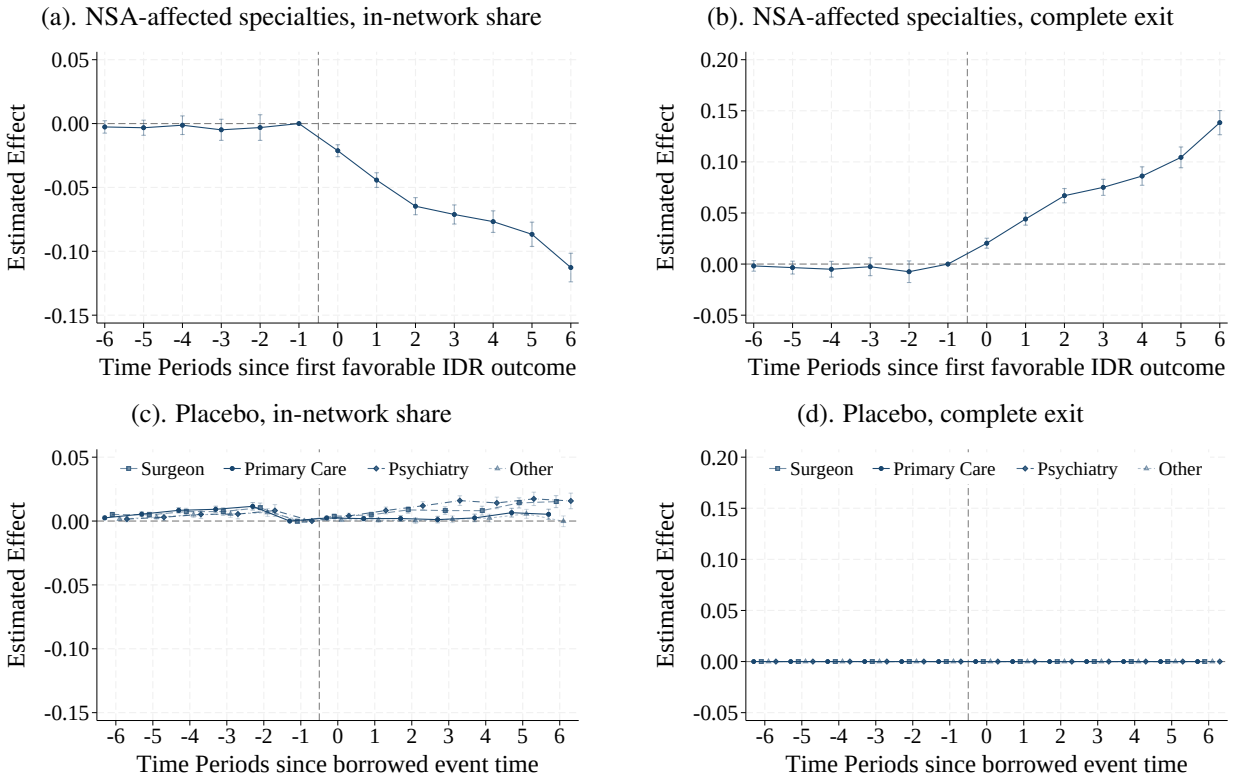


(e). Incremental Payments Relative to Issuer Offers

	Mean	p25	p50	p75	p90
(a). IDR-case level					
Provider win probability (%)	85.9				
Prevailing offer as percent of QPA	1,048.2	209.0	374.0	769.0	1,950.0
Mean/median provider offer as percent of QPA	1,221.7	261.0	436.0	888.0	2,387.0
Mean/median insurer offer as percent of QPA	151.7	100.0	100.0	142.3	269.0
Prevailing offer (dollars)	3,602	518	1,000	2,904	8,000
Provider offer (dollars)	4,319	611	1,196	3,750	9,000
Insurer offer (dollars)	646	145	305	658	1,428
QPA (in thousand dollars)	0.51	0.11	0.21	0.44	1.14
Net provider gain after fees	2,645	17	363	1,958	7,048
Share of disputes with positive net provider gain (%)	75.9				
(b). Provider-year level					
Net provider gain after fees	117,472	-127	1,124	22,315	231,970
Number of IDR disputes	55.3	1.0	2.0	9.0	56.0
(c). State-quarter level					
Number of providers participated in IDR	127.6	10.0	31.0	87.5	297.4
Number of disputes	3,541.6	69.5	469.0	2,586.2	8,741.0

