USC Schaeffer
BROOKINGS

WORKING PAPER

How much can an out-of-network cap reduce in-network prices?

Matthew Fiedler and Bich Ly

June 2023

ABSTRACT

Capping the prices health care providers can collect for out-of-network services is a commonly proposed strategy for reducing the in-network prices negotiated by providers and insurers. We present a model that implies that an out-of-network cap can greatly reduce in-network prices when providers must accept out-of-network patients (i.e., in emergency settings) but may be much less effective in other settings. In these other settings, the achievable price reductions can be bounded using estimates of how much volume a provider retains when it shifts out-of-network—and at what price. We use a large national claims database to examine episodes in which hospitals change network status and estimate that hospitals that shift out-of-network retain only 12% of their non-emergency in-network volume, albeit at prices more than twice as high. Using these estimates to calibrate the model-derived bound implies that an out-of-network cap can reduce in-network prices by at most 19% in non-emergency settings. This bound suggests that policymakers wishing to greatly reduce in-network prices in commercial insurance may need to consider other policy tools and that competition from traditional Medicare, not the presence of an out-of-network cap, may be the main reason that Medicare Advantage plans negotiate lower prices than commercial plans.

ACKNOWLEDGEMENTS AND DISCLOSURES

The authors thank Loren Adler, Erin Duffy, Richard Frank and seminar participants at the Brookings Institution and the Medicare Advantage Data Lab for helpful comments and conversations. We thank Conrad Milhaupt for excellent research assistance and Caitlin Rowley for excellent editorial assistance. This work was supported by a grant from Arnold Ventures. All errors are our own.

This working paper is part of the USC-Brookings Schaeffer Initiative for Health Policy, which is a partnership between Economic Studies at Brookings and the University of Southern California Schaeffer Center for Health Policy & Economics. The Initiative aims to inform the national health care debate with rigorous, evidence-based analysis leading to practical recommendations using the collaborative strengths of USC and Brookings.

Other than the aforementioned, the authors did not receive financial support from any firm or person for this article or from any firm or person with a financial or political interest in this article. The authors are not currently an officer, director, or board member of any organization with a financial or political interest in this article. The Brookings Institution is financed through the support of a diverse array of foundations, corporations, governments, individuals, as well as an endowment. A list of donors can be found in our annual reports published online. The findings, interpretations, and conclusions in this report are solely those of its author(s) and are not influenced by any donation.

One commonly proposed strategy for reducing the prices of health care services is to limit how much health care providers can collect for out-of-network care (e.g., Murray 2013; Berenson et al. 2015; Song 2017; Kane 2019; Chernew, Pany, and Frank 2019; Melnick and Fonkych 2020b). Capping out-of-network prices would directly reduce prices paid for out-of-network services, but many argue that it could also reduce the prices that providers and insurers negotiate for in-network care (e.g., Duffy, Whaley, and White 2020; Chernew, Dafny, and Pany 2020; Prager and Tilipman 2020; Berenson and Murray 2022). Similarly, the out-of-network cap that exists in Medicare Advantage—which limits what providers can collect for out-of-network care to what traditional Medicare would pay for that care—is often cited as a major reason that the in-network prices Medicare Advantage plans negotiate for most services are similar to traditional Medicare's prices, even as the same insurers negotiate much higher prices in their commercial plans (e.g., Berenson et al. 2015; CBO 2017; MedPAC 2017; Trish et al. 2017; Maeda and Nelson 2018; Pelech 2020; Murray and Keane 2022).

The view that an out-of-network cap can markedly reduce in-network prices has an intuitive logic: if an insurer knows that it can access a provider's services at the capped price if network negotiations break down, then it can credibly refuse to sign a network agreement at a higher price and leave the provider no choice but to agree to a price at or below the cap. This logic is compelling if an insurer's enrollees can access a provider's services without a network agreement. This is the case for emergency care, which hospitals must deliver regardless of patients' insurance status under federal law. However, as emphasized by Fiedler (2020), providers can decline to treat out-of-network patients in non-emergency situations. Thus, in these settings, an insurer may jeopardize its enrollees' access to a provider's services if it breaks off negotiations and seeks to rely on the capped price. This, in turn, may limit how much leverage an insurer can derive from an out-of-network cap.

This paper seeks to quantitatively assess how much an out-of-network cap can reduce in-network prices in non-emergency settings. To that end, we rely on a model of provider-insurer bargaining under an out-of-network cap that is adapted from Fiedler (2020). In the model, network negotiations happen in two stages. In the first stage, the insurer and provider announce what actions they will take if negotiations break down; namely, the insurer announces what coverage it will offer for out-of-network services, while the provider announces whether it will accept out-of-network patients (if it is permitted to make that choice) and at what price. In the second stage, the provider and insurer engage in Nash bargaining where the disagreement payoffs reflect the previously announced actions.

Consistent with the informal logic laid out above, equilibrium outcomes under an out-of-network cap depend crucially on whether the provider has the option to turn away out-of-network patients. When the provider cannot turn away out-of-network patients, the negotiated in-network price is always at or below the cap, essentially because the insurer can always opt to break off negotiations and rely on the capped price. By contrast, when a provider can turn away out-of-network patients, the provider opts to do so once the cap becomes sufficiently stringent. Past that point, tightening the cap has no effect on the parties' bargaining positions (since no out-of-network care is delivered) and, thus, no effect on the negotiated in-network price. Correspondingly, the scope for an out-of-network cap to reduce in-network prices is smaller, albeit not zero, since forgoing all out-of-network volume leaves the provider in a somewhat worse bargaining position than without an out-of-network cap.

We use the model to derive an upper bound on how much an out-of-network cap can reduce in-network prices in settings where providers can turn away out-of-network patients. This upper bound depends on two "sufficient statistics": the share of its volume a provider would retain if it shifted out-of-network in a world without an out-of-network cap, and the price it would receive on that retained volume. The role of these two statistics is intuitive. If a provider expects to retain little volume (or receive a low price) if it goes out-of-network in a world *without* a cap, then the provider can put itself in a similar bargaining position in a world *with* a cap by simply refusing to accept out-of-network patients, thereby largely defanging the cap. To supplement this bound, we also derive an exact expression for the maximum amount an out-of-network cap can reduce in-network prices; this expression depends on the same two statistics but also requires an estimate of how the provider's marginal cost compares to the in-network price plus an additional functional form assumption.

The extant literature contains little evidence on how volume and prices change when a provider leaves an insurer's network, with the notable exception of one recent study of a decision by a single California hospital chain to terminate its contracts with all private insurers (Melnick and Fonkych

2020a). This paper aims to help fill this gap by using a large database of health care claims from the Health Care Cost Institute that covers the years 2014-2017 to identify episodes where a hospital enters or leaves a particular plan's network and then estimate how volume and prices change around those network status changes.

Our preferred estimates are that a hospital that shifts out-of-network retains only 12% of its non-emergency in-network volume, while the prices it receives for that care increase to 234% of in-network prices, on average. We use these estimates to calibrate the model-derived upper bound on the effect of an out-of-network cap, which yields an estimate that an out-of-network cap can reduce in-network prices for non-emergency services by at most 19%. Importantly, the maximum achievable reduction may be smaller than this. Indeed, when we calibrate the exact expression for the maximum reduction in in-network prices achievable using an out-of-network cap that was derived under the stronger assumptions described above, we obtain a point estimate of only 13%. Because non-emergency care accounts for the majority of health care spending, including about two-thirds of hospital spending (Fiedler 2020), these results suggest that there are meaningful limits on how much an out-of-network cap can reduce in-network prices, contrary to what analyses that ignore providers' ability to turn away out-of-network patients suggest (Duffy, Whaley, and White 2020; Prager and Tilipman 2020).

Our approach does not account for the greater administrative and reputational costs hospitals may incur when seeking to collect the large patient cost-sharing obligations and "balance bills" commonly associated with out-of-network care (Biener et al. 2021; Cooper, Scott Morton, and Shekita 2020; Pelech 2020; Song et al. 2020). Thus, our approach may overstate how costly it would be to hospitals to forgo out-of-network volume in response to an out-of-network cap and, correspondingly, overstate how much an out-of-network cap would reduce in-network prices for non-emergency services.

On the other hand, it is notable that we find that hospitals that shift out-of-network retain around three-quarters of their *emergency* volume—at prices roughly twice their in-network prices—causing total revenues for emergency services to rise markedly. As noted above, collecting payment for out-of-network care may be burdensome for hospitals, especially in the case of emergency care, so being out-of-network may not be as attractive with respect to emergency services as these results suggest.

Nevertheless, this finding raises the possibility that some of the pricing power hospitals hold with respect to emergency services may be "shifted" into the prices that they negotiate for non-emergency services (Pope 2019; Melnick and Fonkych 2020a). If this were the case, then our estimates would overstate the "true" in-network prices hospitals receive for non-emergency services and, thus, understate how much the prices of these services increase when the hospital shifts out-of-network. This would lead us to understate how much an out-of-network cap could reduce in-network prices for non-emergency services. If shifting occurs, it would also mean that the share of current hospital spending traceable to emergency care—and, thus, the share of spending that occurs in settings where out-of-network caps can be highly effective—is somewhat higher than it appears.

Our results have important implications for policymakers. First, our results suggest that while there is some scope to reduce the prices that commercial plans pay for in-network services by capping out-of-network prices, policymakers wishing to greatly reduce in-network prices may need to consider other approaches. These alternative approaches could include directly limiting in-network prices or coupling an out-of-network cap with a requirement to accept out-of-network patients. Second, our results suggest that the out-of-network cap present in the Medicare Advantage (MA) program is not the primary reason that MA plans are able to negotiate prices that are close to traditional Medicare's (and, correspondingly, far below the prices the same insurers negotiate in the commercial market). Rather, competition from traditional Medicare may play the leading role in disciplining provider prices in MA. This suggests that the prices that MA plans pay providers will begin to rise if traditional Medicare's competitive position becomes sufficiently weak, a plausible scenario in light of the ongoing steady decline in traditional Medicare's market share (Freed et al. 2022).

The rest of the paper proceeds as follows. We first present the model of provider-insurer bargaining under an out-of-network cap that provides the conceptual framework for this paper. Next, we describe our data and empirical methodology, after which we present our results. We then use our empirical results to calibrate the key expressions derived from the model. We conclude by discussing the policy implications of our findings.

1 Model

We begin by describing the model of provider-insurer negotiations under an out-of-network cap that motivates the paper's empirical analyses, which is largely adapted from Fiedler (2020). The model features a provider and an insurer that bargain over a network agreement in two stages. In the first stage, the insurer and provider announce what actions they will take if negotiations break down; the insurer announces what coverage it will offer for out-of-network services, while the provider announces whether it will accept out-of-network patients (if it is permitted to make that choice) and at what price. In the second stage, a network agreement is reached via Nash bargaining where the disagreement payoffs reflect the actions announced in the first stage. The rest of this section first provides a detailed description of the model's primitives, bargaining process, and equilibrium. It then shows how the scope for an out-of-network cap to reduce in-network prices in settings where providers can turn away out-of-network patients can be related to observable statistics.

Other recent work has also modeled how regulating out-of-network prices might affect in-network prices (Duffy, Whaley, and White 2020; Prager and Tilipman 2020). A key difference between the model presented here and the frameworks used in these other analyses is that we allow the provider to choose to turn away out-of-network patients.[1] As will become clear below, allowing the provider to make this choice greatly reduces the scope for an out-of-network cap to reduce in-network prices.

1.1 Model primitives

The provider receives a price $p \in \mathbb{R}$ for its services and makes a choice $a \in \mathcal{A}$ about whether to accept patients (where $\mathcal{A} = \{1\}$ if the provider must accept patients and $\mathcal{A} = \{0,1\}$ otherwise). The insurer offers a level of coverage $l \in [0,1]$ for the provider's services. Setting $l = 1$ corresponds to offering the most generous possible coverage, while setting $l = 0$ corresponds to offering no coverage at all. The level of coverage should be understood to encompass both cost-sharing requirements and other plan design features that affect utilization, like prior authorization procedures.

[1] Our model also explicitly endogenizes the insurer's choice of what level of out-of-network coverage to offer, which Prager and Tilipman (2020) do not and Duffy, Whaley, and White (2020) do only implicitly. This feature of our model tends to *increase* the scope for an out-of-network cap to reduce prices in settings where providers cannot turn away patients.

To specify the parties' payoffs, we follow Gowrisankaran, Nevo, and Town (2015) and assume that the insurer has fixed enrollment (normalized to one) but is subject to long-term competitive pressures that lead it to act as a good agent for enrollees. Thus, the insurer seeks to maximize the value its enrollees receive from the provider's care, net of the premium required to finance that care. To that end, we let $Q(p, l)$ denote the quantity of the provider's services that the insurer's enrollees use when the provider's price is p, the insurer's level of coverage is l, and the provider accepts patients; we let $V(Q)$ denote the value the insurer's enrollees derive from consuming a quantity of services Q. The insurer's payoff W is then given by

$$W(p, l) \equiv V\big(Q(p, l)\big) - pQ(p, l)$$

when the provider accepts patients and zero otherwise. For its part, the provider incurs a marginal cost c to deliver each service and seeks to maximize its profits, so its payoff π is given by

$$\pi(p, l) \equiv Q(p, l)[p - c]$$

when it chooses to accept patients and zero otherwise.

To facilitate the subsequent analysis, we make a few largely standard assumptions about these primitives. We assume that V is twice differentiable, strictly increasing, and strictly concave, with $V(0) = 0$. We further assume that $V'(0) > c$ and that $V'(Q) < c$ for large enough values of Q. This implies that there is a unique "efficient" quantity of care $Q^* > 0$, which satisfies $V'(Q^*) = c$.

We assume that Q is twice differentiable, strictly decreasing in p for $l < 1$ and strictly increasing in l, with $Q_p(p, 1) = 0$. We assume that the insurer's level of coverage affects demand for the provider's services enough to ensure that $Q(p, 0) < Q^* < Q(p, 1)$ for any price $p \in \mathbb{R}$. Additionally, we assume that $Q_{pl}(p, l) \geq 0$, indicating that greater coverage reduces price sensitivity.

We also make a more technical assumption, which we label Assumption QC for later reference.

Assumption QC. *For any $p \in \mathbb{R}$ and $l \in [0,1]$,*

$$\frac{d}{dp}\big[V'\big(Q(p, l)\big)\big] \equiv V''(Q(p, l))Q_p(p, l) \leq 1,$$

with the inequality strict for $l > 0$. Additionally, for any $p \in \mathbb{R}$ and $l < 1$, there exists $\epsilon > 0$ such that

$$\frac{d}{dp}\left[-\frac{Q(p,l)}{Q_p(p,l)}\right] < 1 - V''(Q(p,l))Q_p(p,l) - \epsilon.$$

The first part of this assumption requires that the value enrollees place on the marginal service rises no more than one-for-one with the price; in essence, this requires a limited degree of consistency between enrollees' demand behavior (captured in Q) and their underlying valuation of care (captured in V). The second part of the assumption limits how quickly the price sensitivity of enrollees' demand for care can diminish as the price rises; put another way, it ensures that the demand curve is not "too convex." Together, the two parts of Assumption QC ensure that the parties' best response functions in the bargaining game that we specify below are well-defined and monotonic, which in turn ensures that the model has a well-behaved equilibrium.

1.2 Bargaining process

The provider and insurer negotiate a network agreement in two stages.

In the first stage, the two parties simultaneously announce the actions they will take if network negotiations subsequently break down. Specifically, the provider announces a decision $\tilde{a} \in \mathcal{A}$ about whether it will accept out-of-network patients and an out-of-network price $\tilde{p} \in \mathbb{R}$, with $\tilde{p} \leq \bar{p}$ when the provider is subject to an out-of-network cap of $\bar{p}$. This out-of-network price should be understood to reflect what the provider expects to be able to collect for out-of-network care, net of any discounts or bad debt, so it may be less than the provider's "chargemaster" price. The insurer announces the level of out-of-network coverage it will provide $\tilde{l} \in [0,1]$.

