The Cost of Charitable Care: Resource Externalities and the Value of Insurance *

Xuelin Li† Tong Liu‡ Benjamin Vatter§

September 2026

Abstract

This paper studies the private and public costs of free and discounted care, motivated by its central role in the American healthcare safety net. Leveraging cross-state spillovers from insurance access expansions, we estimate the private cost of this charitable care and find that it crowds out insured care. We develop a model of competitive hospital resource allocations and estimate it using detailed discharge and capacity data. The burden of charitable care reduces emergency care supply, where regulation limits hospitals' ability to discriminate among patients, but increases insured scheduled care. The cost of insuring otherwise charitable care is between two and seven times the benefit it offers recipients. However, over three-quarters of the value of expanding insurance accrues to non-recipients because of its effects on hospital resource allocation. This flips the welfare sign of insuring the uninsured, suggesting an optimal coverage of charitable care at rates comparable to Medicaid. We investigate alternative policies that set minimum requirements for non-profit hospitals and finance charitable care delivery through credit trade. We find a stark trade-off between policy targeting and fiscal spending driven by a misalignment between charitable care needs and non-profit tax exemption benefits.

Keywords: Charity care, uncompensated care, hospital capacity, resource externalities, Medicaid expansion, hospital competition, nonprofit hospitals.
JEL Codes: I11, I13, I18, D24, D62, L13, L31

*†Columbia Business School, Columbia University. Email: xuelin.li@columbia.edu. ‡Sloan School of Management, Massachusetts Institute of Technology. Email: tongl@mit.edu. §Sloan School of Management, Massachusetts Institute of Technology. Email: bvatter@mit.edu. Liu gratefully acknowledges funding support from MIT Sloan Health Systems Initiative (HSI).

1 Introduction

Regulation, technology, and information often constrain firms' ability to discriminate service among consumers. In such environments, when the pool of consumers becomes less profitable in ways not correctable through price or quality adjustments, supply naturally contracts. The freed productive resources, however, are often reallocated rather than left idle, as observed in housing markets (Diamond *et al.*, 2019), education (Bound *et al.*, 2020), and banking (Chakraborty *et al.*, 2018; Bidder *et al.*, 2021). Thus, when the marginal consumer to a firm is less profitable than the opportunity cost of the resources it occupies, its presence induces an externality on others. For those seeking the same output as the marginal consumer, these are standard congestion and selection externalities (Naor, 1969; Akerlof, 1970). In contrast, for those seeking the firm's alternative output, these are positive reallocation externalities from increased supply. We refer to this combination of forces as resource externalities.

This article studies the equilibrium effects of resource externalities on the provision of healthcare services and its implications for the value of insurance. Hospitals routinely provide care for free or at heavily discounted rates, known as charity care. This care constitutes a central part of the American healthcare safety net, treating some of the 23.5 million working-age adults who lack insurance (Cohen and Briones, 2026) and the estimated 45 million adults who are underinsured (Collins and Gupta, 2024). Hospitals provide approximately 1.4 million charitable-care admissions per year and report $31.3 billion in associated costs (Agency for Healthcare Research and Quality, 2025; Centers for Medicare & Medicaid Services, 2026a), which are supported by state and federal governments through tax benefits and indirect compensation policies. Yet charitable care is largely unprofitable, delivered mainly through emergency care, where regulation strictly limits hospitals' ability to deny care. Using hospital discharge data from the period 2010 to 2019, this work examines how demand for charitable care shapes hospitals' incentives to allocate resources across care types, how those decisions channel the resource externalities of uninsurance, and how these externalities affect the public value of insuring those at the margin of charitable care.

We document three key findings regarding the private and public costs of charitable care. First, we find that charitable care crowds out insured care, decreasing hospitals' profits by substituting compensated for uncompensated care. Second, we find that charitable care patients' access to care is limited by their unprofitability. However, compensating hospitals for this care costs between two and seven times as much as the surplus charitable care patients would obtain from this coverage. Third, we find that more than three-quarters of the surplus generated by directly compensating hospitals for otherwise charitable care falls on those already insured. This benefit to non-recipients is large enough to offset the fiscal cost of the coverage, resulting in optimal coverage at or above Medicaid levels across a wide range of regulatory preferences. The impact on the insured quantifies the resource externalities: insured patients seeking scheduled care are harmed by resource reallocation, which is offset by improved access to emergency care.