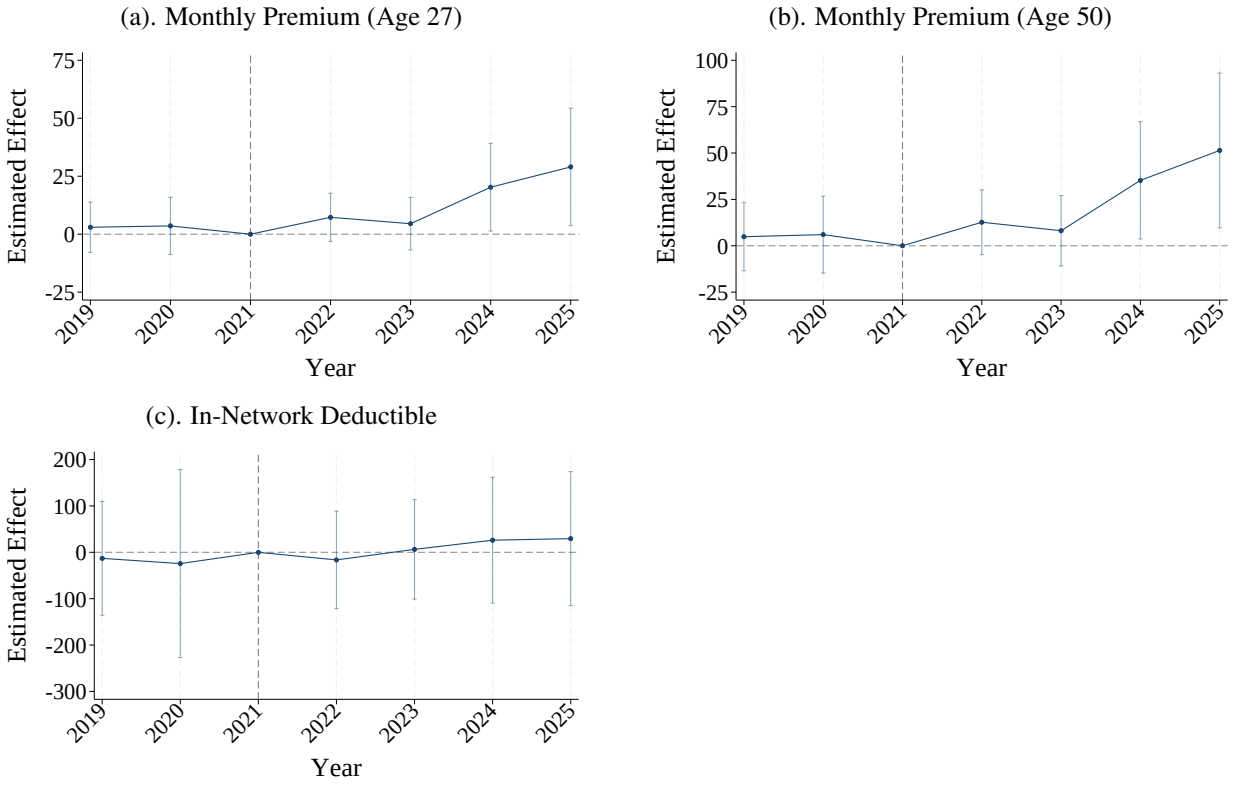
Notes: Panel (a) reports closed federal IDR disputes by quarter and service type. Panel (b) shows the shares resolved for the provider/facility, for the plan/issuer, or by default. Panel (c) reports the quarterly 25th percentile, median, 75th percentile, and 1%-winsorized mean of providers' net IDR gains, defined as the prevailing amount net of expected administrative and arbitration fees, less the issuer's offer. Panel (d) compares issuer offers and prevailing offers as percentages of the QPA; the dashed line marks 100%, and the x-axis is restricted to 0–1000%. Panel (e) summarizes IDR profitability at the case, provider-year, and state-quarter levels. Offer-to-QPA measures are winsorized at the 1st and 99th percentiles, while dollar measures use disputes with recovered payment amounts (Appendix B). Data come from the CMS IDR public-use files, 2023Q1–2025Q2.

Figure 3. Provider-Group Network Participation Around First Favorable IDR Determinations



Notes: Each panel reports an event-study estimate of equation (2), centered on the quarter of a provider group's first favorable IDR determination. A provider group is a provider/facility organization within a state normalized by standardizing capitalization, spacing, and punctuation, and the regression is at the provider group–issuer–quarter level. Panels (a)-(b) include all NSA-exposed providers who have non-missing group information, have ever won a federal arbitration case and have ever appeared in the ACA plan-network data. Panel (a) uses the group–issuer in-network share as the outcome variable, defined as the number of the group's NPIs in the issuer's network that quarter divided by the number of the group's NPIs ever in that issuer's network over the sample period. Panel (b) uses an indicator for complete exit, i.e., when none of the provider group's physicians remain in the issuer's network in a given quarter. Panels (c)-(d) repeat the analysis for non-NSA-exposed (placebo) providers in the same ZIP code as the treated providers, assigned the treated group's event quarter as a placebo event time; A placebo unit is not an observed provider organization: it is the set of physicians of a given placebo specialty practicing in a treated group's ZIP footprint and state, which we treat as a synthetic group and pair with each issuer. Each placebo panel overlays the four control-specialty groups (General Surgery, Primary Care, Psychiatry, and Other), with panel (c) using the in-network share and panel (d) the complete-exit indicator as in panels (a)-(b). Estimates use the imputation estimator of Borusyak et al. (2024). The specification includes provider group–issuer and calendar-quarter fixed effects, as in equation (2). Standard errors are clustered by provider group-issuer; bars are 95% confidence intervals.