In the second stage, the parties bargain over a network agreement that specifies an in-network price $p^* \in \mathbb{R}$ and coverage level $l^* \in [0,1]$.[2] As has become standard in the literature on provider-insurer bargaining (e.g., Gowrisankaran, Nevo, and Town 2015; Clemens and Gottlieb 2016; Ho and Lee 2017; Cooper, Scott Morton, and Shekita 2020), we assume that the outcome of negotiations is determined by Nash bargaining, so the negotiated in-network price p^* and coverage level l^* satisfy

[2] Technically, the network agreement also specifies whether the provider will accept the insurer's patients, but it will always be in the parties' mutual interest for the provider to accept patients, so we omit that detail here to streamline notation.

$$(p^*, l^*) = \underset{\substack{p\in\mathbb{R}, l\in[0,1] \text{ s.t.} \\ W(p,l)\geq\widetilde{W}, \pi(p,l)\geq\tilde{\pi}}}{\operatorname{argmax}} \left[W(p,l) - \widetilde{W}\right]^{\theta} \times [\pi(p,l) - \tilde{\pi}]^{1-\theta}, \tag{1}$$

where $\widetilde{W} \equiv \tilde{a}W(\tilde{p}, \tilde{l})$ and $\tilde{\pi} \equiv \tilde{a}\pi(\tilde{p}, \tilde{l})$ are "disagreement payoffs" that reflect the actions the parties will take if negotiations break down, and $\theta \in (0,1)$ is the insurer's bargaining weight.

This structure gives rise to a one-shot non-cooperative game in which the strategies are the disagreement actions that the parties announce prior to bargaining ($\tilde{p}$, $\tilde{l}$, and $\tilde{a}$), and the parties' payoffs are determined by the second-stage Nash bargaining problem specified in equation (1).

Importantly, this framework assumes that the actions announced in the first stage will actually be implemented if negotiations end up breaking down. This assumption may seem unappealing since it will often be in the parties' interest to announce actions that they would prefer to renege on. For example, as we show below, the provider can often benefit from announcing that it will turn away the insurer's patients in the absence of a network agreement but prefer to accept those patients if negotiations actually break down since they are profitable (for any $\tilde{p} > c$). However, it is natural to think that reputational considerations would lead the parties to follow through on their threatened actions in practice, and this intuition can be formalized. Indeed, the model considered here is a member of the broader class of "Nash bargaining with threats" models due to Nash (1953). Abreu and Pearce (2007) show that the Nash bargaining with threats framework can be viewed as the reduced form of a rich class of multi-period contract negotiation models in which parties build reputations based on their bargaining demands and actions.

1.3 Equilibrium

We now characterize the model's equilibrium. To start, we solve for the outcome of the second-stage Nash bargaining problem specified in equation (1). Appendix A shows that this problem has a unique solution that satisfies two intuitive conditions:

$$Q(p^*, l^*) = Q^* \tag{2}$$

$$p^*Q^* = \theta c Q^* + (1-\theta)V(Q^*) + \theta\tilde{\pi} - (1-\theta)\widetilde{W}. \tag{3}$$

Equation (2) says that, regardless of the disagreement payoffs, the parties always agree on a combination of a price p^* and coverage terms l^* that result in delivery of the efficient quantity of care Q^*. If this was not the case, then there would be an alternative agreement with higher total surplus, which the parties could split between them by suitably adjusting the negotiated price.

Because the second-stage agreement always results in a quantity Q^*, the provider will select its first-stage actions to maximize the negotiated price, while the insurer will seek to minimize this price. Equation (3) shows that the negotiated price is the sum of: a term $\theta c Q^* + (1-\theta)V(Q^*)$ that is a weighted average of the provider's cost of delivering the efficient quantity of care and the value the insurer's enrollees derive from that care; and a term $\theta\tilde{\pi} - (1-\theta)\widetilde{W}$ that depends on the disagreement payoffs. Crucially, the form of this second term shows that the provider can secure a higher in-network price by *either* increasing its disagreement payoff $\tilde{\pi}$ or reducing the insurer's disagreement payoff $\widetilde{W}$. Similarly, the insurer can secure a lower in-network price by either increasing $\widetilde{W}$ or reducing $\tilde{\pi}$.

We next characterize the equilibrium actions that the provider and insurer will announce in the first stage and, in turn, the equilibrium negotiated outcomes. The proposition below summarizes how the equilibrium outcomes depend on the presence of an out-of-network cap and whether the provider has the option to reject out-of-network patients. A proof is in Appendix B.

Proposition 1. *With no out-of-network cap or a cap $\bar{p} \geq c$, the model has at least one pure strategy Nash equilibrium, and all equilibria have the same in-network price. Furthermore, the equilibrium in-network prices and out-of-network actions satisfy the following:*

(i) *If there is no cap: the provider always accepts out-of-network patients; the out-of-network quantity $\tilde{Q}_{\text{nocap}} \equiv Q(\tilde{p}_{\text{nocap}}, \tilde{l}_{\text{nocap}})$ associated with the (unique) equilibrium actions $\tilde{p}_{\text{nocap}}$ and $\tilde{l}_{\text{nocap}}$ satisfies $\tilde{Q}_{\text{nocap}} < Q^*$; and the in-network price p^*_{nocap} satisfies $p^*_{\text{nocap}} < \tilde{p}_{\text{nocap}}$.*

(ii) *If there is a cap and the provider cannot reject out-of-network patients, the in-network price p^*_{accept} is a differentiable function of $\bar{p}$ that satisfies: $p^*_{\text{accept}}(\bar{p}) = p^*_{\text{nocap}}$ for $\bar{p} \geq \tilde{p}_{\text{nocap}}$; $(p^*_{\text{accept}})'(\bar{p}) > 0$ and $p^*_{\text{accept}}(\bar{p}) < \bar{p}$ for $\bar{p} \in (c, \tilde{p}_{\text{nocap}})$; and $p^*_{\text{accept}}(c) = c$.*

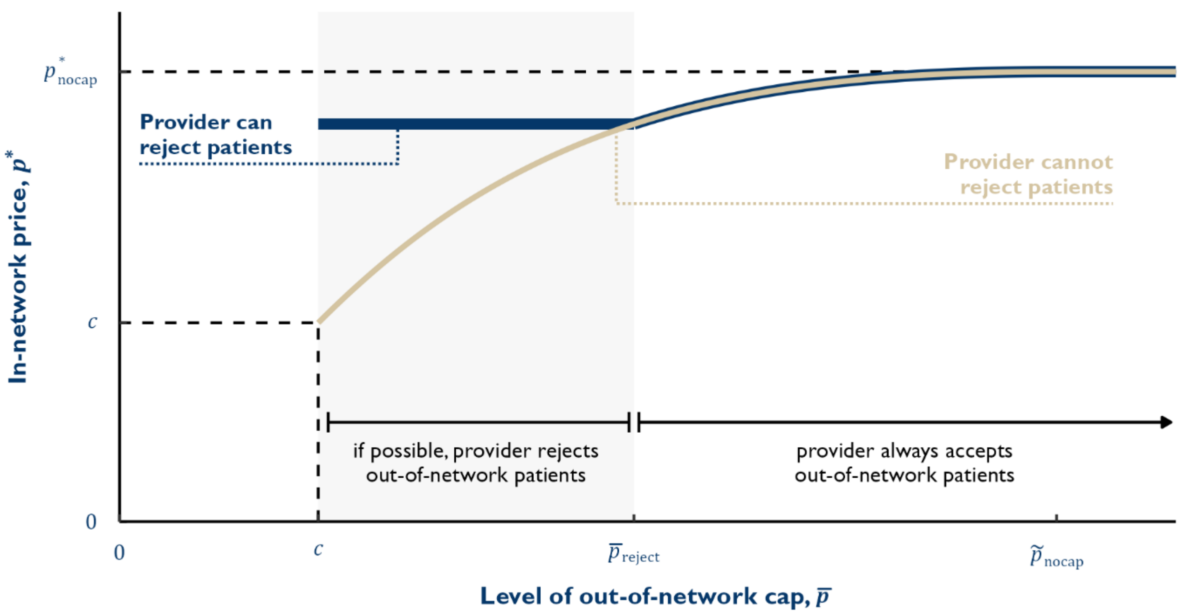
(iii) *If there is a cap and the provider can reject out-of-network patients, there is a critical level of the out-of-network cap $\bar{p}_{\text{reject}} \in (c, \tilde{p}_{\text{nocap}})$ such that:*

a. *For any $\bar{p} > \bar{p}_{\text{reject}}$, the provider does not reject out-of-network patients, and the in-network price is $p^*_{\text{accept}}(\bar{p})$*

b. *For any $\bar{p} \leq \bar{p}_{\text{reject}}$, the provider rejects out-of-network patients in some equilibria and always does when this inequality is strict. The in-network price is $p^*_{\text{accept}}(\bar{p}_{\text{reject}})$.*

The proposition says that the model has at least one equilibrium and that this equilibrium is effectively unique, in the sense that all equilibria have the same in-network price.[3] Parts (i)-(iii) of the proposition then state the properties of that (effectively) unique equilibrium for several specific scenarios. The equilibrium in-network prices are illustrated schematically in Figure 1.

Figure 1. In-Network Prices Under an Out-of-Network Cap



[3] In fact, the equilibrium is fully unique except when there is an out-of-network cap $\bar{p} \leq \bar{p}_{\text{reject}}$, and the provider has the option to turn away out-of-network patients. For $\bar{p} < \bar{p}_{\text{reject}}$, the out-of-network price $\tilde{p}$ and coverage terms $\tilde{l}$ are irrelevant to the disagreement payoffs, so there are many functionally identical equilibria. In the knife-edge case where $\bar{p} = \bar{p}_{\text{reject}}$, the provider is indifferent between rejecting and accepting out-of-network patients, so there is one equilibrium in which the provider accepts these patients and various other equilibria in which it does not.

Part (i) of the proposition says that when there is no out-of-network cap, the provider sets a high out-of-network price $\tilde{p}_{\text{nocap}}$, and the insurer responds by offering restrictive coverage that results in an out-of-network quantity below the efficient quantity; that is, $\tilde{Q}_{\text{nocap}} < Q^*$. The provider never wishes to turn away out-of-network patients because the profit it earns at this price outweighs the value the insurer derives from giving its enrollees limited access to the provider's services. Against the backdrop of these threatened out-of-network actions, the parties sign a network agreement that takes a familiar form: the provider accepts a price $p^*_{\text{nocap}} < \tilde{p}_{\text{nocap}}$, and, in exchange, the insurer offers more generous coverage for the provider's services that increases the provider's volume to the efficient quantity Q^*.

Part (ii) of the proposition (which corresponds to the gold line in Figure 1) shows that introducing an out-of-network cap can change this equilibrium markedly—at least when the provider cannot turn away out-of-network patients. Once the level of the out-of-network cap $\bar{p}$ falls below the out-of-network price without a cap $\tilde{p}_{\text{nocap}}$, the in-network price begins to gradually decline, and it does so fast enough to ensure that the in-network price is always strictly below the capped price; that is, $p^*_{\text{accept}}(\bar{p}) < \bar{p}$. In effect, the cap allows the insurer to unilaterally impose an agreement at the capped price by breaking off network negotiations, offering relatively generous out-of-network coverage for the provider's services, and directing its enrollees to see the provider on an out-of-network basis. As such, the insurer has no reason to agree to an in-network price above the capped price. In fact, the insurer is typically able to secure an in-network price below the capped price by offering the provider a negotiated agreement that features somewhat more generous coverage for the provider's services (and thus somewhat higher volume) than the provider would receive when out-of-network.

Part (iii) of the proposition (which corresponds to the dark blue line in Figure 1) shows that situation changes if the provider can turn away out-of-network patients. Now, once the cap falls below a threshold level $\bar{p}_{\text{reject}}$, the provider threatens to turn away the insurer's enrollees if they seek out-of-network care. This eliminates the insurer's ability to facilitate out-of-network access to the provider's services at the capped price and thereby limits how much leverage the insurer can derive from a cap. The result is that an out-of-network cap cannot reduce the in-network price below $p^*_{\text{accept}}(\bar{p}_{\text{reject}})$.

Even in this case, an out-of-network cap can still reduce in-network prices to some degree; that is $p^*_{\text{accept}}(\bar{p}_{\text{reject}}) < p^*_{\text{nocap}}$. This is because delivering *no* care if the parties fail to reach agreement, rather than delivering a small volume of care at a high price, results in smaller disagreement profits for the provider (without commensurately reducing the insurer's disagreement payoff). Thus, the provider still ends up in a somewhat weaker bargaining position than it held without cap. Nevertheless, the deterioration in its bargaining position (and, thus, the resulting reduction in the in-network price) may be smaller—often considerably smaller—than if it could not turn away patients.

1.4 Maximum price reduction under an out-of-network cap

The preceding discussion shows that an out-of-network cap can reduce in-network prices by at most $p^*_{\text{nocap}} - p^*_{\text{accept}}(\bar{p}_{\text{reject}})$ in settings where providers can turn away out-of-network patients. We now seek to gauge how large this reduction may be. We first derive a bound on the reduction that is a function of two "sufficient statistics." We then derive an exact expression for the reduction that holds under an additional functional form assumption and depends on an additional statistic.

To begin, we note that when the provider rejects out-of-network patients, $\tilde{\pi} = \widetilde{W} = 0$ and, thus, $\theta\tilde{\pi} - (1-\theta)\widetilde{W} = 0$. If we let $h(Q) \equiv \theta c Q + (1-\theta)V(Q)$, the weighted average of the costs the provider incurs to deliver a quantity of care Q and the value the insurer's enrollees derive from that care, equation (3) then implies that the maximum achievable price reduction satisfies

$$\left[p^*_{\text{nocap}} - p^*_{\text{accept}}(\bar{p}_{\text{reject}})\right]Q^* = \tilde{p}_{\text{nocap}}\tilde{Q}_{\text{nocap}} - h(\tilde{Q}_{\text{nocap}}). \tag{4}$$

The first term on the right-hand side of equation (4) reflects the revenue the provider can earn from the insurer's patients if it goes out-of-network; it depends solely on prices and volume, each of which can be directly measured in the claims data we use below. The second term, $h(\tilde{Q}_{\text{nocap}})$, depends on the provider's marginal cost and enrollees' valuation of the provider's services, neither of which is directly observable in our data; however, the term as a whole can be related to observable quantities.