Studying the externalities of charitable care requires identifying its associated cost and crowd-out. This requires a change in demand for charitable care that is plausibly exogenous to unobserved changes in hospitals' production technology and insured patients' preferences. This, however, poses a significant challenge as those seeking charitable care are unresponsive to changes in networks, prices, or other regulatory shocks commonly used as shifters of insured demand. Moreover, charity care lacks a price that can be inverted to recover hospitals' costs, as used in the literature on the industrial organization of healthcare markets (Gaynor and Vogt, 2003; Gowrisankaran *et al.*, 2015). Our strategy instead relies on a shift produced by a reallocation of demand following a shock to some hospitals' supply incentives. In particular, we build on the observation that non-profit hospital systems might have incentives (altruistic or otherwise) to maintain stable system-wide levels of charitable care (Almond *et al.*, 2011). This hypothesis is not unfounded: In tax filings, the IRS explicitly asks non-profit systems whether they established a budget for charitable care, which 93% of systems confirm. In addition, we document significant elasticity of donations and grants with respect to system-level charitable care spending. Thus, if the hypothesis holds, a system that sees a decrease in charitable care demand in one market might be incentivized to increase supply in another market. If aggregate demand for charitable care is invariant to moderate supply changes, as documented in prior literature (Garthwaite *et al.*, 2018; Moran, 2025), this increase will translate into a reallocation of charitable care, reducing demand for all other hospitals in the same market.

Following this logic, we uncover a cross-state spillover from Medicaid access expansions following the Affordable Care Act (ACA). We show that, consistent with a reduction in uninsurance rate, hospitals in states that expanded Medicaid lose 1,543.8 charitable care bed-days per year. In contrast, hospitals in non-expansion states belonging to the same non-profit systems with member hospitals in expansion states gain 1,252.8 charitable care bed-days per year. We show that the aggregate demand for charitable care in non-expansion states is not affected by the expansion of coverage in other states and that the additional charitable care patients served by these multistate member hospitals—which we call *spillover* hospitals—is driven by a reallocation of patients across hospitals within the non-expansion state. We also show that this increase in charitable care at spillover hospitals minimizes the negative impact of Medicaid expansion on their system-wide charitable care spending. Given the timing of the effects, findings on the quality and types of care delivered, and the non-profit charitable objective hypothesis, we interpret this reallocation as driven by spillovers from Medicaid expansions in other states. Crucially, this shock provides the variation needed to resolve the identification challenge—hospitals near spillover hospitals experience a plausibly exogenous (to their technology of production and insured population preferences) contraction in demand for charitable care.

These spillovers allow us to quantify the impact of charitable care on hospital costs, revenue, and private care. We estimate that an additional charitable care admission entails an average service cost of $13,819.6. Extrapolating to the population and multiplying by their admission risk, each uninsured individual in the country carries an expected uncompensated hospital cost

of $1,177.8, which closely matches the (inflation-adjusted) $1,125 estimate from Garthwaite *et al.* (2018), identified from changes in insurance coverage within states during 1988 and 2011. Broadly, our estimates imply an expected total spending on charitable care in 2019 of $42.3 billion, comparable to the realized value of $40.5 billion. In addition, our evidence shows that every additional charitable care admission decreases hospital revenue by $295,000 per patient and displaces 12.7 insured patient bed-days. The revenue impact shows that hospitals cannot fully shift the burden of charitable care to insured patients, which contributes to a broad literature on cost-shifting and cross-subsidization (Cutler, 1998; Frakt, 2011; White, 2013; Clemens and Gottlieb, 2017; Dranove *et al.*, 2017). The crowding out of insured patients shows the direct effects of the resource externality imposed by charitable care patients. Uncovering the full extent of these externalities, however, requires simulating the outcomes that would result if care for those patients were compensated, which necessitates a model of hospital resource allocation.

We develop an equilibrium model of hospital competition and inpatient demand. Hospitals in the model allocate limited bed-day resources across service lines (e.g., emergency care) and manage admission requests in each. Bed-days serve as an anchor resource shared across lines and tie production to hospitals' key capital constraint: capacity. Admission decisions succinctly capture how hospital queues, rental resources, and flexible staffing are used. We show that if a hospital line manages a first-come, first-served queue with guaranteed maximum wait times, it effectively controls the fraction of requested admissions it accepts in every period. We use this logic to simplify the line problem such that each hospital line chooses the rate at which it accepts patients who request admission. This choice must, in expectation, not exceed the line's allocated resources. The model thus accommodates both observed hospital overflows and the variety of physical and regulatory constraints dictated by their resources. In turn, for patients seeking care among hospitals, these decisions determine which hospitals can admit them. We show that despite the constraints on hospital behavior, the model is well-behaved: A pure-strategy equilibrium always exists, and under some technical qualifications, it is unique. In addition, as long as hospital-insurer negotiations do not condition on hospital resource allocations, the model nests within the standard workhorse framework for hospital-insurer competition (Handel and Ho, 2021).

We estimate the model using patient-level discharge data from Florida from 2011 to 2019. The model provides a clear argument for identifying hospital costs from variation in demand. Intuitively, a hospital with multiple service lines allocates resources to equate its marginal profitability across lines. As data on prices and demand identify revenue differences across service lines, changes in those elements reveal cost differences across lines. We show that this argument, together with the appropriate data variation and the hospital's total resource constraint, identifies hospital costs. This identification argument yields a natural estimator of hospital costs based on the likelihood of patients being admitted to each service line. The estimator imposes no assumptions about how hospital prices are determined, other than that they are negotiated before resource allocations and that insurers do not contract over hospital capacity, as implicitly assumed

in the literature. Our approach thus identifies the rate at which patients are admitted to hospitals, following the logic of Agarwal and Somaini (2026), and then combines this with observed total capacity and marginal revenues in each quarter to estimate costs.