Figure 4. Plan Design Changes by Predicted IDR Exposure

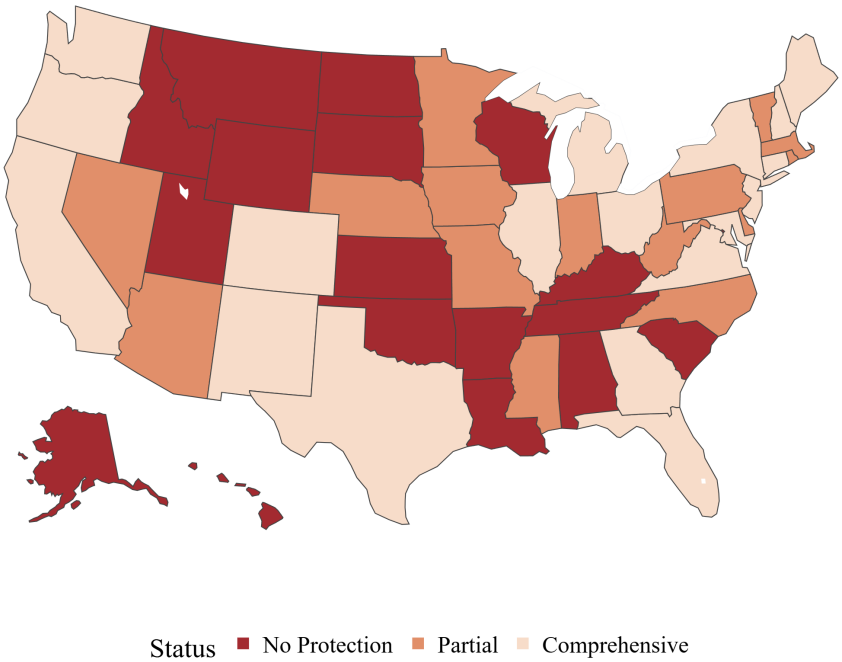


Notes: Each panel plots the event-study analog of equation (3), replacing Post_t with event-time indicators and normalizing the year before treatment to zero. High-exposure issuers are issuer-states at or below the within-state median of the pre-period staffing-firm affiliation share (the share of pre-period in-network NSA-exposed-specialty physicians affiliated with TeamHealth, SCP Health, Envision Healthcare, or Radiology Partners). The outcome variables are the monthly premium for a 27-year-old (panel (a)) and a 50-year-old (panel (b)), and annual in-network deductibles (panel (c)), in dollars. The sample is a plan-year-level panel from the Centers for Medicare and Medicaid Services (CMS) Health Insurance Exchange Public Use Files for 2017-2025. The specification includes issuer-state, state-year, and rating area fixed effects. The specification is estimated on data collapsed to the plan-cell-by-year level and includes plan-cell (issuer by state by metal tier by rating area by market segment) and state-year fixed effects. Estimates use the imputation estimator of Borusyak et al. (2024). Standard errors are clustered by issuer-state; bars are 95% confidence intervals.

Online Appendix

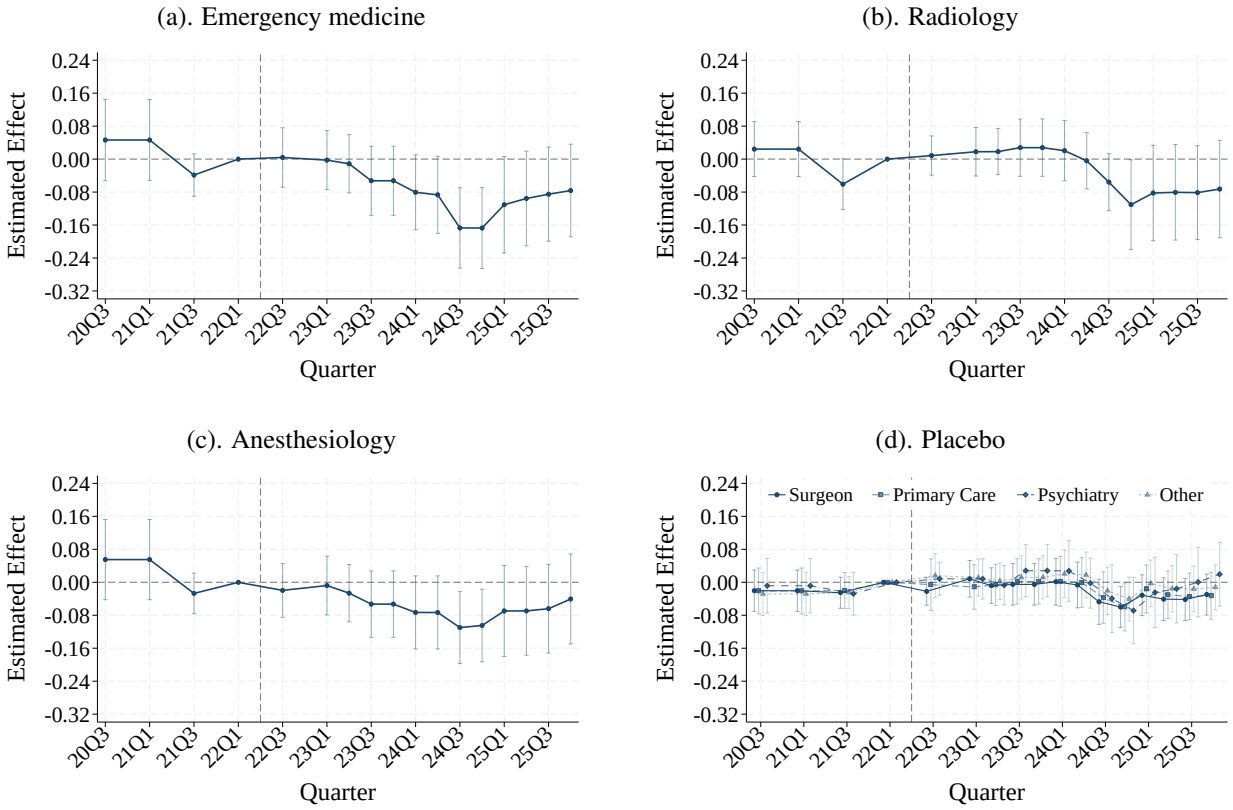
A. Additional Exhibits

Figure A1. State Classification by Pre-NSA Surprise-Billing Protections



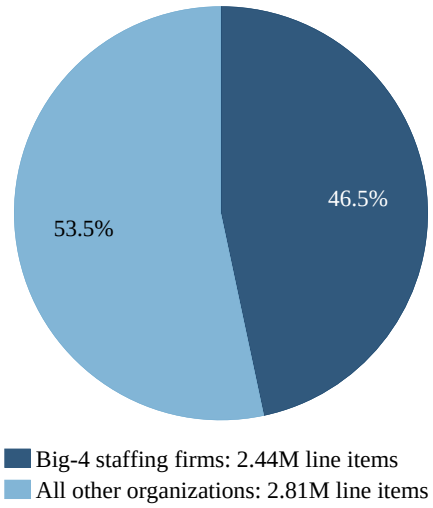
Notes: The map classifies states by their pre-NSA surprise-billing protections, following Ruhter et al. (2023). No-prior-protection states are the primary treated group. Comprehensive-protection states are the main control group. Partial-protection states and states with late pre-NSA changes are excluded from the baseline treatment-control comparison unless otherwise noted. White areas within state boundaries reflect inland water bodies.

Figure A2. Provider Network Participation: Specialty-Specific Difference-in-Differences Event Studies



Notes: Each panel plots event-study estimates from a difference-in-differences design comparing no-prior-protection states to comprehensive-protection states before and after the NSA. The specialty is indicated in each panel. The regression sample and outcome variables (the issuer–state–specialty in-network share) are the same as those in Figure 1. We include insurer-state and calendar quarter fixed effects in the regression. Estimates use the imputation estimator of Borusyak et al. (2024). Standard errors are clustered by issuer–state; bars are 95% confidence intervals.