In particular, equation (3) can be rewritten as $h(Q^*) - h(\tilde{Q}_{\text{nocap}}) = p^*_{\text{nocap}}Q^* - \tilde{p}_{\text{nocap}}\tilde{Q}_{\text{nocap}}$, which implies that the average slope of h on the interval $[\tilde{Q}_{\text{nocap}}, Q^*]$ can be estimated using the difference between the provider's in- and out-of-network revenue in a world without an out-of-network cap. Since h is strictly concave, $h(0) = 0$, and $\tilde{Q}_{\text{nocap}} \in (0, Q^*)$, it then follows that

$$h(\tilde{Q}_{\text{nocap}}) > \frac{h(Q^*) - h(\tilde{Q}_{\text{nocap}})}{Q^* - \tilde{Q}_{\text{nocap}}} \tilde{Q}_{\text{nocap}} = \left[p^*_{\text{nocap}} Q^* - \tilde{p}_{\text{nocap}} \tilde{Q}_{\text{nocap}}\right] \left[\frac{\tilde{Q}_{\text{nocap}}/Q^*}{1 - \tilde{Q}_{\text{nocap}}/Q^*}\right], \tag{5}$$

Combining equations (4) and (5) and doing some straightforward algebra then yields the following upper bound on the maximum price reduction achievable with an out-of-network cap:

$$\underbrace{\frac{p^*_{\text{nocap}} - p^*_{\text{accept}}(\bar{p}_{\text{reject}})}{p^*_{\text{nocap}}}}_{\substack{\text{maximum proportional} \\ \text{reduction in in-network price}}} < \underbrace{\left[\frac{\tilde{Q}_{\text{nocap}}/Q^*}{1 - \tilde{Q}_{\text{nocap}}/Q^*}\right]}_{\substack{\text{out-of-network} \\ \text{volume term}}} \cdot \underbrace{\left[\frac{\tilde{p}_{\text{nocap}}}{p^*_{\text{nocap}}} - 1\right]}_{\substack{\text{out-of-network} \\ \text{price term}}}. \tag{6}$$

Both terms on the right-hand side of equation (6) have a straightforward intuition behind them. The first term depends on the share of its volume the provider can retain if it goes out-of-network in a world without an out-of-network cap, $\tilde{Q}_{\text{nocap}} / Q^*$. If this share is small, then the provider can closely replicate the out-of-network outcomes that would have occurred without a cap by turning away out-of-network patients, thereby ensuring that the cap has little effect on negotiated outcomes. By contrast, if this share is large, then turning away out-of-network patients may make failing to reach agreement much less attractive to the provider than it was in a world without a cap (albeit still more attractive than if it allowed the insurer's enrollees to access its services at the capped price), resulting in a meaningfully lower in-network price than in a world without a cap.

The second term depends on the ratio between the price the provider receives for out-of-network services and the in-network price in an environment without an out-of-network cap, $\tilde{p}_{\text{nocap}} / p^*_{\text{nocap}}$. Naturally, the higher the out-of-network price the provider can secure in a world without a cap, the more being forced to turn away out-of-network patients worsens the provider's bargaining position relative to the world without a cap, and the more a cap can reduce prices.

A limitation of equation (6) is that it provides only an upper bound on the reduction in prices achievable using an out-of-network cap, not a point estimate of the maximum achievable price reduction. The tightness of this bound depends on the shape of the function h. If h is strongly concave (which will be the case if V is strongly concave), then the bound on $h(\tilde{Q}_{\text{nocap}})$ in equation (5) will be loose, so the bound on $p^*_{\text{nocap}} - p^*_{\text{accept}}(\bar{p}_{\text{reject}})$ in equation (6) will be loose as well. On the other hand, as h approaches linearity, the bounds in equations (5) and (6) become tight.

This discussion implies that obtaining an exact expression for $h(\tilde{Q}_{\text{nocap}})$ and, in turn, the maximum achievable price reduction, requires estimating the *concavity* of h. That can be done at the cost of imposing a functional form on h and requiring an additional statistic for calibration. In particular, it was noted above that $h(Q^*) - h(\tilde{Q}_{\text{nocap}})$, which corresponds to the average slope of h on $[\tilde{Q}_{\text{nocap}}, Q^*]$, is a function of observable quantities. Additionally, $h'(Q^*) = c$ since $V'(Q^*) = c$. Thus, given the provider's marginal cost c, we can estimate the local (average) concavity of h on the interval $[\tilde{Q}_{\text{nocap}}, Q^*]$. If we also assume that h is a quadratic function, we can extrapolate this local concavity estimate to obtain an exact expression for $h(\tilde{Q}_{\text{nocap}})$ as a function of observable quantities.[4]

Appendix C shows that under the assumption that h is quadratic,

$$h(\tilde{Q}_{\text{nocap}}) = \left[p^*_{\text{nocap}} Q^* - \tilde{p}_{\text{nocap}} \tilde{Q}_{\text{nocap}} + \frac{\pi^*_{\text{nocap}} - \tilde{\pi}_{\text{nocap}}}{1 - \tilde{Q}_{\text{nocap}}/Q^*}\right] \left[\frac{\tilde{Q}_{\text{nocap}}/Q^*}{1 - \tilde{Q}_{\text{nocap}}/Q^*}\right], \tag{7}$$

where $\pi^*_{\text{nocap}} \equiv Q^*(p^*_{\text{nocap}} - c)$ and $\tilde{\pi}_{\text{nocap}} \equiv \tilde{Q}_{\text{nocap}}(\tilde{p}_{\text{nocap}} - c)$. Combining equations (7) and (4) then implies that the maximum price reduction achievable with an out-of-network cap is given by:

$$\frac{p^*_{\text{nocap}} - p^*_{\text{accept}}(\bar{p}_{\text{reject}})}{p^*_{\text{nocap}}} = \left[\frac{\tilde{Q}_{\text{nocap}}/Q^*}{1 - \tilde{Q}_{\text{nocap}}/Q^*}\right] \cdot \left[\frac{\tilde{p}_{\text{nocap}}}{p^*_{\text{nocap}}} - 1 - \frac{\dfrac{\pi^*_{\text{nocap}}}{p^*_{\text{nocap}} Q^*} - \dfrac{\tilde{\pi}_{\text{nocap}}}{p^*_{\text{nocap}} Q^*}}{1 - \tilde{Q}_{\text{nocap}}/Q^*}\right]. \tag{8}$$

The right-hand side of equation (8) is similar to the right-hand side of equation (6), except for the addition of a new term that depends on the difference between in- and out-of-network profits. Because $\frac{\pi^*_{\text{nocap}}}{p^*_{\text{nocap}} Q^*} = 1 - \frac{c}{p^*_{\text{nocap}}}$ and $\frac{\tilde{\pi}_{\text{nocap}}}{p^*_{\text{nocap}} Q^*} = \left[\frac{\tilde{Q}_{\text{nocap}}}{Q^*}\right]\left[\frac{\tilde{p}_{\text{nocap}}}{p^*_{\text{nocap}}} - \frac{c}{p^*_{\text{nocap}}}\right]$, calibrating these two terms requires knowing c / p^*_{nocap}, the ratio of the provider's marginal cost to the in-network price without a cap, which did not appear in equation (6). Because providers' costs are not observable in the claims data we use here, we calibrate this ratio using estimates from the literature.

[4] Technically, we assume only that h is quadratic on the interval $[0, Q^*]$ since if h were globally quadratic then that would contradict our maintained assumption that V is globally strictly increasing. Since we are exclusively interested in the interval $[0, Q^*]$, this distinction is irrelevant for our purposes, so we ignore it in what follows to avoid tedious qualifiers.

2 Empirical Methods

We now turn to estimating the statistics needed to calibrate equations (6) and (8). To do so, we examine how hospital volume and prices change during episodes in which a provider transitions from being inside an insurer's network to outside the insurer's network, or vice versa. Our main focus is on trends in volume and prices of non-emergency care since these are the situations in which providers have the option to turn away patients, but we also report results for emergency care. We are aware of one other paper that studies similar transitions: Melnick and Fonkych (2020a), who examine an episode in which a single hospital chain left all private insurers' networks.[5]

We note that an alternative approach would be to use cross-sectional (or other) variation to estimate a demand system that predicts hospital volume as a function of network status and other characteristics, which is the approach taken by Prager and Tilipman (2020) in their analysis of regulating out-of-network prices. If paired with comparable estimates of prices as a function of network status, these estimates could be used to calibrate equations (6) and (8). Our approach has the virtue of being simple and transparent, but this alternative approach has advantages as well, as it can potentially provide estimates applicable to all hospitals (or even specific hospitals), whereas our estimates only directly apply to the sample of hospitals that we observe changing network status.

2.1 Data

We use claims data from the Health Care Cost Institute (HCCI) that encompass the universe of claims from three large national insurers (Aetna, Humana, and United Healthcare) for 2014-2017. Each claims record contains various information about the relevant encounter, including: diagnosis, procedure, and revenue codes; an encrypted enrollee identifier; an encrypted version of the provider's national provider identifier (NPI); the provider's billed charge; the plan's allowed amount; an indicator for whether the provider was in the plan's network when the service was delivered; and an indicator for whether the plan was the primary payer for that encounter. The data also include a file that reports

[5] The network status transitions studied by Melnick and Fonkych occurred during the 2005-2007 period. The transitions we study all occurred during 2015 or 2016, so there is no overlap between our samples.

information about each enrollee, including: the market segment of the enrollee's plan (e.g., group, nongroup, or Medicare Advantage); an encrypted group identifier for the employer offering the plan (for group plans); and the enrollee's age, months of enrollment, and zip code of residence. Throughout, we limit our analysis to employer plans, exclude enrollees ages 65 and older, and exclude claims where the plan is functioning as a secondary payer.

Our empirical strategy requires us to group together claims associated with a particular hospital as well as claims associated with a particular employer plan. To group claims at the hospital level, we use a crosswalk produced by HCCI that maps all facility NPIs associated with a given hospital to a single "consolidated" NPI (HCCI 2020). To group claims at the plan level, we use the encrypted group identifier reported for each enrollee. There may be cases where a single group identifier encompasses multiple plans offered by the same employer. For our purposes, this will be a problem only if those plans offer different networks and are available in the same geographic areas, which is likely relatively rare. Moreover, even when this does occur, the likely consequence is that we would fail to identify some instances where a hospital changed network status and thus lose the ability to analyze those particular episodes, which would not pose an obvious risk of bias. Thus, in what follows, we abstract from this complication and describe each group identifier as representing a single "plan."

2.2 Identifying network status transitions

We seek to identify episodes where a hospital enters or leaves the network of one or more plans. We do not directly observe plan networks, so we instead use the following algorithm to identify episodes empirically using the network status indicator reported on each claim.

Step 1: Identify hospital-plan pairs with adequate expected volume

To begin, we limit our attention to hospital-plan pairs that are expected to account for at least 10 emergency department visits in each year covered by our data given the distribution of the plan's enrollment across zip codes and average hospital utilization patterns in each zip code. To calculate expected volume, we first use the enrollment file to tabulate the number of person-years of enrollment accounted for by plan p in zip code z in year y, which we denote n_{pzy}. We then use the claims file to calculate the number of emergency department visits to hospital h by people in zip code z in year y, which we denote v_{hzy}.[6] Our measure of expected volume in year y for the hospital-plan pair consisting of hospital h and plan p is then simply $e_{hpy} = \sum_z n_{pzy} \cdot [v_{hzy}/N_{zy}]$, where $N_{zy} = \sum_p n_{pzy}$.

The purpose of this expected volume restriction is to limit our sample to hospital-plan pairs where we are likely to observe enough emergency department visits to obtain an accurate picture of how the hospital's network status under a plan evolves over time. We use emergency department volume (rather than total volume) for this purpose because the flow of these visits is likely less affected by network status (something that is borne out in our results) and because our primary interest is in changes in volume, revenue, and prices for *non-emergency* visits. Thus, using emergency department volume minimizes the risk that hospital-plan pairs will be selected into our initial sample based on how much non-emergency volume the hospital actually retains when out-of-network.[7]

Step 2: Identify candidate transition episodes

We next identify hospital-plan pairs that may have experienced a suitable network status transition by limiting the sample to hospital-plan pairs with at least two quarters where the majority of emergency department visits are out-of-network and at least two quarters where the majority of emergency department visits are in-network.

Step 3: Combine hospital-plan pairs experiencing the same transition event into a single episode

In some cases, we find that more than one hospital-plan pair associated with a particular hospital appears to have experienced a network status transition, which may often reflect instances where an insurer has a contract with the hospital that applies across multiple employer plans. In these cases, we apply a chi-squared test of whether all of the hospital-plan pairs associated with the hospital exhibit the same mix of network statuses by quarter. If this test generates a p-value greater than or equal to

[6] We identify emergency department claims as claims that HCCI classifies as outpatient facility claims that report a place of service code of 22, 23, or missing and either: (1) a Current Procedural Terminology code of 99281-99285 or 99291-99292; or (2) a revenue code of 0450-0459 or 0981. This definition is similar to definitions used in prior work (e.g., Venkatesh et al. 2017). We use a visit identifier populated by HCCI to group all claims associated with an encounter into a single "visit."

[7] While they are not our main focus, we do present estimates of how emergency volume, revenue, and prices change around network status transitions. In principle, these estimates could also be distorted by the type of sample selection bias we are discussing here. In practice, the summary statistics presented in Table 1 suggest that our episodes typically have ample volume when both in- and out-of-network, suggesting that few episodes are near the margin of sample inclusion.

0.05, then we consolidate the hospital-plan pairs into a single transition episode.[8] Consolidating pairs in this way reduces the number of hospital-plan-quarter tuples for which we do not observe a claim and, thus, cannot ascertain the hospital's network status under the plan.

Step 4: Assign quarterly network status and select episodes of interest

For each potential transition episode, we determine the network status of the hospital under the relevant plan (or plans) based on the network status of the emergency department visits observed for that quarter. Each quarter's assigned network status can be either in-network, out-of-network, mixed (which may occur if a network status transition occurs in the middle of a quarter), or missing (which may occur if no claims are observed for a quarter). We then select only the episodes for which we observe the hospital to have had a single network status (either in- or out-of-network) for at least four quarters and then the opposite network status for the subsequent six quarters (or, alternatively, mixed network status for next quarter and the opposite status for the subsequent five quarters).

2.3 Analytic dataset

After identifying the transition episodes, we construct an analytic dataset that consists of ten quarterly entries for each transition episode: four pre-transition entries and six follow-up entries. For each quarter, we compute several aggregate amounts across all plans associated with the episode, including the aggregate number of (combined) inpatient and outpatient visits to the hospital in question by enrollees in the relevant plan (or plans), the aggregate facility charges associated with those visits, and the aggregate facility allowed amounts associated with those visits.[9] We calculate each of these aggregates separately for emergency and non-emergency encounters.[10]

---

[8] In simulations, we found that this test had high power to detect instances where two hospital-plan pairs had different network status transition patterns, so we are unlikely to inappropriately consolidate hospital-plan pairs into a single episode.

[9] We are unable to report results for inpatient and outpatient visits separately because we often observe very few non-emergency inpatient visits when hospitals are out-of-network, which caused some of our regression coefficients to correspond to cells with fewer than 11 claims, the minimum cell size for which HCCI permits researchers to export results.

[10] We classify an outpatient visit as emergent if the underlying claim (or claims) meet the criteria we used to make our initial tallies of emergency department visits. We classify an inpatient visit as emergent if the underlying claim (or claims) reports: (1) a type of admission code of "emergency" or "trauma center"; or (2) one of the revenue codes (or, rarely, CPT codes) we use to identify emergency outpatient visits. Once again, we use a field pre-populated by HCCI to consolidate all claims associated with an encounter into a single "visit."

Table 1. Summary Statistics

Transition sample:	In-Network to Out-of-Network		Out-of-Network to In-Network	
	Pre	Post	Pre	Post
Mean quarterly emergency volume				
Visits	21.6	17.9	16.4	17.8
Charges (1000s of $)	139.8	135.8	122.3	158.1
Allowed amounts (1000s of $)	46.6	82.8	81.6	55.8
Mean quarterly non-emergency volume				
Visits	42.0	12.0	2.9	18.4
Charges (1000s of $)	271.2	70.1	19.6	129.1
Allowed amounts (1000s of $)	162.6	53.9	14.5	41.4
Number of observations				
Episode-quarters	24	36	80	120
Episodes	6	6	20	20

Table 1 presents descriptive statistics for our final episode-by-quarter analytic sample. The sample contains a total of 26 episodes in which a hospital changed its network status under one or more plans. Of those episodes, 6 involve a hospital transitioning from in-network to out-of-network, while 20 involve a hospital transitioning from out-of-network to in-network.