Figure A3. Share of Federal IDR Dispute Line Items Attributable to Four Large Staffing Firms



Notes: The figure counts dispute line items in the federal IDR public-use files, 2023Q1–2025Q2 (5.25 million line items). A line item is attributed to the big-4 staffing firms (TeamHealth, SCP Health, Envision Healthcare, and Radiology Partners; Fiedler and Adler, 2024) when it carries a firm's email domain or any of its practice entities' group names.

Table A1. Summary Statistics for the Network and Plan Samples

(a). Network estimation sample, by specialty

	Emergency Medicine	Anesthesi-ology	Radiology	General Surgery	Primary Care	Psychiatry	Other
Issuer–state pairs, no protection	89	94	115	144	164	89	156
Issuer–state pairs, comprehensive	91	100	117	121	139	83	133
Ever-present NPIs (Ideon), no prot.	4,319	4,065	5,831	5,363	59,248	3,217	18,896
Ever-present NPIs (Ideon), comp.	9,932	12,484	14,397	10,494	144,111	10,238	47,774
Ever-in NPPES, no protection	14,314	12,670	11,902	12,036	124,498	12,071	38,144
Ever-in NPPES, comprehensive	31,687	35,119	30,425	25,363	309,061	38,257	100,395

(b). Ever-in-network physicians per network, all networks

Specialty	Mean	p10	Median	p90
Emergency Medicine	422	25	205	1,023
Anesthesiology	470	32	228	1,093
Radiology	566	50	307	1,376
Pathology	139	8	68	347
General Surgery	645	111	436	1,505
Primary Care	7,630	1,168	5,052	17,562
Psychiatry	462	27	230	1,200
Other	2,487	384	1,704	5,667

(c). Plan level summary stats

	Mean/N	p25	p50	p75
Gross monthly premium, age 27	411.6	326.4	395.2	478.7
Gross monthly premium, age 50	695.7	552.5	668.9	809.6
In-network deductible	3,828	2,000	3,500	5,500
Number of unique issuer-state pairs	340			
Number of unique plans	62,840			

Notes: Panel (a) summarizes the issuer–state network estimation sample from the Ideon data and NPPES data, all entries are counts of issuer-state pairs of distinct physicians;in-network shares are ever-present-weighted means of the issuer–state in-network share. Panel (b) reports the across-network distribution of the number of distinct physicians ever in-network per specialty; the distribution is strongly right-skewed, and the 10th percentile of emergency-services breadth (25) motivates the minimum-network-size screen. Across the 847 issuer–state–market networks, all but a handful field every analyzed specialty in some quarter; pathology is the only specialty with non-trivial non-coverage (about 2% of networks). Panel (c) reports plan-level summary statistics over the plan-design estimation sample. Premiums are gross monthly premiums (unsubsidized) for a 27- and a 50-year-old; the deductible is the in-network individual deductible.

Table A2. Effect of the NSA on Provider Network Participation: Triple-Difference Estimates

	(1) Emergency Medicine	(2) Anesthesiology	(3) Radiology	(4) General Surgery	(5) Primary Care	(6) Psychiatry
(a). Pooled across years						
NP × Focal × Post	-0.0470	-0.0336	-0.0181	-0.0064	-0.0046	-0.0082
	(0.0277)	(0.0269)	(0.0219)	(0.0048)	(0.0044)	(0.0166)
	[0.090]	[0.212]	[0.409]	[0.184]	[0.299]	[0.622]
Baseline mean	0.348	0.345	0.512	0.595	0.578	0.521
Relative to mean (%)	-13.5	-9.7	-3.5	-1.1	-0.8	-1.6
(b). Separate by years						
NP × Focal × Post 0-3 quarters	-0.0020	-0.0074	0.0132	0.0013	0.0012	-0.0109
	(0.0191)	(0.0179)	(0.0132)	(0.0027)	(0.0028)	(0.0096)
	[0.917]	[0.679]	[0.317]	[0.630]	[0.668]	[0.257]
NP × Focal × Post 4-7 quarters	-0.0631	-0.0547	0.0038	-0.0097	-0.0025	-0.0165
	(0.0314)	(0.0295)	(0.0209)	(0.0064)	(0.0052)	(0.0221)
	[0.044]	[0.064]	[0.856]	[0.130]	[0.631]	[0.455]
NP × Focal × Post 8+ quarters	-0.0662	-0.0369	-0.0536	-0.0094	-0.0098	-0.0009
	(0.0420)	(0.0409)	(0.0362)	(0.0064)	(0.0060)	(0.0262)
	[0.115]	[0.367]	[0.139]	[0.142]	[0.102]	[0.973]
Observations	5,733	5,904	6,406	6,834	7,348	5,490

Notes: Each column reports point estimates, standard errors (in parentheses, clustered by issuer-state), and p-values (in brackets) from the estimation of equation (1) for the listed focal specialty. Estimation uses the imputation estimator of Borusyak et al. (2024). The regression sample and outcome variables (the issuer–state–specialty in-network share) are the same as in Figure 1. Panel (a) pools all post-NSA quarters. Panel (b) replaces Post_t with indicators for 0–3, 4–7, and 8 or more quarters after the NSA.