There are likely multiple reasons that we identify relatively few transition episodes. First, limiting our sample to hospital-plan pairs with at least 10 expected emergency department visits annually has the effect of limiting our attention to pairs where the plan is of at least moderate size *and* has meaningful enrollment in the hospital's catchment area. Since larger employers tend to prefer plans with very broad networks (KFF 2020), there may be relatively few out-of-network hospitals in these plans that could transition to being in-network, and insurers may be loath to allow in-network hospitals to transition out-of-network (especially for long enough to satisfy our inclusion criteria). Second, because we hold data for 2014 through 2017 and must observe at least four pre-transition quarters and six follow-up quarters means, our final sample can only include transitions that occurred during 2015 or the first three quarters of 2016, a relatively short time period.

2.4 Regression specification

We use the resulting dataset to run a series of (quasi-)Poisson event study regressions of the form:

$$\mathrm{E}[y_{eq}] = \exp(\gamma_e + \delta_q), \tag{9}$$

where e indexes transition episodes, $q \in \{-4..5\}$ indexes quarters relative to the transition quarter, y_{eq} is the outcome of interest, $\{\gamma_e\}$ are a set of episode fixed effects, and $\{\delta_q\}$ are a set of time-to-transition effects. We estimate these regressions by conditional maximum likelihood and cluster standard errors at the episode level. In what follows, the parameters of interest are generally the exponentiated coefficients $\exp \delta_q$ (or combinations thereof), which reflect proportional differences across periods. We construct standard errors and confidence intervals for these parameters using the delta method.

We estimate equation (9) separately for the samples of in-network to out-of-network transition episodes and out-of-network to in-network transition episodes. We normalize δ_{-4} to zero (equivalently, $\exp \delta_{-4} = 1$) when analyzing in-network to out-of-network transitions and normalize δ_5 to zero (equivalently, $\exp \delta_5 = 1$) when analyzing out-of-network to in-network transitions. This ensures that the base quarter is always an in-network quarter, which facilitates graphical display.

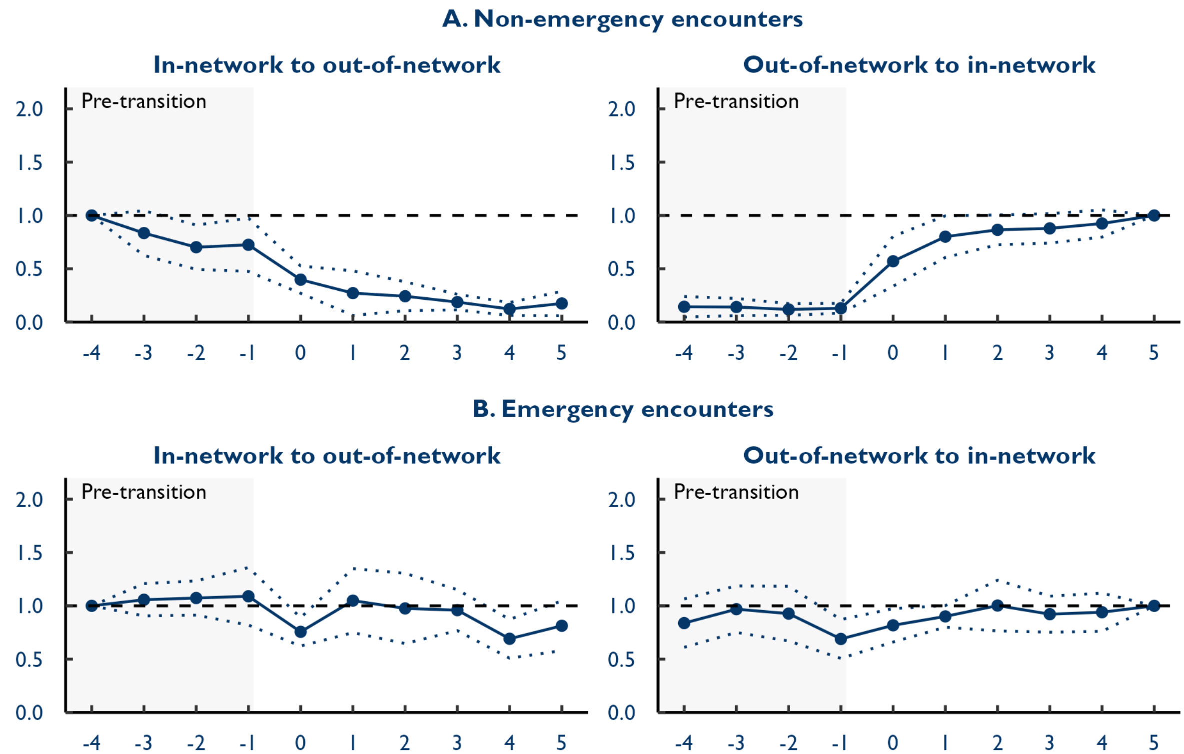
Our ultimate goal is to estimate the ratio of (steady state) out-of-network outcomes to (steady state) in-network outcomes. To do so, we make the identifying assumptions that post-transition outcomes: (a) would have matched the pre-transition average without the change in network status; and (b) remained at the level observed in the final post-transition quarter in later quarters. The estimand of interest can then be expressed as $\exp \delta_5 \,/\, \left[(1/4) \sum_{q \in \{-4..-1\}} \exp \delta_q\right]$ for in-network to out-of-network transitions and as the reciprocal, $\left[(1/4) \sum_{q \in \{-4..-1\}} \exp \delta_q\right] / \exp \delta_5$, for out-of-network to in-network transitions. We also report weighted average estimates in which the estimates obtained for each type of transition episode are weighted by the number of episodes of that type.

3 Empirical Results

Figure 2 reports the estimates of how hospital visit volume changes around a change in network status that are obtained from equation (9); formally, each point corresponds to the relevant estimate of $\exp \delta_q$. Panel A shows that the number of non-emergency visits changes sharply around a change in network status. Visit volume falls sharply when an in-network hospital moves out-of-network and rises sharply when an out-of-network hospital moves in-network. By contrast, Panel B shows that there is

Figure 2. Number of Visits Around Network Transitions

Number of visits relative to base quarter



Note: Figure plots estimates of $\exp \delta_q$ from equation (9). Dotted lines depict 95% confidence intervals. Transition quarter is q = 0. Base quarter is q = -4 for in-network to out-of-network episodes and q = 5 for out-of-network to in-network episodes.

much less, if any, change in emergency volume around a network status transition, plausibly because patients seeking emergency care often just choose a nearby hospital (e.g., Brown, Decker, and Selck 2015) or are steered to hospitals by factors beyond their control, like ambulance company preferences (e.g., Doyle, Graves, and Gruber 2019). Panels A.1 and B.1 of Table 2 present corresponding summary estimates of how out-of-network volume compares to in-network volume. When averaging across the two transition directions, we estimate that out-of-network non-emergency volume is 15.2% of in-network volume, while out-of-network emergency volume is 83.6% of in-network volume.

The observed volume trends offer some support for our identifying assumptions. The relative stability in non-emergency volume prior to the network status transition and the sharp changes immediately thereafter suggests the observed volume changes are not driven by secular trends in plan

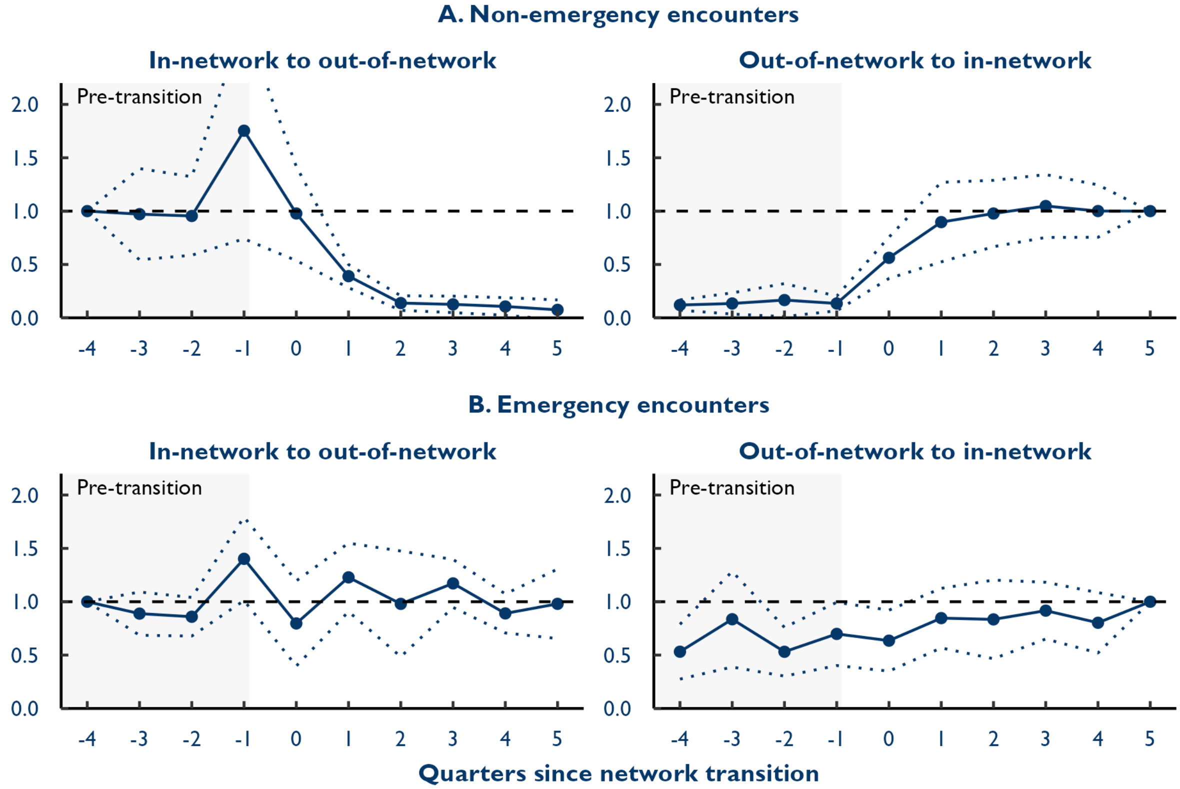
Table 2. Out-of-Network Volume, Revenue, and Prices

Transition type:	In-Network to Out-of-Network	Out-of-Network to In-Network	Weighted Average
A. Non-emergency encounters			
A.1 Volume (% of in-network volume)			
Visits	21.4	13.4	15.2
	(8.8)	(3.2)	(3.2)
Charges	6.4	13.9	12.2
	(5.1)	(3.0)	(2.6)
A.2 Revenue (% of in-network revenue)			
Collect allowed amount only	4.6	31.7	25.5
	(2.6)	(7.6)	(5.9)
Collect full charge	10.6	42.9	35.4
	(9.0)	(8.7)	(7.0)
Collect allowed amount + 30% of balance bill	6.4	35.1	28.5
	(4.6)	(7.9)	(6.2)
A.3 Prices (% of in-network prices)			
Collect allowed amount only	72.6	228.4	209.5
Collect full charge	165.5	308.9	291.5
Collect allowed amount + 30% of balance bill	100.7	252.6	234.1
B. Emergency encounters			
B.1 Volume (% of in-network volume)			
Visits	77.1	85.6	83.6
	(10.5)	(8.8)	(7.2)
Charges	94.5	64.9	71.8
	(17.7)	(13.6)	(11.2)
B.2 Revenue (% of in-network revenue)			
Collect allowed amount only	177.6	107.8	123.9
	(32.2)	(29.9)	(24.1)
Collect full charge	281.3	159.8	187.9
	(55.8)	(45.2)	(37.0)
Collect allowed amount + 30% of balance bill	208.9	123.5	143.2
	(38.0)	(34.1)	(27.7)
B.3 Prices (% of in-network prices)			
Collect allowed amount only	188.0	166.1	172.7
Collect full charge	297.7	246.2	261.8
Collect allowed amount + 30% of balance bill	221.1	190.2	199.6

Note: For in-network to out-of-network transitions, the volume and revenue point estimates are calculated as $\exp \delta_5 / \left[(1/4)\sum_{q\in\{-4..-1\}} \exp \delta_q\right]$ using estimates from the relevant version of equation (9), while estimates for out-of-network to in-network transitions are the reciprocal. Weighted average estimates weight each transition type by the number of episodes of that type. Standard errors are calculated via the delta method from clustered standard errors of the underlying parameters. Price estimates are derived by dividing the relevant revenue estimate by the relevant charge-based volume estimate.

Figure 3. Aggregate Charges Around Network Transitions

Aggregate charges relative to base quarter



Note: Figure plots estimates of $\exp \delta_q$ from equation (9). Dotted lines depict 95% confidence intervals. Transition quarter is q = 0. Base quarter is q = -4 for in-network to out-of-network episodes and q = 5 for out-of-network to in-network episodes.

enrollment or patient demand for the transitioning hospitals. The relative stability in emergency volume over the entire period also implies that network status transitions do not typically coincide with large changes in the number of plan enrollees who live near the transitioning hospital. It also offers some additional evidence that patient demand for different hospitals is relatively stable over time, although shocks to patient demand for particular hospitals might have a muted effect on emergency volume for the same reasons that network status has little effect on those volumes.

A downside of using visit counts to assess volume trends is that they do not account for variation in the resource intensity of different visits. To address this concern, Figure 3 presents trends in aggregate provider charges around network status transitions. While provider charges are an admittedly crude measure of intensity, they are available directly from the claims record and are likely computed comparably for both in- and out-of-network visits.

The trends in aggregate charges depicted in Figure 3 are qualitatively similar to the trends in raw visit volume depicted in Figure 2, albeit a bit noisier. As with raw visits, charge-weighted non-emergency volume falls sharply when a hospital moves out-of-network and rises sharply after a transition in the opposite direction. Once again, however, emergency volume changes much less, if at all, surrounding changes in network status. Panels A.1 and B.1 of Table 2 report the corresponding summary estimates. When averaging across the two transition directions, out-of-network charge-weighted non-emergency volume is 12.2% of in-network volume, while out-of-network charge-weighted emergency volume is 71.8% of in-network volume.

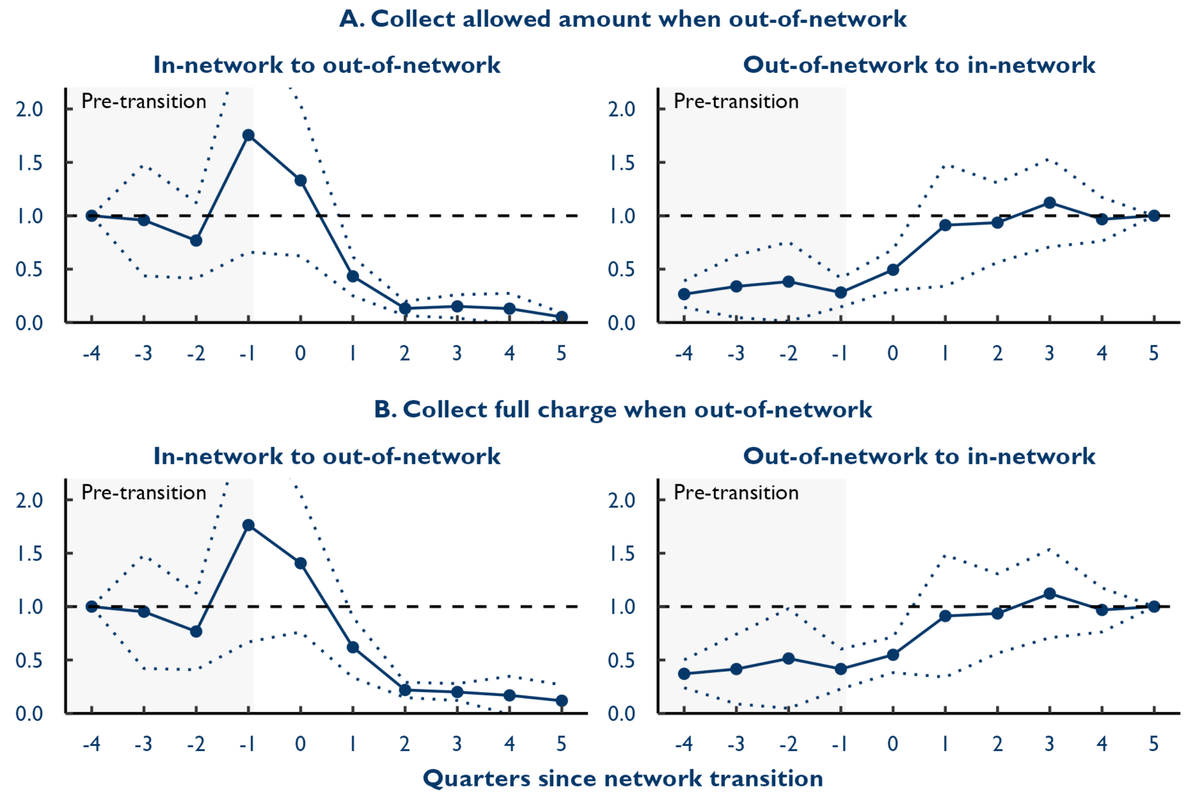
We next examine trends in hospital revenues and prices around network transitions. A challenge in analyzing revenue trends is that we observe only the plan's allowed amount and the hospital's charge, not how much the hospital actually succeeds in collecting. In particular, we do not observe whether the hospital successfully collects the portion of the allowed amount that is due from the patient as cost-sharing. Additionally, for out-of-network care, providers can bill the patient for the amount that the provider's charge exceeds the plan's allowed amount, but we do not observe how much hospitals collect from these "balance bills."[11] Since hospitals' ability to collect from patients is imperfect (Ippolito and Vabson 2023), these omissions create uncertainty about how much hospitals ultimately collect, especially for out-of-network care that tends to carry larger cost-sharing obligations and can precipitate balance bills (Biener et al. 2021; Song et al. 2020; Pelech 2020).