Table A3. Network-Participation Estimates under Alternative Treatments of Provider-Directory Churn

	(1) Emergency Medicine	(2) Anesthesiology	(3) Radiology	(4) General Surgery	(5) Primary Care	(6) Psychiatry
(a). No bridge (baseline)						
NP × Focal × Post	-0.0470	-0.0336	-0.0181	-0.0064	-0.0046	-0.0082
	(0.0277)	(0.0269)	(0.0219)	(0.0048)	(0.0044)	(0.0166)
	[0.090]	[0.212]	[0.409]	[0.184]	[0.299]	[0.622]
Relative to mean (%)	-13.5	-9.7	-3.5	-1.1	-0.8	-1.6
(b). Bridge gaps (2 quarters)						
NP × Focal × Post	-0.0539	-0.0403	-0.0235	-0.0067	-0.0037	-0.0096
	(0.0277)	(0.0272)	(0.0221)	(0.0049)	(0.0045)	(0.0169)
	[0.052]	[0.138]	[0.287]	[0.166]	[0.405]	[0.572]
Relative to mean (%)	-13.9	-9.8	-4.3	-1.1	-0.6	-1.8
(c). Bridge gaps (4 quarters)						
NP × Focal × Post	-0.0609	-0.0433	-0.0297	-0.0080	-0.0054	-0.0151
	(0.0281)	(0.0273)	(0.0231)	(0.0046)	(0.0046)	(0.0168)
	[0.031]	[0.113]	[0.198]	[0.085]	[0.240]	[0.367]
Relative to mean (%)	-14.9	-9.9	-5.2	-1.3	-0.9	-2.9
(d). Exclude large directory updates (5%)						
NP × Focal × Post	-0.0517	-0.0367	-0.0211	-0.0068	-0.0039	-0.0105
	(0.0272)	(0.0263)	(0.0216)	(0.0049)	(0.0044)	(0.0164)
	[0.057]	[0.163]	[0.327]	[0.167]	[0.377]	[0.524]
Relative to mean (%)	-14.9	-10.6	-4.1	-1.1	-0.7	-2.0
(e). Exclude large directory updates (1%)						
NP × Focal × Post	-0.0467	-0.0323	-0.0180	-0.0064	-0.0044	-0.0077
	(0.0278)	(0.0269)	(0.0219)	(0.0048)	(0.0044)	(0.0167)
	[0.092]	[0.230]	[0.412]	[0.186]	[0.316]	[0.644]
Relative to mean (%)	-13.4	-9.3	-3.5	-1.1	-0.8	-1.5
(f). Exclude large directory updates (change larger than 25%)						
NP × Focal × Post	-0.0488	-0.0345	-0.0202	-0.0069	-0.0040	-0.0079
	(0.0274)	(0.0267)	(0.0215)	(0.0049)	(0.0044)	(0.0162)
	[0.075]	[0.196]	[0.347]	[0.158]	[0.354]	[0.624]
Relative to mean (%)	-14.0	-10.0	-4.0	-1.2	-0.7	-1.5
(g). Exit only if the physician remains active in NPPES						
NP × Focal × Post	-0.0421	-0.0365	-0.0105	-0.0123	-0.0025	-0.0190
	(0.0218)	(0.0202)	(0.0169)	(0.0056)	(0.0046)	(0.0180)
	[0.053]	[0.071]	[0.532]	[0.027]	[0.585]	[0.289]
Baseline mean	0.560	0.567	0.655	0.701	0.582	0.673
Relative to mean (%)	-7.5	-6.4	-1.6	-1.8	-0.4	-2.8
(h). Exit only if the physician remains in another issuer's directory in the state						
NP × Focal × Post	-0.0300	-0.0295	-0.0088	-0.0056	-0.0061	-0.0180
	(0.0213)	(0.0200)	(0.0181)	(0.0063)	(0.0051)	(0.0198)
	[0.160]	[0.141]	[0.626]	[0.374]	[0.231]	[0.364]
Baseline mean	0.733	0.731	0.771	0.814	0.709	0.799
Relative to mean (%)	-4.1	-4.0	-1.1	-0.7	-0.9	-2.3
Baseline mean	0.348	0.345	0.512	0.595	0.578	0.521
Observations	5,733	5,904	6,406	6,834	7,348	5,490

Notes: Each column reports point estimates, standard errors (in parentheses, clustered by issuer-state), and p-values (in brackets) from the estimation of equation (1) for the listed focal specialty. Estimation uses the imputation estimator of Borusyak et al. (2024). The regression sample and outcome variables (the issuer–state–specialty in-network share) are defined similarly as in Figure 1, with the following exception. Panel (a) is the baseline (no bridging). Panels (b)-(c) bridge short directory gaps by treating absences shorter than two or four consecutive quarters as continued in-network participation. Panels (d)-(f) exclude issuer–state–quarter cells with large directory changes, defined as changes in the top 5 percent, changes in the top 1 percent, or changes larger than 25 percent. Panels (g)-(h) address workforce attrition: a physician absent from an issuer's directory counts as an exit only if, in that quarter, the physician (g) still appears in the NPPES registry or (h) is still listed in at least one other issuer's directory in the same state; physicians failing the condition leave the denominator in that quarter, so these panels report their own baseline means.

Table A4. Network-Participation Estimates Excluding IDR-Participating NPIs and Major Staffing Firms

	(1) Emergency Medicine	(2) Anesthesiology	(3) Radiology	(4) General Surgery	(5) Primary Care	(6) Psychiatry
(a). Baseline						
NP × Focal × Post	-0.0470	-0.0336	-0.0181	-0.0064	-0.0046	-0.0082
	(0.0277)	(0.0269)	(0.0219)	(0.0048)	(0.0044)	(0.0166)
	[0.090]	[0.212]	[0.409]	[0.184]	[0.299]	[0.622]
Relative to mean (%)	-13.5	-9.7	-3.5	-1.1	-0.8	-1.6
(b). Exclude IDR-participating NPIs						
NP × Focal × Post	-0.0456	-0.0341	-0.0178	-0.0052	-0.0044	-0.0082
	(0.0282)	(0.0270)	(0.0219)	(0.0048)	(0.0044)	(0.0166)
	[0.106]	[0.207]	[0.417]	[0.280]	[0.319]	[0.623]
Baseline mean	0.3792	0.4013	0.5271	0.5917	0.5802	0.521
Relative to mean (%)	-12.0	-8.5	-3.4	-0.9	-0.8	-1.6
(c). Exclude TeamHealth						
NP × Focal × Post	-0.0469	-0.0335	-0.0181	-0.0064	-0.0046	-0.0082
	(0.0278)	(0.0269)	(0.0219)	(0.0048)	(0.0044)	(0.0166)
	[0.091]	[0.213]	[0.409]	[0.184]	[0.299]	[0.622]
Baseline mean	0.3775	0.4010	0.5273	0.5916	0.5800	0.521
Relative to mean (%)	-12.4	-8.4	-3.4	-1.1	-0.8	-1.6
(d). Exclude SCP Health						
NP × Focal × Post	-0.0465	-0.0336	-0.0181	-0.0064	-0.0046	-0.0082
	(0.0278)	(0.0269)	(0.0219)	(0.0048)	(0.0044)	(0.0166)
	[0.095]	[0.212]	[0.409]	[0.184]	[0.299]	[0.622]
Baseline mean	0.3775	0.4010	0.5273	0.5916	0.5800	0.521
Relative to mean (%)	-12.3	-8.4	-3.4	-1.1	-0.8	-1.6
(e). Exclude Envision						
NP × Focal × Post	-0.0455	-0.0327	-0.0181	-0.0064	-0.0046	-0.0082
	(0.0280)	(0.0270)	(0.0219)	(0.0048)	(0.0044)	(0.0166)
	[0.104]	[0.226]	[0.409]	[0.186]	[0.299]	[0.622]
Baseline mean	0.3804	0.4012	0.5273	0.5916	0.5801	0.521
Relative to mean (%)	-12.0	-8.1	-3.4	-1.1	-0.8	-1.6
(f). Exclude Radiology Partners						
NP × Focal × Post	-0.0470	-0.0336	-0.0180	-0.0064	-0.0046	-0.0082
	(0.0277)	(0.0269)	(0.0219)	(0.0048)	(0.0044)	(0.0166)
	[0.090]	[0.212]	[0.411]	[0.184]	[0.299]	[0.622]
Baseline mean	0.3776	0.4010	0.5273	0.5916	0.5800	0.521
Relative to mean (%)	-12.4	-8.4	-3.4	-1.1	-0.8	-1.6
(g). Exclude American Physician Partners						
NP × Focal × Post	-0.0470	-0.0336	-0.0181	-0.0064	-0.0046	-0.0082
	(0.0277)	(0.0269)	(0.0219)	(0.0048)	(0.0044)	(0.0166)
	[0.090]	[0.212]	[0.409]	[0.184]	[0.299]	[0.622]
Baseline mean	0.3776	0.4010	0.5273	0.5916	0.5800	0.521
Relative to mean (%)	-12.4	-8.4	-3.4	-1.1	-0.8	-1.6

Notes: Each column reports point estimates, standard errors (in parentheses, clustered by issuer-state), and p-values (in brackets) from the estimation of equation (1) for the listed focal specialty. Estimation uses the imputation estimator of Borusyak et al. (2024). The regression sample and outcome variables (the issuer–state–specialty in-network share) are defined similarly as in Figure 1, with the following exception: Panel (b) drops all NPIs that ever appear as IDR-participating providers. Panels (c)-(g) drop, respectively, NPIs affiliated with any of TeamHealth, SCP Health, Envision Healthcare, or Radiology Partners, or American Physician Partners. The former four high-volume filers are selected following Fiedler and Adler (2024). Envision and APP are included to address the possibility that 2023 staffing-firm distress, rather than the IDR outside option, drives the estimated network decline. Firm affiliation is assigned using provider email domains reported in the IDR records, supplemented by normalized provider-group names.

Table A5. Federal IDR Payment Outcomes and Estimated Net Provider Gains

	Mean	Median
IDR-case level		
Provider win probability (%)	85.9	–
Prevailing offer as % of QPA	1,048	374
Net provider gain after fees ($)	2,645	363
Share of disputes with positive net gain (%)	75.9	–
Provider-year level		
Net provider gain after fees ($)	117,472	1,124

Notes: Short-form summary of IDR profitability, corresponding to Panel (e) of Figure 2. "Provider win probability" and "Share of disputes with positive net gain" are binary at the case level, so their medians are uninformative and shown as "–". Net provider gain after fees is the arbitrated amount net of administrative fees and (if the provider loses) the arbitration fee, less the issuer's offer; the 1%-winsorized mean is reported. Data are from the CMS Independent Dispute Resolution public-use files, 2023Q1–2025Q2.

Table A6. Network Participation Around First Favorable IDR Determinations

	(1) NSA-exposed	(2) General Surgery	(3) Primary Care	(4) Other	(5) Psychiatry
(I). Provider-group level, event being group's first favorable determination					
(a). In-network share					
Post	-0.0682	0.0090	0.0031	0.0016	0.0125
	(0.0034)	(0.0013)	(0.0011)	(0.0011)	(0.0017)
	[0.000]	[0.000]	[0.005]	[0.161]	[0.000]
Baseline mean	0.344	0.614	0.563	0.573	0.560
Relative to mean (%)	-19.8	1.5	0.6	0.3	2.2
(b). Complete exit					
Post	0.0765	0.0000	0.0000	0.0000	0.0000
	(0.0036)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
	[0.000]	[·]	[·]	[·]	[·]
Baseline mean	0.600	0.000	0.000	0.000	0.000
Relative to mean (%)	12.8	0.0	0.0	0.0	0.0
Observations	728,194	1,053,260	1,420,577	1,361,170	809,146

Notes: Each cell reports the pooled post-period estimate, its standard error (in parentheses), and p-value (in brackets) from the imputation estimator of Borusyak et al. (2024), replacing the event-time dummies of equation (2) with a single post-period indicator. *Relative to mean* is the estimate as a percentage of the pre-event baseline mean; a bracketed · marks an estimate mechanically fixed at zero (no within-cell variation among the placebo groups). Panels (a)-(b) report the provider-group design. The unit is the provider-group–issuer–quarter and the event is the group's first favorable IDR determination (the earliest quarter the group prevails, in its state); SE clustered by provider group–issuer. A provider group is a provider/facility organization within a state, normalized by standardizing capitalization, spacing, and punctuation. Panel (a)'s outcome is the group–issuer in-network share (the group's NPIs in the issuer's network that quarter divided by the number ever in that issuer's network); panel (b)'s outcome is a complete-exit indicator (one when none of the group's NPIs remain in the issuer's network). Column (1) is all NSA-exposed groups that ever won a federal arbitration case and ever appeared in the ACA plan-network data; columns (2)–(5) are non-NSA-exposed (placebo) groups in the same ZIP code, assigned the nearby treated group's event quarter. Unlike column (1), a placebo unit is not an observed provider organization but the set of physicians of a given placebo specialty practicing in a treated group's ZIP footprint and state, treated as a synthetic group and paired with each issuer.