We proceed by considering three scenarios for provider collections. The three scenarios all assume that the hospital collects the plan's full allowed amount when in-network but incorporate different assumptions about out-of-network collections. The first scenario assumes that hospitals collect only the plan's allowed amount when out-of-network, while the second assumes that hospitals collect their full charge when out-of-network. Neither of these extreme scenarios is likely realistic, so we also consider a third scenario, which is our preferred scenario in what follows. This scenario assumes that hospitals collect the allowed amount plus 30% of the difference between the hospital's charge and the

[11] During our study period, some states had banned balance billing in certain circumstances, but those laws generally did not apply to facilities. As of January 2022, federal law bars facilities from balance billing for emergency care.

Figure 4. Revenues for Non-Emergency Care Around Network Transitions

Expected collections relative to base quarter



Note: Figure plots estimates of $\exp \delta_q$ from equation (9). Dotted lines depict 95% confidence intervals. Transition quarter is q = 0. Base quarter is q = -4 for in-network to out-of-network episodes and q = 5 for out-of-network to in-network episodes.

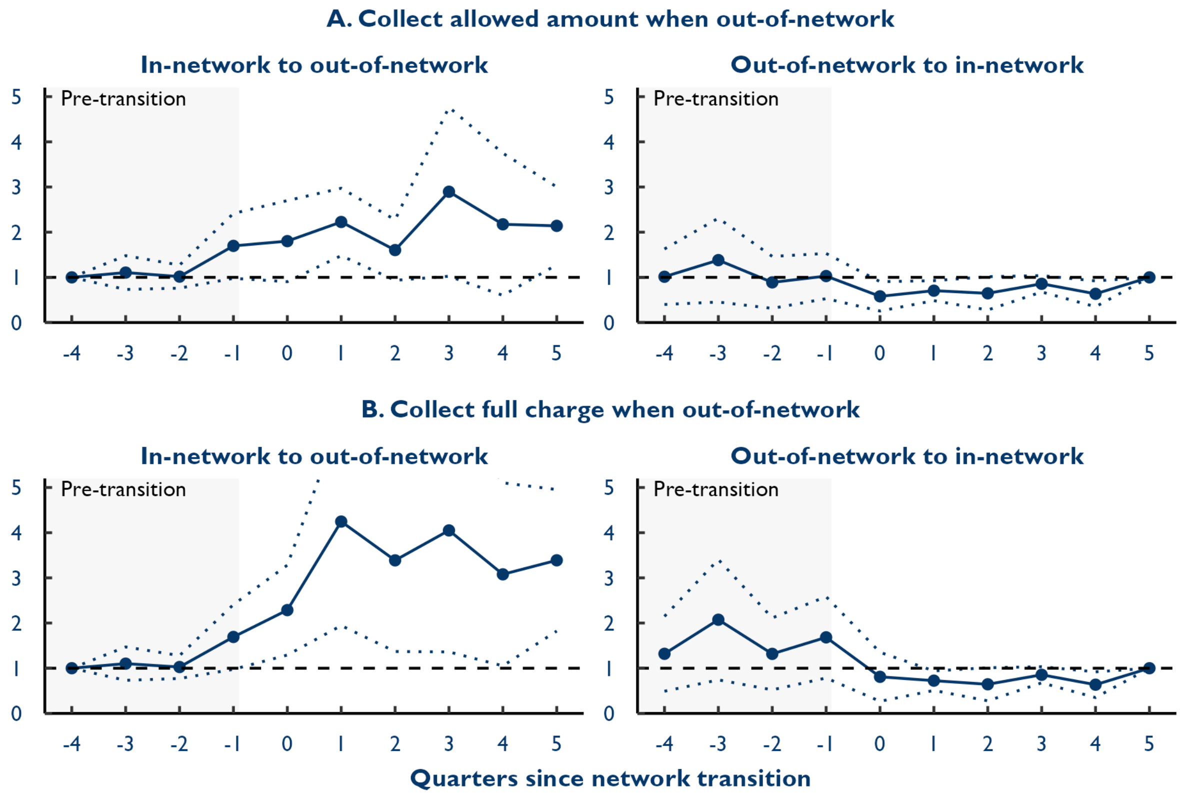
plan's allowed amount (that is, 30% of the potential balance bill); this aligns with evidence on how much physicians collect for out-of-network emergency services (Biener et al. 2021).[12]

Figure 4 reports estimated trends in revenue for non-emergency care around network transitions under the two extreme scenarios for providers' out-of-network collections. Like volumes, revenues for non-emergency care are also lower when out-of-network. However, the revenue declines depicted in Figure 4 are smaller than the volume declines depicted in Figures 2 and 3, at least in the sample of out-of-network to in-network transitions. Panel A.2 of Table 2 shows that, in our preferred scenario where hospitals collect the allowed amount plus 30% of the potential balance bill, the weighted average

[12] For emergency visits occurring in 2011 through 2016 involving potential balance bills, Biener et al. report an average charge of $789, an average plan payment of $168, and an average enrollee payment of $219. If the patient's cost-sharing obligation averages 20% of the allowed amount, then the average cost-sharing payment is $42 (=$168*[0.2/0.8]), which together with the above implies an average balance bill of $579 (=$789 - $168 - $42) and an average balance bill collection of $177 (=$219 - $42). Thus, the total average collection is the allowed amount plus 31% (=$177/$579).

Figure 5. Revenues for Emergency Care Around Network Transitions

Expected collections relative to base quarter



Note: Figure plots estimates of $\exp \delta_q$ from equation (9). Dotted lines depict 95% confidence intervals. Transition quarter is q = 0. Base quarter is q = -4 for in-network to out-of-network episodes and q = 5 for out-of-network to in-network episodes.

estimate is that out-of-network non-emergency revenue is 28.5% of in-network revenue, whereas out-of-network non-emergency volume is 12.2% of in-network volume (using the charge-based measure).

This pattern indicates that hospitals collect notably higher prices for out-of-network care than they collect for in-network care, at least under our preferred assumptions about out-of-network collections. To formalize this, Panel A.3 of Table 2 calculates estimates of weighted average prices by dividing the revenue estimates from Panel A.2 by the charge-based volume estimate from Panel A.1. Under our preferred assumption that hospitals collect 30% of potential balance bills, the weighted average estimate is that the out-of-network price for non-emergency services is 234.1% of the in-network price.

In closing, we note that revenue trends differ markedly for emergency care. Figure 5 demonstrates that under the full range of assumptions we consider, expected collections for emergency care are actually *higher* when a hospital is out-of-network than when it is in-network. Indeed, Panel B.2 of

Table 2 shows that, in our preferred scenario, the weighted average estimate is that emergency out-of-network revenue is 143.2% of in-network revenue. The difference relative to non-emergency care primarily reflects the fact emergency volume is relatively insensitive to provider network status, as our estimates of the ratio of out-of-network prices to in-network prices are similar for emergency and non-emergency care.

4 Model Calibration

We now use our empirical results to calibrate the upper bound on how much an out-of-network cap can reduce in-network prices that appears in equation (6), as well as the exact expression for the maximum achievable reduction in in-network prices that appears in equation (8).

In doing so, we rely on the estimates in Table 2 that average across the two transition directions. For volume, our preferred estimate (which reflects our charge-based volume measure) is that hospitals that shift out-of-network retain 12.2% of their in-network non-emergency volume; thus, we set $\tilde{Q}_{\text{nocap}}/Q^* = 0.122$. For prices, our preferred estimate (which assumes that hospitals collect the plan's allowed amount plus 30% of the potential balance bill when out-of-network) is that the prices hospitals collect for out-of-network non-emergency care are 234% of in-network prices; thus, we set $\tilde{p}_{\text{nocap}}/p^*_{\text{nocap}} = 2.34$. We note that these parameter estimates are broadly similar to the corresponding estimates reported by Melnick and Fonkych (2020a) in their study of a single hospital chain. The hospitals they studied appear to have retained around 8% of their in-network non-emergency volume when they shifted out-of-network and also received markedly higher out-of-network prices.[13]

Calibrating the exact expression for the maximum achievable reduction in in-network prices that appears in equation (8) also requires an estimate of the ratio of the provider's marginal cost to the in-network price, $c \,/\, p^*_{\text{nocap}}$. We do not observe hospital costs in our data, so we combine an estimate from the Medicare Payment Advisory Commission (2023) that hospitals' marginal cost of delivering care averaged 92% of Medicare's prices in 2021 with an estimate from a Congressional Budget Office (2022)

[13] See footnote 28 of Fiedler (2020) for how to derive this 8% estimate from Melnick and Fonkych's estimates.

review that the prices commercial insurers pay for hospital services (including both inpatient and outpatient services) average 223% of Medicare's prices. We obtain $c/p^*_{\text{nocap}} = 0.412$

Using these estimates to calibrate the bound in equation (6), we find that an out-of-network cap can reduce in-network prices for non-emergency hospital services by at most 19%. This estimate is an *upper bound* on the potential effect of an out-of-network cap in this setting, so the actually achievable price reductions may be smaller. Indeed, our point estimate from calibrating the exact expression for this maximum achievable reduction presented in equation (8) is only 13%, although we note that this estimate relies on an assumption that h is quadratic, whereas the bound holds more generally. Because non-emergency care accounts for a large majority of health care spending, including around two-thirds of hospital spending (Fiedler 2020), these results suggest that there are meaningful limits on how much an out-of-network cap can reduce in-network prices, contrary to what an analysis that ignored providers' ability to turn away out-of-network patients would suggest.

There are a several caveats to these conclusions worth noting. First, our results would look at least somewhat different if we focused exclusively on the sample of in-network to out-of-network transitions or the sample of out-of-network to in-network transitions, rather than weighted averages across the samples. While both samples exhibit broadly similar volume trends, the price trends differ markedly. For the in-network to out-of-network transitions, our preferred estimate of $\tilde{p}_{\text{nocap}}/p^*_{\text{nocap}}$ is close to one, which implies the that upper bound in equation (6) is close to zero. For out-of-network to in-network transitions, on the other hand, our estimates of $\tilde{p}_{\text{nocap}} \,/\, p^*_{\text{nocap}}$ are higher; using those estimates implies an upper bound from equation (6) of 25% and a point estimate from equation (8) of 19%, higher than the bound and point estimate we obtain when using the weighted average estimates. The reason for the differences between the two samples is not clear, but one possible explanation is that they reflect idiosyncratic differences between the plan-hospital pairs that happen to be included in the two samples. If this is the case, it would suggest that the scope for an out-of-network cap to reduce in-network prices may vary across plans and providers.

Second, our model does not allow failure to reach a network agreement to affect the parties' payoffs except through its effects on prices and volume. However, as shown formally in Appendix D, if being out-of-network imposes additional costs on hospitals, then our approach will overstate how much a hospital's decision to give up its out-of-network volume will worsen its bargaining position relative to a world without a cap and, correspondingly, overstate the scope for an out-of-network cap to reduce in-network prices. The reverse will be true if insurers bear additional costs when out-of-network.

In practice, there is reason to suspect that collecting payment is more costly for hospitals when they are out-of-network. A much larger share of the revenue hospitals receive for out-of-network care must be collected directly from patients since patient cost-sharing obligations typically represent a larger fraction of the plan's allowed amount and "balance bills" are, by definition, collected from patients (Biener et al. 2021; Pelech 2020; Song et al. 2020). In practice, collecting large amounts directly from patients may involve larger administrative and reputational costs (Cooper, Scott Morton, and Shekita 2020). Thus, these considerations offer some reason to believe that our results overstate the scope for an out-of-network cap to reduce in-network prices.[14]

Third, it is notable that we find that the revenues and prices hospitals receive for *emergency* services actually rise when a hospital shifts out-of-network, a finding that again echoes Melnick and Fonkych (2020a). As noted above, collecting from out-of-network patients may be burdensome for hospitals, so being out-of-network may not be as attractive with respect to emergency services as our results suggest. Nevertheless, these findings raise the possibility that some of the pricing power hospitals hold is "shifted" into the in-network prices of non-emergency services rather than being fully reflected in the prices of emergency services (Pope 2019; Melnick and Fonkych 2020a). This could be because provider-insurer contracts often specify prices as a multiple of some other fee schedule, like a Medicare fee schedule or the hospital's chargemaster (Clemens and Gottlieb 2016; Cooper et al. 2019), rather than separately specifying prices for all relevant services.

If some of the pricing power hospitals derive from their emergency services is indeed reflected in the prices negotiated for non-emergency services, then the in-network prices for non-emergency services that we observe on claims are higher than the "true" prices of those services; similarly, the in-

[14] If the out-of-network collection process imposes additional non-financial costs on *patients* (e.g., disutility from negotiating a balance bill, being exposed to unpleasant collection tactics, or failing to pay) and insurers take those non-financial costs in account when assessing the value of reaching a network agreement, then that could work in the opposite direction.

network prices that we observe for emergency services are lower than the "true" prices of those services. That would suggest that our estimate of the ratio $\tilde{p}_{\text{nocap}}$ / p^*_{nocap} is biased downward and, thus, that our results understate the scope for an out-of-network cap to reduce in-network prices. It would also suggest that the "true" share of hospital spending attributable to emergency care is higher than it appears, which would imply in turn that more spending happens in settings where out-of-network caps can be highly effective in reducing in-network prices.[15]

Finally, our analysis does not account for the possibility that hospitals that turn away out-of-network patients incur reputational costs by doing so. If this is the case, then our approach will overstate how easy it is for hospitals to turn away out-of-network patients and, correspondingly, understate the scope for an out-of-network cap to reduce in-network prices. Importantly, however, this is only the case for reputational costs that are directly tied to turning way out-of-network patients, rather than just to going out-of-network. (As noted above, if hospitals incur reputational costs by going out-of-network, then our approach will actually tend to *overstate* the scope for an out-of-network cap to reduce prices.) Our finding that a hospital that goes out-of-network loses the large majority of its non-emergency volume suggests that there may be little scope for turning away out-of-network patients to impose substantial additional reputational costs.