Table A7. Network Participation Around First Favorable IDR Determinations, by Size of the Event

	(1) In-network share	(2) Complete exit
(a). Provider − insurer offer, % of QPA		
Post × Above median	-0.0866	0.0921
	(0.0054)	(0.0056)
	[0.000]	[0.000]
Post × Below median	-0.0586	0.0632
	(0.0060)	(0.0062)
	[0.000]	[0.000]
Baseline mean (above / below)	0.369 / 0.335	0.592 / 0.615
Relative to mean (%)	-23.5 / -17.5	15.6 / 10.3
Observations	309,643 / 278,540	
(b). Provider − insurer offer, dollars		
Post × Above median	-0.0941	0.1000
	(0.0065)	(0.0068)
	[0.000]	[0.000]
Post × Below median	-0.0409	0.0478
	(0.0066)	(0.0069)
	[0.000]	[0.000]
Baseline mean (above / below)	0.354 / 0.324	0.602 / 0.620
Relative to mean (%)	-26.6 / -12.6	16.6 / 7.7
Observations	219,317 / 212,078	
(c). Provider offer, % of QPA		
Post × Above median	-0.0886	0.0941
	(0.0054)	(0.0056)
	[0.000]	[0.000]
Post × Below median	-0.0580	0.0626
	(0.0060)	(0.0062)
	[0.000]	[0.000]
Baseline mean (above / below)	0.369 / 0.334	0.591 / 0.616
Relative to mean (%)	-24.0 / -17.4	15.9 / 10.2
Observations	311,847 / 277,115	
(d). Provider offer, dollars		
Post × Above median	-0.0930	0.1003
	(0.0065)	(0.0068)
	[0.000]	[0.000]
Post × Below median	-0.0427	0.0489
	(0.0066)	(0.0069)
	[0.000]	[0.000]
Baseline mean (above / below)	0.356 / 0.322	0.599 / 0.624
Relative to mean (%)	-26.1 / -13.3	16.7 / 7.8
Observations	220,476 / 210,919	

Notes: Each cell reports the pooled post-period estimate, its standard error (in parentheses), and p-value (in brackets) from the imputation estimator of Borusyak et al. (2024), replacing the event-time dummies of equation (2) with a single post-period indicator, estimated separately for provider groups above and below the median size of their first favorable event. The unit, event, sample, and outcomes are those of Table A6, column (1); SE clustered by provider group–issuer. Baseline mean, Relative to mean, and Observations are reported for the above- and below-median subsamples, respectively. Event size is measured on the group's favorable disputes in its first favorable quarter as the gap between the provider's and the insurer's offers (panels (a)–(b)) or the provider's offer (panels (c)–(d)), in percent of the qualifying payment amount (panels (a) and (c)) or in dollars (panels (b) and (d), QPA-recovered sample); dollar amounts are summed across the line items of a dispute and understate the amounts at stake in batched disputes.

Table A8. ACA Marketplace Premiums by Insurers' Predicted IDR Exposure

	(1) Premium (Age 27)	(2) Premium (Age 50)
(a). Pool across years		
High Exposure × Post	15.25	26.83
	(7.45)	(12.34)
	[0.041]	[0.030]
Baseline mean	420.51	705.95
Relative to mean (%)	3.6	3.8
(b). Separate by years		
High Exposure × Post 0-1 years	5.87	10.34
	(5.24)	(8.78)
	[0.263]	[0.239]
High Exposure × Post 2+ years	24.63	43.31
	(10.96)	(18.13)
	[0.025]	[0.017]
Baseline mean	420.51	705.95
Relative to mean, 0-1 years	1.4	1.5
Relative to mean, 2+ years	5.9	6.1
Observations	59,067	59,067

Notes: Each column reports point estimates, standard errors (in parenthesis, clustered by issuer-state), and p-values (in brackets) from the estimation of equation (3) for the listed plan characteristics. "High Exposure" indicates issuer-states at or below the within-state median of the pre-period staffing-firm affiliation share (the share of pre-period in-network affected-specialty physicians affiliated with TeamHealth, SCP Health, Envision Healthcare, or Radiology Partners), treated at the 2022 plan year. Panel (a) reports the pooled post-period estimate (the average of the post-period imputation coefficients). Panel (b) separates it by event time: "Post 0–1 years" is the average effect over the first two post-treatment years (2022–2023) and "Post 2+ years" the average over the later post years (2024–2025). Estimation uses the imputation estimator of Borusyak et al. (2024). The regression sample and outcome variables are the same as in Figure 4.

B. Sample Construction

This appendix describes how we recover the dollar-valued Qualifying Payment Amount (QPA) and party offers that underlie our measure of the gain from arbitration in the dispute-level analysis sample from the Centers for Medicare & Medicaid Services (CMS) Federal Independent Dispute Resolution (IDR) Public Use Files (PUFs). The core difficulty is that the dollar amounts and the dispute characteristics are reported in separate tables that share no common identifier. We resolve this through a probabilistic record-linkage procedure that exploits the offer ratios reported in both tables.

Raw Data. Our primary source is the set of quarterly Federal IDR PUFs released by CMS, covering disputes with payment determinations from the first quarter of 2023 through the second quarter of 2025. Each quarterly release comprises three tables: (1) the Emergency and Non-Emergency table, recording out-of-network disputes for emergency and certain non-emergency items and services; (2) the Air Ambulance table, recording out-of-network air ambulance disputes; (3) the QPA and Offers table, recording the dollar-valued QPA, provider, plan, and prevailing offers.

The two dispute tables are the unit of analysis. Each record identifies the disputing parties (the provider National Provider Identifier (NPI), and the health plan or issuer), the service (procedure code and place-of-service code), the location of service, the type of dispute (single line item or batched), the initiating party, the payment determination outcome, the arbitrator (IDRE) compensation, and the administrative characteristics of the case. Critically, the dispute tables report each party's offer only as a percentage of the QPA and do not report the underlying dollar amounts. Those dollar amounts appear only in the QPA and Offers table with anonymized dispute number. Recovering the dollar QPA and offers for each dispute therefore requires a matching procedure.

The Recovery Problem and Linking Variables Our objective is to recover, for each dispute line item, the dollar-valued QPA and the corresponding provider and issuer offers reported in the QPA and Offers table. The key linking information is the pair of offer ratios reported in the dispute tables. Let r^P and r^H denote the provider and issuer offers as fractions of the QPA. In the QPA and Offers table, these ratios can be reconstructed as

$$r^P = \frac{O^P}{Q}, \qquad r^H = \frac{O^H}{Q}$$

where Q is the QPA and O^P and O^H are the provider and issuer dollar offers. Records referring to the same line item should therefore agree, up to rounding, on both offer ratios and on the relevant categorical dispute characteristics.

We match at the line-item level. For batched disputes, the unit is the dispute line item; for single

disputes, it is the dispute itself. Air ambulance disputes are matched separately from emergency and non-emergency disputes.

Matching Algorithm. We implement the match separately by quarter using a sequential linkage procedure. At each step, a line item is assigned recovered dollar amounts only if the matched candidate records imply a unique QPA and unique provider and issuer offers. Otherwise, the line item remains unresolved and proceeds to the next step.

The baseline match requires agreement on the categorical key

$$K = (\text{procedure code},\ \text{place of service},\ \text{selected offer},\ \text{default decision},\ \text{line-item type})$$

and approximate agreement on both offer ratios:

$$K_i = K_j, \qquad |r_i^P - r_j^P| \le \epsilon, \qquad |r_i^H - r_j^H| \le \epsilon,$$

with $\epsilon = 0.01$. Candidate matches are retained only when they imply a unique set of recovered dollar amounts.

For line items not resolved by the baseline match, we successively add additional restrictions. First, we augment the key with the provider's geographic region, constructed from the provider NPI, NPPES practice address history, and a ZIP-to-region crosswalk. Second, for the remaining unresolved observations, we impose within-stratum rank consistency in offers, comparing the relative position of offers within cells defined by procedure code and region. This final restriction is used only to break residual ambiguities among otherwise observationally similar candidates.

The procedure yields a line-item file with recovered QPA, provider offer, issuer offer, and the matching step at which the recovery was made.

Collapse to Disputes and Usable Sample. We collapse recovered line items to the dispute level. For single-item disputes, the dispute-level QPA and offers are the recovered line-item amounts. For batched disputes, QPA and offers are aggregated across line items. We retain a dispute in the analysis sample only when the recovered amounts are internally coherent. In particular, batched disputes must have sufficient line-item information to aggregate QPA and offers without ambiguity.

Aggregating to the dispute level, the overall recovery rate is 54.3 percent.

Constructing the Gain from Dispute. For each usable dispute, we measure each party's realized gain from arbitration relative to accepting the opposing party's offer. Let O^P and O^H denote the recovered provider and issuer offers, and let

$$O^* = \begin{cases} O^P, & \text{if the provider prevails,} \\ O^H, & \text{if the issuer prevails.} \end{cases}$$

Each party pays the non-refundable administrative fee A, and the losing party pays the IDRE fee F. Let $F^P = F$ if the provider loses and zero otherwise, and let $F^H = F$ if the issuer loses and zero otherwise. The provider's net receipt and the issuer's net payment are

$$R = O^* - A - F^P, \qquad P = O^* + A + F^H.$$

We define the provider's and issuer's gains from dispute as

$$G^{\text{prov}} = R - O^H, \qquad G^{\text{iss}} = O^P - P.$$

Thus, the provider's gain is the net arbitration receipt relative to the issuer's offer, while the issuer's gain is the avoided payment relative to the provider's offer, net of arbitration costs. These measures form the basis for the arbitration-value outcomes used in the dispute-level analysis.

C. Staffing-Firm Affiliation

This appendix describes how physicians are matched to the staffing firms used to construct the insurer exposure measure in Section 5 and the affiliation-based robustness checks in Section 3.

Attributing Dispute Records to Firms. The source is the Emergency and Non-Emergency table of the quarterly federal IDR PUFs, 2023Q1–2025Q2, the same files as in Appendix B. Each dispute line item reports the provider NPI, the provider email domain, and the provider/facility group name. A line item is attributed to a firm if either (i) its email domain belongs to the firm's set of known domains, following Fiedler and Adler (2024) and including domains of the firms' revenue-cycle-management subsidiaries, or (ii) its group name, normalized by uppercasing and collapsing non-alphanumeric characters, contains the firm's name. The domain criterion does most of the work: the firms file through many subsidiary group names, so name matching alone would substantially understate their volume. American Physician Partners publishes no email domain and is matched on group name only.

NPI Rosters. An NPI is affiliated with a firm if any of its dispute line items is attributed to that firm. Of 17,332 distinct disputing NPIs, 9,980 are affiliated with at least one of the four firms. Because the rosters are constructed from arbitration records, they contain only physicians who appear in at least one dispute; they are a lower bound on the firms' rosters, but capture exactly the physicians who generate arbitration volume.

Issuer Exposure. For each issuer-state, the staffing-firm affiliation share in Section 5 is the share of the issuer's pre-period in-network physicians in NSA-exposed specialties whose NPI appears on any of the four rosters, with the pre-period network measured over 2020Q1–2022Q2. High-exposure issuer-states are those at or below the within-state median of this share.