5 Conclusion

We use a model of provider-insurer bargaining adapted from Fiedler (2020) to show that the amount of leverage an insurer derives from an out-of-network cap hinges on whether the provider is able to turn away out-of-network patients. We then use the model to derive an upper bound on the amount an out-of-network cap can reduce in-network prices in settings where providers can turn away patients; this bound depends on how much volume hospitals retain when they shift out-of-network—and at what price. By examining episodes where hospitals change network status, we estimate that

[15] Notably, this view would suggest that the No Surprises Act—a recent federal law that limits prices for out-of-network emergency services using an arbitration system in which arbitrators are directed to base their decisions in part on each insurer's median in-network price before enactment—is more stringent than it appears and, thus, has correspondingly greater potential to reduce in-network prices for hospital services.

hospitals that shift out-of-network retain only 12% of their non-emergency in-network volume, albeit at a price more than double what they receive when in-network. When we use these estimates to calibrate the model-derived bound, we conclude that an out-of-network cap can reduce in-network prices for non-emergency hospital services by at most 19%. Because non-emergency care accounts for a large majority of all health care spending, our results suggest that there are meaningful limits on how much an out-of-network cap can reduce the prices of health care services, contrary to what an analysis that ignored providers' ability to turn away out-of-network patients would suggest.

Our findings have a couple of policy implications. First, our results suggest that while there is some scope to reduce the prices that commercial plans pay for in-network services by capping out-of-network prices, policymakers wishing to greatly reduce in-network prices may need to consider other approaches. These could include regulating both in- and out-of-network prices or combining an out-of-network cap with limits on providers' ability to turn away out-of-network patients in non-emergency settings (Fiedler 2020). These alternative approaches may also be less likely to reduce access to out-of-network services; regulating in-network prices does not create the same incentives to turn away out-of-network patients, and the latter policy directly prevents them from doing so.

Second, our estimates offer insights into the forces that determine provider prices in the Medicare Advantage (MA) program. A striking feature of the MA market is that the prices MA plans negotiate for hospital and physician services are close to traditional Medicare's (e.g., Berenson et al. 2015; Maeda and Nelson 2018; Pelech 2020), even as commercial plans pay much higher prices, including around twice as much for inpatient services and even more for outpatient facility services (e.g., Blumberg et al. 2020; Chernew, Hicks, and Shah 2020; Cooper et al. 2019; Maeda and Nelson 2018; Whaley et al. 2020). Some prior work (e.g., Berenson et al. 2015; Maeda and Nelson 2018; Pelech 2020) has suggested that the MA out-of-network cap is a major driver of the lower prices observed in MA.

Our bound on the amount an out-of-network cap can reduce in-network prices implies that the MA out-of-network cap is likely not the primary reason that MA plans pay providers so much less than commercial plans.[16] Rather, other aspects of the MA landscape likely play the central role in disciplining provider prices in MA. One particularly important factor may be MA plans' need to remain competitive with traditional Medicare, which may allow plans to credibly refuse to pay prices much above traditional Medicare's (Berenson et al. 2015; Maeda and Nelson 2018; Fiedler 2020; Pelech 2020). However, other factors could also play a role. For example, the fact that plan choices typically occur at the individual level in MA, but at the employer level in commercial plans, could reduce how much plans are willing to pay to lure additional providers into their plans by magnifying risk selection pressures (Shepard 2022) or reducing the weight assigned to consumers who place a high value on having broad provider networks (Tilipman 2022).

If competition from traditional Medicare is playing an important role in disciplining provider prices in MA, then this has major implications for the future of the Medicare program. MA plans have captured steadily higher market share in recent years and now account for around half of all program enrollment (Freed et al. 2022), which suggests that traditional Medicare is becoming a steadily weaker competitor for MA plans. Traditional Medicare's ability to discipline the prices that MA plans pay providers likely hinges on the amount of competitive pressure it applies to MA plans, not its mere presence in the market (Fiedler 2020). In the future, that competitive pressure may become weak enough that the out-of-network cap is no longer sufficient to bring the prices negotiated by MA plans the rest of the way down to traditional Medicare's prices. If MA plans began to pay providers higher prices, this would put upward pressure on plan bids, which would increase federal payments to plans, erode supplemental benefits for MA enrollees, and raise premiums for all Medicare beneficiaries, while also potentially slowing or stopping the decline in traditional Medicare's market share.

This possibility also has implications for the effects of reforms to the Medicare program that would change traditional Medicare's competitive position relative to private plans. For example, in modeling proposals to adopt a "premium support" system in Medicare, the Congressional Budget Office (2017) has assumed that the prices paid by private plans would remain close to the prices paid by traditional

[16] The existence of limits on how much an out-of-network cap can reduce in-network prices can also help explain why MA plans sometimes fail to negotiate in-network prices close to traditional Medicare's. Notably, Lin et al. (2022) estimate that MA plans pay 127% of traditional Medicare's prices for dialysis services, on average.

Medicare (as long as the existing out-of-network cap remained in place and traditional Medicare remained available as an option) even though traditional Medicare's market share would fall sharply, reflecting a sharp deterioration in its competitive position. Our results and the discussion above cast some doubt on this assumption, suggesting that this type of system might result in higher plan bids and, correspondingly, higher federal outlays, than projected.

Appendix A: Solution of Nash Bargaining Problem

This appendix shows that the solution to the Nash bargaining problem specified in equation (1) satisfies equations (2) and (3). To start, we show that any solution must satisfy equation (2): $Q(p^*, l^*) = Q^*$. To that end, consider any (p, l) with $Q(p, l) \neq Q^*$. Define $\delta = [V(Q^*) - cQ^*] - \left[V\big(Q(p, l)\big) - cQ(p, l)\right]$, and observe that $\delta > 0$. Next, define $p' = c + [\pi(p, l) + \delta/2] \,/\, Q^*$ and select $l' \in [0,1]$ such that $Q(p', l') = Q^*$. Straightforward algebra demonstrates that $\pi(p', l') = \pi(p, l) + \delta \,/\, 2$ and $W(p', l') = W(p, l) + \delta \,/\, 2$, which implies that (p, l) does not satisfy equation (1).

We now consider the restricted maximization problem

$$p_r^* = \underset{\substack{p \in \mathbb{R} \text{ s.t.} \\ W(p,l) \geq \widetilde{W}, \pi(p,l) \geq \tilde{\pi}}}{\operatorname{argmax}} \left[V(Q^*) - pQ^* - \widetilde{W}\right]^{\theta} [(p - c)Q^* - \tilde{\pi}]^{1-\theta}.$$

This problem's domain can be rewritten as $p \in \left[c + \tilde{\pi}/Q^*, \big(V(Q^*) - \widetilde{W}\big)/Q^*\right]$. This domain is non-empty since $\widetilde{W} = \tilde{\alpha} W(\tilde{p}, \tilde{l})$ and $\tilde{\pi} = \tilde{\alpha}\pi(\tilde{p}, \tilde{l})$ for some $(\tilde{p}, \tilde{l}, \tilde{\alpha})$ in our setting and, thus,

$$\frac{V(Q^*) - \widetilde{W}}{Q^*} - \left\{c + \frac{\tilde{\pi}}{Q^*}\right\} = \frac{1}{Q^*}\left[\{V(Q^*) - cQ^*\} - \tilde{\alpha}\left\{V\left(Q\big(\tilde{p}, \tilde{l}\big)\right) - cQ\big(\tilde{p}, \tilde{l}\big)\right\}\right] \geq 0.$$

If the domain consists of a single point, then straightforward algebra shows that this single point satisfies $pQ^* = cQ^* + (1 - \theta)[V(Q^*) - cQ^*] + \theta\tilde{\pi} - (1 - \theta)\widetilde{W}$, so the restricted problem has a unique solution, and this solution satisfies equation (3). Alternatively, if the domain is non-degenerate, let f denote the maximand of the restricted problem, and observe that

$$f'(p) = f(p)Q^*\left[\frac{\theta}{V(Q^*) - pQ^* - \widetilde{W}} - \frac{1 - \theta}{(p - c)Q^* - \tilde{\pi}}\right]$$

for $p \in \left(c + \tilde{\pi}/Q^*, \big(V(Q^*) - \widetilde{W}\big)/Q^*\right)$. It is easy to show that $f'(p)$ is strictly decreasing with a unique interior zero that satisfies $pQ^* = cQ^* + (1 - \theta)[V(Q^*) - cQ^*] + \theta\tilde{\pi} - (1 - \theta)\widetilde{W}$. The maximand is strictly positive at this point versus zero on the boundary of the domain, so this point must be the unique solution of the restricted problem and that solution satisfies equation (3). It follows immediately that the full problem has a unique solution with $p^* = p_r^*$ and l^* equal to the unique value of l such that $Q(p_r^*, l) = Q^*$ and that this solution satisfies equations (2) and (3).

Appendix B: Proof of Proposition 1

This section presents a proof of Proposition 1. We begin by making a few definitions that are convenient for what follows. To start, we observe that equations (2) and (3) imply that the provider chooses out-of-network actions $\tilde{p}$ and $\tilde{a}$ that maximize:

$$g(\tilde{p},\tilde{l},\tilde{a}) \equiv \theta\tilde{a}\pi(\tilde{p},\tilde{l}) - (1-\theta)\tilde{a}W(\tilde{p},\tilde{l}) = \tilde{a}[\tilde{p}\tilde{Q} - \theta c\tilde{Q} - (1-\theta)V(\tilde{Q})],$$

where $\tilde{Q} \equiv Q(\tilde{p},\tilde{l})$. Similarly, the insurer chooses the out-of-network coverage level $\tilde{l}$ to minimize $g(\tilde{p},\tilde{l},\tilde{a})$. The function g is essentially just a normalized version of the in-network price $p^*(\tilde{p},\tilde{l},\tilde{a})$ that arises from actions $(\tilde{p},\tilde{l},\tilde{a})$, as $p^*(\tilde{p},\tilde{l},\tilde{a}) = [1/Q^*][\theta cQ^* + (1-\theta)V(Q^*) + g(\tilde{p},\tilde{l},\tilde{a})]$. Working with g is more convenient than working directly with p^* or the associated payoff functions.

For later reference, we state the derivatives of g with respect to the actions $\tilde{p}$ and $\tilde{l}$ in the case where the provider has decided to accept out-of-network patients (so $\tilde{a} = 1$):

$$g_{\tilde{p}}(\tilde{p},\tilde{l},1) = \tilde{Q} + \tilde{Q}_p[\tilde{p} - \theta c - (1-\theta)V'(\tilde{Q})] \tag{B1}$$

$$g_{\tilde{l}}(\tilde{p},\tilde{l},1) = \tilde{Q}_l[\tilde{p} - \theta c - (1-\theta)V'(\tilde{Q})], \tag{B2}$$

where $\tilde{Q}_p \equiv Q_p(\tilde{p},\tilde{l})$ and $\tilde{Q}_l \equiv \tilde{Q}_l(\tilde{p},\tilde{l})$.

We now prove a pair of lemmas showing that $g(\tilde{p},\tilde{l},1)$ is suitably quasi-concave in $\tilde{p}$ and quasi-convex in $\tilde{l}$ and, thus, that these objective functions give rise to well-defined best response functions.

Lemma B1 (Provider). *For any $\tilde{l} < 1$, $g(\tilde{p},\tilde{l},1)$ is strictly quasi-concave as a function of $\tilde{p}$, the provider's best response function $r^P(\tilde{l}) = \operatorname{argmax}_{p\in\mathbb{R}}\, g(\tilde{p},\tilde{l},1)$ is well-defined, and r^P is strictly increasing in $\tilde{l}$. For $\tilde{l} = 1$, $g(\tilde{p},\tilde{l},1)$ is strictly increasing in $\tilde{p}$.*

Proof. Fix some $\tilde{l} < 1$, and observe that equation (B1) can be rewritten as

$$g_{\tilde{p}}(\tilde{p},\tilde{l},1) = \tilde{Q}_p\left[\frac{\tilde{Q}}{\tilde{Q}_p} + \tilde{p} - \theta c - (1-\theta)V'(\tilde{Q})\right].$$

The expression in brackets on the right-hand side is negative for sufficiently small $\tilde{p}$ (e.g., $\tilde{p} = 0$). Furthermore, differentiating this expression with respect to $\tilde{p}$ and applying Assumption QC implies that it is strictly increasing in $\tilde{p}$ with a slope bounded below by some $\epsilon > 0$. Since this expression is also continuous, it must cross zero exactly once at some critical value $\tilde{p}'$ and do so from below. Because $\tilde{Q}_p < 0$, it follows that $g_{\tilde{p}}(\tilde{p}, \tilde{l}, 1)$ has its unique zero at $\tilde{p}'$ and crosses zero from above. Thus, $g(\tilde{p}, \tilde{l}, 1)$ is strictly quasi-concave and $r^P(\tilde{l}) = \tilde{p}'$.

To establish that r^P is strictly increasing, observe first that

$$g_{\tilde{p}\tilde{l}}(\tilde{p}, \tilde{l}, 1) = \tilde{Q}_l[1 - (1-\theta)V''(\tilde{Q})\tilde{Q}_p] + \tilde{Q}_{pl}[\tilde{p} - \theta c - (1-\theta)V'(\tilde{Q})],$$

where $\tilde{Q}_{pl} \equiv Q_{pl}(\tilde{p}, \tilde{l})$. Applying Assumption QC and the other maintained assumptions shows that the first term on the right-hand side is always strictly positive, while the second is weakly positive if $\tilde{p} \geq \theta c + (1-\theta)V'(\tilde{Q})$, which must be the case at $\tilde{p} = r^P(\tilde{l})$. It follows that $g_{\tilde{p}\tilde{l}}(r^P(\tilde{l}), \tilde{l}, 1) > 0$. Since $g_{\tilde{p}}(r^P(\tilde{l}), \tilde{l}, 1) = 0$, this implies that for any $\tilde{l}' > \tilde{l}$ sufficiently close to $\tilde{l}$, we have $g_{\tilde{p}}(r^P(\tilde{l}), \tilde{l}', 1) > 0$ and thus $r^P(\tilde{l}') > r^P(\tilde{l})$. The conclusion that r^P is globally strictly increasing follows immediately.

To see that $g(\tilde{p}, 1,1)$ is strictly increasing in $\tilde{p}$, simply recall that $Q_p(\tilde{p}, 1) = 0$, which implies that $g_{\tilde{p}}(\tilde{p}, 1,1) = Q(\tilde{p}, 1) > 0$, as desired.□

Lemma B2 (Insurer). *For any $\tilde{p} \in \mathbb{R}$, the function $g(\tilde{p}, \tilde{l}, 1)$ is strictly quasi-convex as a function of $\tilde{l}$, and the insurer's best response $r^I(\tilde{p}) = \operatorname{argmin}_{\tilde{l} \in [0,1]} g(\tilde{p}, \tilde{l}, 1)$ is well-defined. Furthermore, r^I is decreasing in $\tilde{p}$.*

Proof. To start, fix some $\tilde{p} \in \mathbb{R}$, and observe that differentiating the term in brackets in equation (B2) with respect to $\tilde{l}$ yields $-(1-\theta)V''(\tilde{Q})\tilde{Q}_l > 0$, which indicates that the term in brackets is strictly increasing in $\tilde{l}$. Since $\tilde{Q}_l > 0$, it follows that $g_{\tilde{l}}(\tilde{p}, \tilde{l}, 1)$ is either everywhere positive, everywhere negative, or has exactly one zero where it crosses zero from below. Thus, $g(\tilde{p}, \tilde{l}, 1)$ is strictly quasi-convex in $\tilde{l}$. Since the interval [0,1] is compact, it follows that r^I is well-defined.

To establish that r^I is decreasing, we first define $\Omega(\tilde{p}) = \{\tilde{l} \in [0,1) : g_{\tilde{l}}(\tilde{p}, \tilde{l}, 1) \geq 0\} \cup \{1\}$ and note that the discussion above implies that $r^I(\tilde{p}) = \inf \Omega(\tilde{p})$. Differentiating the term in brackets in equation (B2) yields $1 - (1-\theta)V''(\tilde{Q})\tilde{Q}_p$, which is strictly positive by Assumption QC, which in turn implies that $\Omega(\tilde{p}) \subset \Omega(\tilde{p}')$ whenever $\tilde{p}' \geq \tilde{p}$. It follows immediately that $r^I(\tilde{p}') \leq r^I(\tilde{p})$ whenever $\tilde{p}' \geq \tilde{p}$.□

We now present the proof of Proposition 1. In doing so, we rely heavily on the facts established in the preceding lemmas, but we omit references to the lemmas to streamline the prose.

Proof of Proposition 1. We proceed by separately considering the scenario corresponding to each of parts (i) through (iii) of the proposition. For each part, we establish that an equilibrium exists, that all equilibria give rise to the same in-network price, and that the other specified properties hold.

Part (i): No out-of-network cap

For this part, we begin with the case where the provider cannot reject out-of-network patients (i.e., when the provider must set $\tilde{a} = 1$). To start, observe that for any $(\tilde{p}, \tilde{l})$ such that $g_{\tilde{p}}(\tilde{p}, \tilde{l}, 1) = 0$, equation (A1) implies that $\tilde{p} - \theta c - (1 - \theta)V'(\tilde{Q}) > 0$, which in combination with equation (A2) implies that $g_{\tilde{l}}(\tilde{p}, \tilde{l}, 1) > 0$. It follows that $\tilde{l} = 0$ is the insurer's best response to $r^P(0)$, so that $(r^P(0), 0)$ is an equilibrium. Furthermore, since $g_{\tilde{p}}(\tilde{p}, \tilde{l}, 1) = 0$ whenever the provider plays a best response, it follows that any equilibrium must have $\tilde{l} = 0$, so this equilibrium is unique.

We now consider the case where the provider can reject out-of-network patients. Observe that $g(\theta c + (1 - \theta)V'(0), \tilde{l}, 1) > 0$ for any $\tilde{l}$ since V is strictly concave. Since $g(\tilde{p}, \tilde{l}, 0) = 0$ for any $\tilde{p}$ and $\tilde{l}$, it follows that setting $\tilde{a} = 0$ can never be a best response for the provider. Thus, the unique equilibrium when the provider cannot reject patients is also the unique equilibrium when it can reject patients.

The preceding discussion establishes that $(0, r^P(0))$ are the desired $(\tilde{p}_{\text{nocap}}, \tilde{l}_{\text{nocap}})$ and that p^*_{nocap} is the corresponding in-network price. The assumed properties of Q imply that $\tilde{Q}_{\text{nocap}} \equiv Q(\tilde{p}_{\text{nocap}}, 0) < Q^*$. To see that $p^*_{\text{nocap}} < \tilde{p}_{\text{nocap}}$, then note that

$$\begin{aligned} p^*_{\text{nocap}} Q^* &= \tilde{p}_{\text{nocap}} \tilde{Q}_{\text{nocap}} + \theta c \left[Q^* - \tilde{Q}_{\text{nocap}}\right] + (1 - \theta)\left[V(Q^*) - V(\tilde{Q}_{\text{nocap}})\right] \\ &< \tilde{p}_{\text{nocap}} \tilde{Q}_{\text{nocap}} + \left[\theta c + (1 - \theta)V'(\tilde{Q}_{\text{nocap}})\right]\left[Q^* - \tilde{Q}_{\text{nocap}}\right] \\ &< \tilde{p}_{\text{nocap}} Q^*, \end{aligned}$$

where: the equality follows from equation (3); the first inequality follows from the strict concavity of V and the fact that $\tilde{Q}_{\text{nocap}} \equiv Q(\tilde{p}_{\text{nocap}}, 0) < Q^*$; and the second inequality follows from the fact that $g_{\tilde{p}}(\tilde{p}_{\text{nocap}}, 0,1) = 0$, which in turn implies that $\theta c + (1 - \theta)V'(\tilde{Q}_{\text{nocap}}) < \tilde{p}_{\text{nocap}}$.

Part (ii): Out-of-network cap and provider cannot reject patients

For this part, we first characterize the equilibrium strategies and then use them to establish the relevant facts about the in-network price. To start, we note that the provider's best response function with an out-of-network cap is $\bar{r}^P(\tilde{l}) = \operatorname{argmax}_{\tilde{p}\in(\infty,\bar{p}]} g(\tilde{p},\tilde{l},1)$. Since $g(\tilde{p},\tilde{l},1)$ is strictly quasi-concave in $\tilde{p}$ for $\tilde{l} < 1$ and strictly increasing in $\tilde{p}$ for $\tilde{l} = 1$, it follows that $\bar{r}^P(\tilde{l}) = \min\{\bar{p}, r^P(\tilde{l})\}$ for $\tilde{l} < 1$ and $r^P(1) = \bar{p}$. We now consider two cases.

We first consider the case $\bar{p} \geq \tilde{p}_{\text{nocap}}$. The equilibrium without an out-of-network cap derived in part (i) remains an equilibrium since $\bar{r}^P(0) = \tilde{p}_{\text{nocap}}$. It also remains true that any equilibrium must have $\tilde{l} = 0$ since either the provider opts for an interior solution with $g_{\tilde{p}}(\tilde{p},\tilde{l},1) = 0$, in which case the arguments from part (i) imply that $\tilde{l} = 0$, or the provider sets $\tilde{p} = \bar{p} \geq \tilde{p}_{\text{nocap}}$, in which case the fact that r^I is a decreasing in $\tilde{p}$ implies that $\tilde{l} = r^I(\tilde{p}) \leq r^I(\tilde{p}_{\text{nocap}}) = 0$. Thus, this equilibrium remains unique.

Now, consider the case where $\bar{p} < \tilde{p}_{\text{nocap}}$. Observe first that $\bar{r}^P(r^I(\tilde{p})) = \bar{p}$ for any $\tilde{p} \leq \bar{p}$. This is immediate if $r^I(\tilde{p}) = 1$; otherwise, it follows from the fact that r^I is decreasing, so $r^I(\tilde{p}) \geq r^I(\tilde{p}_{\text{nocap}}) = 0$, which together with the fact that r^P is increasing implies that $r^P(r^I(\tilde{p})) \geq r^P(0) = \tilde{p}_{\text{nocap}} > \bar{p}$. This implies that $(\bar{p}, r^I(\bar{p}))$ is an equilibrium and, in fact, the unique equilibrium.

We now let $p^*_{\text{accept}}(\bar{p})$ denote the in-network price that arises in the unique equilibrium under an out-of-network cap $\bar{p}$ and show that it has the desired properties. It is immediate that $p^*_{\text{accept}}(\bar{p}) = p^*_{\text{nocap}}$ for $\bar{p} \geq \tilde{p}_{\text{nocap}}$. To handle the cases $\bar{p} \in (c, \tilde{p}_{\text{nocap}})$ and $\bar{p} = c$, note that any out-of-network actions $(\tilde{p},\tilde{l})$ such that $Q(\tilde{p},\tilde{l}) = Q^*$ result in an in-network price of $\tilde{p}$ by equation (3). Thus, for any $\bar{p} \in (c, \tilde{p}_{\text{nocap}})$, we can choose $\bar{l}$ such that $Q(\bar{p},\bar{l}) = Q^*$ and note that $\bar{l} > r^I(\bar{p})$ since $g_{\tilde{l}}(\bar{p},\bar{l},1) = Q_l(\bar{p},\bar{l})[\bar{p} - c] > 0$; it follows that $\bar{p}Q^* - p^*_{\text{accept}}(\bar{p})Q^* = g(\bar{p},\bar{l},1) - g(\bar{p}, r^I(\bar{p}),1) > 0$, so $p^*_{\text{accept}}(\bar{p}) < \bar{p}$. Similarly, for $\bar{p} = c$, we note that $g(c,\tilde{l},1) = (1-\theta)\left[cQ(c,\tilde{l}) - V\left(Q(c,\tilde{l})\right)\right]$, so $Q(c, r^I(c)) = Q^*$ and $p^*_{\text{accept}}(c) = c$.

Finally, we note that a suitable envelope theorem (e.g., Corollary 4 in Milgrom and Segal 2002) implies that the function $g(\bar{p}, r^I(\bar{p}),1)$ is differentiable as a function of $\bar{p}$ on the domain $[c, \tilde{p}_{\text{nocap}}]$, with derivative $g_{\tilde{p}}(\bar{p}, r^I(\bar{p}),1)$, so p^*_{accept} is differentiable on this domain with $\left(p^*_{\text{cap}}\right)'(\bar{p}) = g_{\tilde{p}}(\bar{p}, r^I(\bar{p}),1)\;/$

Q^*. Since $g_{\tilde{p}}\big(\tilde{p}_{\text{nocap}}, r^I(\tilde{p}_{\text{nocap}}), 1\big) = 0$, it follows that p^*_{accept} is, in fact, differentiable on $[c, \infty)$. Additionally, for $\bar{p} \in [c, \tilde{p}_{\text{nocap}})$, recall that $r^P\big(r^I(\bar{p})\big) > \bar{p}$, which implies that $g_{\tilde{p}}(\bar{p}, r^I(\bar{p}), 1) > 0$, so $\big(p^*_{\text{accept}}\big)'(\bar{p}) > 0$.

Part (iii): Out-of-network cap and provider can reject patients

For this part, we begin by showing that there is a unique $\bar{p} \in (c, \tilde{p}_{\text{nocap}})$ such that $g(\bar{p}, r^I(\bar{p}), 1) = 0$. The fact that the provider never wishes to turn away out-of-network patients without an out-of-network cap, which was established in part (i), implies that $g\big(\tilde{p}_{\text{nocap}}, r^I(\tilde{p}_{\text{nocap}}), 1\big) > 0$. Similarly, it was shown in part (ii) that $Q\big(c, r^I(c)\big) = Q^*$, which implies that $g(c, r^I(c), 1) = (1-\theta)[cQ^* - V(Q^*)] < 0$. Since $g(\bar{p}, r^I(\bar{p}), 1)$ is differentiable (hence continuous) and strictly decreasing as a function of $\bar{p}$ on the domain $[c, \tilde{p}_{\text{nocap}}]$, the desired $\bar{p}$ must exist.

We label this unique value $\bar{p}_{\text{reject}}$ and verify that it has the desired properties. When $\bar{p} > \bar{p}_{\text{reject}}$, observe that for any $\tilde{l}$, $g\big(\bar{p}, \tilde{l}, 1\big) \geq g(\bar{p}, r^I(\bar{p}), 1) > 0$, where the first inequality follows because r^I is the insurer's best response and the second follows from the definition of $\bar{p}_{\text{reject}}$, so rejecting out-of-network patients is never a best response for the provider. It follows that the equilibrium is unchanged from the case where the provider cannot reject patients.

When $\bar{p} < \bar{p}_{\text{reject}}$, the definition of $\bar{p}_{\text{reject}}$ implies that $g(\bar{p}, r^I(\bar{p}), 1) < 0$, so playing $\tilde{a} = 1$ is no longer a best response for the provider when the insurer plays $r^I(\bar{p})$. Since $\big(\bar{p}, r^I(\bar{p})\big)$ was the only equilibrium when the provider was required to play $\tilde{a} = 1$, it follows that any equilibrium must have $\tilde{a} = 0$. There are many such equilibria, but it is easy to see that $(\bar{p}, r^I(\bar{p}), 0)$ is one of them and that all lead to disagreement payoffs $\tilde{\pi} = \widetilde{W} = 0$ and, thus, an in-network price $p^*_{\text{accept}}(\bar{p}_{\text{reject}})$.

When $\bar{p} = \bar{p}_{\text{reject}}$, similar arguments imply that $(\bar{p}, r^I(\bar{p}), 1)$ is an equilibrium, and that there are also equilibria with $\tilde{a} = 0$. All have $g\big(\tilde{p}, \tilde{l}, \tilde{a}\big) = 0$ and, thus, an in-network price $p^*_{\text{accept}}(\bar{p}_{\text{reject}})$.□

Appendix C: Derivation of Equation (7)

This appendix derives equation (7) from the main text. Since h is assumed to be a quadratic function, h' is a linear function, and $h'([Q_0+Q_1]/2) = \big(h(Q_1)-h(Q_0)\big)/(Q_1-Q_0)$ for any Q_0 and Q_1.

Using these properties and the fact that $h(0)=0$, we obtain

$$\begin{aligned}
\frac{h\big(\tilde{Q}_{\text{nocap}}\big)}{\tilde{Q}_{\text{nocap}}} - \frac{h(Q^*) - h\big(\tilde{Q}_{\text{nocap}}\big)}{Q^* - \tilde{Q}_{\text{nocap}}} &= h'\left(\frac{\tilde{Q}_{\text{nocap}}}{2}\right) - h'\left(\frac{\tilde{Q}_{\text{nocap}} + Q^*}{2}\right) \\
&= \left[h'\left(\frac{\tilde{Q}_{\text{nocap}} + Q^*}{2}\right) - h'(Q^*)\right]\left[\frac{Q^*}{Q^* - \tilde{Q}_{\text{nocap}}}\right] \\
&= \left[\frac{h(Q^*) - h\big(\tilde{Q}_{\text{nocap}}\big)}{Q^* - \tilde{Q}_{\text{nocap}}} - h'(Q^*)\right]\left[\frac{Q^*}{Q^* - \tilde{Q}_{\text{nocap}}}\right].
\end{aligned}$$

Recalling that $h(Q^*) - h\big(\tilde{Q}_{\text{nocap}}\big) = p^*_{\text{nocap}}Q^* - \tilde{p}_{\text{nocap}}\tilde{Q}_{\text{nocap}}$ and $h'(Q^*) = c$, we obtain

$$\frac{h\big(\tilde{Q}_{\text{nocap}}\big)}{\tilde{Q}_{\text{nocap}}} = \frac{p^*_{\text{nocap}}Q^* - \tilde{p}_{\text{nocap}}\tilde{Q}_{\text{nocap}}}{Q^* - \tilde{Q}_{\text{nocap}}} + \left[\frac{\pi^*_{\text{nocap}} - \tilde{\pi}_{\text{nocap}}}{Q^* - \tilde{Q}_{\text{nocap}}}\right]\left[\frac{Q^*}{Q^* - \tilde{Q}_{\text{nocap}}}\right],$$

from which equation (7) follows immediately.

Appendix D: Adding Costs of Being Out-of-Network

This appendix considers how our conclusions would change if the provider and insurer bear additional costs in the absence of a network agreement. Let $k_P(\tilde{Q})$ and $k_I(\tilde{Q})$ denote the additional costs borne by the provider and insurer, respectively. For convenience, we also let $\bar{k}(\tilde{Q}) \equiv \theta k_P(\tilde{Q}) - (1-\theta)k_I(\tilde{Q})$, the weighted difference between the additional costs borne by the two parties.

We assume that the functions k_P and k_I are specified so that the equilibrium of the modified model has the same basic structure as the equilibrium of the original model. That is: (1) the model has an (essentially) unique equilibrium; (2) without an out-of-network cap, the provider does not turn away patients and the parties' actions result in $\tilde{Q}_{\text{nocap}} \in (0, Q^*)$; and (3) with an out-of-network cap, the provider turns away out-of-network patients once the cap becomes stringent enough that it can obtain a higher negotiated price by doing so. These conditions will be satisfied for many different specifications of k_P and k_I, including specifications in which the parties bear a fixed cost of being out-of-network and specifications in which these costs vary linearly with sufficiently small slope.

Under these conditions, equation (4) is replaced by an analogue with an additional term:

$$\left[p^*_{\text{nocap}} - p^*_{\text{accept}}(\bar{p}_{\text{reject}})\right]Q^* = \tilde{p}_{\text{nocap}}\tilde{Q}_{\text{nocap}} - h(\tilde{Q}_{\text{nocap}}) - \left[\bar{k}(\tilde{Q}_{\text{nocap}}) - \bar{k}(0)\right], \tag{4$'$}$$

Similarly, equation (3) now implies that $h(Q^*) - h(\tilde{Q}_{\text{nocap}}) = p^*_{\text{nocap}}Q^* - \tilde{p}_{\text{nocap}}\tilde{Q}_{\text{nocap}} + \bar{k}(\tilde{Q}_{\text{nocap}})$, so the analogue to equation (5) in the modified model is

$$h(\tilde{Q}_{\text{nocap}}) > [p^*_{\text{nocap}}Q^* - \tilde{p}_{\text{nocap}}\tilde{Q}_{\text{nocap}} + \bar{k}(\tilde{Q}_{\text{nocap}})]\left[\frac{\tilde{Q}_{\text{nocap}}/Q^*}{1 - \tilde{Q}_{\text{nocap}}/Q^*}\right]. \tag{5$'$}$$

Combining equations (4') and (5'), we obtain a modified upper bound

$$\frac{p^*_{\text{nocap}} - p^*_{\text{accept}}(\bar{p}_{\text{reject}})}{p^*_{\text{nocap}}} < \left[\frac{\tilde{Q}_{\text{nocap}}/Q^*}{1 - \tilde{Q}_{\text{nocap}}/Q^*}\right] \cdot \left[\frac{\tilde{p}_{\text{nocap}}}{p^*_{\text{nocap}}} - \frac{\left[\bar{k}(\tilde{Q}_{\text{nocap}}) - \bar{k}(0)\right]/\tilde{Q}_{\text{nocap}} + \bar{k}(0)/Q^*}{p^*_{\text{nocap}}} - 1\right]. \tag{6$'$}$$

The key difference between this upper bound and the original upper bound is the presence of a term that reflects the additional costs the parties incur when out-of-network. When the costs borne by the provider are large relative to the costs borne by the insurer, the scope for an out-of-network cap to reduce in-network prices will be smaller than implied by our original bound. When the costs borne by the insurer are relatively large, the opposite will be true.

Similarly, we can derive a modified version of the exact expression for the maximum achievable price reduction presented in equation (8). The analogue to equation (7) in the modified model is

$$h(\tilde{Q}_{\text{nocap}}) = \left[p^*_{\text{nocap}} Q^* - \tilde{p}_{\text{nocap}} \tilde{Q}_{\text{nocap}} + \bar{k}(\tilde{Q}_{\text{nocap}}) + \frac{\pi^*_{\text{nocap}} - \tilde{\pi}_{\text{nocap}} + \bar{k}(\tilde{Q}_{\text{nocap}})}{1 - \tilde{Q}_{\text{nocap}}/Q^*}\right]\left[\frac{\tilde{Q}_{\text{nocap}}/Q^*}{1 - \tilde{Q}_{\text{nocap}}/Q^*}\right]. \tag{7'}$$

Combining equations (4') and (7'), we obtain an analogue to equation (8):

$$\frac{p^*_{\text{nocap}} - p^*_{\text{accept}}(\bar{p}_{\text{reject}})}{p^*_{\text{nocap}}} = \left[\frac{\tilde{Q}_{\text{nocap}}/Q^*}{1 - \tilde{Q}_{\text{nocap}}/Q^*}\right] \cdot \left[\frac{\tilde{p}_{\text{nocap}}}{p^*_{\text{nocap}}} - \frac{[\bar{k}(\tilde{Q}_{\text{nocap}}) - \bar{k}(0)]/\tilde{Q}_{\text{nocap}} + \bar{k}(0)/Q^*}{p^*_{\text{nocap}}} - 1 - \frac{\frac{\pi^*_{\text{nocap}}}{p^*_{\text{nocap}} Q^*} - \frac{\tilde{\pi}_{\text{nocap}}}{p^*_{\text{nocap}} Q^*} + \frac{\bar{k}(\tilde{Q}_{\text{nocap}})}{p^*_{\text{nocap}} Q^*}}{1 - \tilde{Q}_{\text{nocap}}/Q^*}\right]. \tag{8'}$$

Once again, the key difference relative to the original expression is the presence of expressions related to the additional costs borne by the parties when they are out-of-network. As above, when the costs borne by the provider are large relative to the costs borne by the insurer, the scope for an out-of-network cap to reduce in-network prices will be smaller than implied by the original version. When the costs borne by the insurer are relatively large, the opposite will be the case.

References

Abreu, Dilip, and David Pearce. 2007. "Bargaining, Reputation, and Equilibrium Selection in Repeated Games with Contracts." *Econometrica* 75 (3): 653–710. https://doi.org/10.1111/j.1468-0262.2007.00765.x.

Berenson, Robert A., and Robert B. Murray. 2022. "How Price Regulation Is Needed To Advance Market Competition." *Health Affairs* 41 (1): 26–34. https://doi.org/10.1377/hlthaff.2021.01235.

Berenson, Robert A., Jonathan H. Sunshine, David Helms, and Emily Lawton. 2015. "Why Medicare Advantage Plans Pay Hospitals Traditional Medicare Prices." *Health Affairs* 34 (8): 1289–95. https://doi.org/10.1377/hlthaff.2014.1427.

Biener, Adam I., Benjamin L. Chartock, Christopher Garmon, and Erin Trish. 2021. "Emergency Physicians Recover A Higher Share Of Charges From Out-Of-Network Care Than From In-Network Care." *Health Affairs* 40 (4): 622–28. https://doi.org/10.1377/hlthaff.2020.01471.

Blumberg, Linda J., John Holahan, Stacey McMorrow, and Michael Simpson. 2020. "Estimates of the Implications of Public Option and Capped Provider Payment Rate Reforms." Urban Institute. https://www.urban.org/research/publication/estimates-implications-public-option-and-capped-provider-payment-rate-reforms.

Brown, Amy M., Sandra L. Decker, and Frederic W. Selck. 2015. "Emergency Department Visits and Proximity to Patients' Residences, 2009-2010." *NCHS Data Brief*, no. 192 (March): 1–8.

Chernew, Michael E., Leemore S. Dafny, and Maximilian J. Pany. 2020. "A Proposal to Cap Provider Prices and Price Growth in the Commercial Health-Care Market." Brookings Institution. https://www.hamiltonproject.org/assets/files/CDP_PP_WEB_FINAL.pdf.

Chernew, Michael E., Andrew L. Hicks, and Shivani A. Shah. 2020. "Wide State-Level Variation In Commercial Health Care Prices Suggests Uneven Impact Of Price Regulation." *Health Affairs* 39 (5): 791–99. https://doi.org/10.1377/hlthaff.2019.01377.

Chernew, Michael E., Maximilian J. Pany, and Richard G. Frank. 2019. "The Case for Market-Based Price Caps." *Health Affairs Blog* (blog). September 3, 2019. https://www.healthaffairs.org/do/10.1377/hblog20190826.369708/full/.

Clemens, Jeffrey, and Joshua D. Gottlieb. 2016. "In the Shadow of a Giant: Medicare's Influence on Private Physician Payments." *Journal of Political Economy* 125 (1): 1–39. https://doi.org/10.1086/689772.

Congressional Budget Office. 2022. "The Prices That Commercial Health Insurers and Medicare Pay for Hospitals' and Physicians' Services." https://www.cbo.gov/publication/57422.

Congressional Budget Office (CBO). 2017. "A Premium Support System for Medicare: Updated Analysis of Illustrative Options." https://www.cbo.gov/publication/53077.

Cooper, Zack, Stuart V. Craig, Martin Gaynor, and John Van Reenen. 2019. "The Price Ain't Right? Hospital Prices and Health Spending on the Privately Insured." *The Quarterly Journal of Economics* 134 (1): 51–107. https://doi.org/10.1093/qje/qjy020.

Cooper, Zack, Fiona Scott Morton, and Nathan Shekita. 2020. "Surprise! Out-of-Network Billing for Emergency Care in the United States." *Journal of Political Economy* 128 (9): 3626–3677. https://doi.org/10.1086/708819.

Doyle, Joseph, John Graves, and Jonathan Gruber. 2019. "Evaluating Measures of Hospital Quality: Evidence from Ambulance Referral Patterns." *The Review of Economics and Statistics* 101 (5): 841–52. https://doi.org/10.1162/rest_a_00804.

Duffy, Erin, Christopher Whaley, and Chapin White. 2020. "The Price and Spending Impacts of Limits on Payments to Hospitals for Out-of-Network Care." RAND Corporation. https://doi.org/10.7249/RR4378.

Fiedler, Matthew. 2020. "Capping Prices or Creating a Public Option: How Would They Change What We Pay for Health Care?" Brookings Institution.

https://www.brookings.edu/research/capping-prices-or-creating-a-public-option-how-would-they-change-what-we-pay-for-health-care/.

Freed, Meredith, Jeannie Fuglesten Biniek, Anthony Damico, and Tricia Neuman. 2022. "Medicare Advantage in 2022: Enrollment Update and Key Trends." https://www.kff.org/medicare/issue-brief/medicare-advantage-in-2022-enrollment-update-and-key-trends/.

Gowrisankaran, Gautam, Aviv Nevo, and Robert Town. 2015. "Mergers When Prices Are Negotiated: Evidence from the Hospital Industry." *American Economic Review* 105 (1): 172–203. https://doi.org/10.1257/aer.20130223.

Health Care Cost Institute (HCCI). 2020. "2020 Healthy Marketplace Index, Technical Appendix." https://healthcostinstitute.org/images/pdfs/hmi_2020_technical_appendix.pdf.

Ho, Kate, and Robin S. Lee. 2017. "Insurer Competition in Health Care Markets." *Econometrica* 85 (2): 379–417. https://doi.org/10.3982/ECTA13570.

Ippolito, Benedic, and Boris Vabson. 2023. "How Should Policymakers Respond To Growth In Cost Sharing That Often Goes Unpaid?" *Health Affairs Forefront*, January. https://doi.org/10.1377/forefront.20230113.900882.

Kaiser Family Foundation (KFF). 2020. "2020 Employer Health Benefits Survey." Kaiser Family Foundation. https://www.kff.org/health-costs/report/2020-employer-health-benefits-survey/.

Kane, Arielle. 2019. "Affordable Health Care for All: An American Solution to High Costs and Coverage Gaps." Progressive Policy Institute. https://www.progressivepolicy.org/wp-content/uploads/2021/07/PPI_UniversalHealthcare_v3.pdf.

Lin, Eugene, Bich Ly, Erin Duffy, and Erin Trish. 2022. "Medicare Advantage Plans Pay Large Markups To Consolidated Dialysis Organizations." *Health Affairs* 41 (8): 1107–16. https://doi.org/10.1377/hlthaff.2021.02009.

Maeda, Jared Lane K., and Lyle Nelson. 2018. "How Do the Hospital Prices Paid by Medicare Advantage Plans and Commercial Plans Compare with Medicare Fee-for-Service Prices?" *INQUIRY: The Journal of Health Care Organization, Provision, and Financing* 55 (January): 1–8. https://doi.org/10.1177/0046958018779654.

Medicare Payment Advisory Commission (MedPAC). 2017. "Medicare and the Health Care Delivery System." https://www.medpac.gov/wp-content/uploads/import_data/scrape_files/docs/default-source/reports/jun17_reporttocongress_sec.pdf.

———. 2023. "Medicare Payment Policy." https://www.medpac.gov/wp-content/uploads/2023/03/Mar23_MedPAC_Report_To_Congress_SEC.pdf.

Melnick, Glenn, and Katya Fonkych. 2020a. "An Empirical Analysis of Hospital ED Pricing Power." *AJMC* 26 (3). https://www.ajmc.com/journals/issue/2020/2020-vol26-n3/an-empirical-analysis-of-hospital-ed-pricing-power.

———. 2020b. "Regulating Out-Of-Network Hospital Emergency Prices: Problem And Potential Benchmarks." *Health Affairs Blog* (blog). March 23, 2020. https://www.healthaffairs.org/do/10.1377/hblog20200320.866552/full/.

Milgrom, Paul, and Ilya Segal. 2002. "Envelope Theorems for Arbitrary Choice Sets." *Econometrica* 70 (2): 583–601. https://doi.org/10.1111/1468-0262.00296.

Murray, Robert. 2013. "Hospital Charges and the Need for a Maximum Price Obligation Rule for Emergency Department & Out-of-Network Care." *Health Affairs Blog* (blog). May 16, 2013. https://www.healthaffairs.org/do/10.1377/hblog20130516.031255/full/.

Murray, Robert, and Jack Keane. 2022. "Setting Caps on Out-of-Network Hospital Payments: A Low-Intensity Regulatory Intervention for Reducing Hospital Prices Overall." Commonwealth Fund. https://doi.org/10.26099/tvdr-1w88.

Nash, John. 1953. "Two-Person Cooperative Games." *Econometrica* 21 (1): 128–40. https://doi.org/10.2307/1906951.

Pelech, Daria M. 2020. "Prices for Physicians' Services in Medicare Advantage and Commercial Plans." *Medical Care Research and Review* 77 (3): 236–48. https://doi.org/10.1177/1077558718780604.

Pope, Chris. 2019. "End Hospitals' Right to a Blank Check for Emergency Care." *Economics21* (blog). May 3, 2019. https://economics21.org/end-hospital-emergency-medical-billing-blank-check.

Prager, Elena, and Nicholas Tilipman. 2020. "Regulating Out-of-Network Hospital Payments: Disagreement Payoffs, Negotiated Prices, and Access." https://sites.google.com/view/eprager/research.

Shepard, Mark. 2022. "Hospital Network Competition and Adverse Selection: Evidence from the Massachusetts Health Insurance Exchange." *American Economic Review* 112 (2): 578–615. https://doi.org/10.1257/aer.20201453.

Song, Zirui. 2017. "Using Medicare Prices — Toward Equity and Affordability in the ACA Marketplace." *New England Journal of Medicine* 377 (24): 2309–11. https://doi.org/10.1056/NEJMp1710020.

Song, Zirui, William Johnson, Kevin Kennedy, Jean Fuglesten Biniek, and Jacob Wallace. 2020. "Out-Of-Network Spending Mostly Declined In Privately Insured Populations With A Few Notable Exceptions From 2008 To 2016." *Health Affairs* 39 (6): 1032–41. https://doi.org/10.1377/hlthaff.2019.01776.

Tilipman, Nicholas. 2022. "Employer Incentives and Distortions in Health Insurance Design: Implications for Welfare and Costs." *American Economic Review* 112 (3): 998–1037. https://doi.org/10.1257/aer.20181917.

Trish, Erin, Paul Ginsburg, Laura Gascue, and Geoffrey Joyce. 2017. "Physician Reimbursement in Medicare Advantage Compared with Traditional Medicare and Commercial Health Insurance." *JAMA Internal Medicine* 177 (9): 1287–95. https://doi.org/10.1001/jamainternmed.2017.2679.

Venkatesh, Arjun K., Hao Mei, Keith E. Kocher, Michael Granovsky, Ziad Obermeyer, Erica S. Spatz, Craig Rothenberg, Harlan M. Krumholz, and Zhenqui Lin. 2017. "Identification of Emergency Department Visits in Medicare Administrative Claims: Approaches and Implications." *Academic Emergency Medicine* 24 (4): 422–31. https://doi.org/10.1111/acem.13140.

Whaley, Christopher M., Brian Briscombe, Rose Kerber, Brenna O'Neill, and Aaron Kofner. 2020. "Nationwide Evaluation of Health Care Prices Paid by Private Health Plans: Findings from Round 3 of an Employer-Led Transparency Initiative." RAND Corporation. https://www.rand.org/pubs/research_reports/RR4394.html.