Our estimates show how markups and access differ across patients and over time. As a fraction of the median hospital's costs, we find that Medicare markups fell from -1.3% in 2013 to -11.4% in 2019, while private insurance markups fell from 229.3% to 210.1% during the same period. Overall hospital margins on inpatient care fell from 12.3% in 2013 to 6.3% in 2019, evidence that contributes to the extensive debate on the evolution of markups in the US economy (Basu, 2019; Berry *et al.*, 2019; De Loecker *et al.*, 2020; Grieco *et al.*, 2024; Shapiro and Yurukoglu, 2026), as hospital care constitutes over five percent of gross domestic product (Centers for Medicare & Medicaid Services, 2026b). Notably, our estimates of hospital markups and trends closely match hospital financial data and regulatory analyses of hospital margins, when available. The estimates also show that emergency care accepts almost all incoming patients (85.9%) and accounts for over half of all available and accessible resources. In contrast, approximately one in two charitable care admission requests is rejected, and only 3% of bed capacity is made available to them. These admission constraints stem from profitability, as fewer than 5% of hospital-quarters operate at capacity in expectation. Appealing to the connection between stable matching mechanisms and price equilibrium (Hatfield and Milgrom, 2005; Bulow and Levin, 2006; Hatfield *et al.*, 2013; Azevedo and Leshno, 2016), we find that constraints on patient selection and the lack of a market coordination mechanism protect Medicaid patients and those who seek charitable care at the expense of Medicare patients and the privately insured. Holding supply fixed, we find that the average privately insured Floridian loses the equivalent of one monthly marketplace premium per year due to these factors, while the average uninsured Floridian gains the equivalent of 1.49 market premiums.

To quantify the full resource externality of charitable care, we simulate the state's equilibrium under pseudo-insurance policies for the uninsured. These policies differ in the compensation they offer hospitals for otherwise charitable care but do not alter other aspects of patients' access, such as hospitals' ability to discriminate in providing scheduled care or patients' networks. We find that across a range of compensation levels for charitable care, the average surplus of all patient types (by insurance source) improves, with over three-quarters of the gains accruing to the insured. At a Medicaid-equivalent policy level, uninsured patient surplus increases by $1,235 per patient per quarter from direct access improvements, while the insured gain between $467 and $617 on average, depending on their insurance. However, we also document losses to those seeking insured scheduled care due to the reallocation of resources. In particular, at the Medicaid-equivalent policy, the insured lose access to 5,358 bed-days of scheduled care, while the uninsured gain access to an additional 103,931 bed-days of scheduled care, and emergency care gains 155,423. The net increase in available capacity is driven by a 7.46-percentage-point decrease in idle capacity.

We find that a policymaker seeking to maximize patient surplus net of fiscal spending would

choose to leave the uninsured without coverage only if it ignored the resource externalities to the insured. Intuitively, the fiscal cost of insuring the uninsured exceeds the surplus it delivers because the government must compensate hospitals for all the care that patients would have received regardless. The cost of this inframarginality is large relative to the surplus gains to the uninsured, as their gains come not from extensive-margin adjustment but from improved sorting across hospitals, shorter travel times, and better match quality. However, after accounting for positive externalities to the insured, the surplus-optimal coverage to the uninsured is only slightly more generous than Medicaid and well below Medicare levels, which helps rationalize existing coverage levels. We show that under a wide array of regulatory preferences, including valuing hospital profits, the optimal policy compensates hospitals at or above Medicaid rates, and that in any such outcome most of the surplus would accrue to non-beneficiaries.

Direct compensation policies, however, are costly and have faced significant political challenges. An alternative of bipartisan interest is to impose minimum requirements on charitable care provided by non-profit hospitals. The challenge, however, is that the benefit of the tax exemption afforded to non-profit hospitals decreases as charitable care delivery increases, because a larger uncompensated pool implies fewer profit dollars are untaxed. This correlation misaligns the benefits and burdens of non-profit status. To address this, proposals have suggested establishing a market that allows hospitals to trade charity care credits to meet minimum requirements (Dranove *et al.*, 2015), an idea echoed in congressional testimony and state legislative deliberations (Levinson, 2023; Georgia House of Representatives, 2019).

We solve for the optimal floor-and-trade policy and evaluate its market impact. To be conservative, we assume that only a 10% tax is at risk for non-complying hospitals and seek a policy that makes it incentive compatible for all existing non-profit hospitals in Florida to retain their status. The resulting trade map lets us visualize the misalignment between non-profit status benefits and burdens and how financial resources must flow to improve outcomes. We find that the optimal floor-and-trade policy achieves 88.3% of the surplus gain from a Medicaid-equivalent pseudo-insurance counterfactual at no additional fiscal cost. However, our results also show how targeting is delicate and necessary. A regulator that must assign a uniform minimum requirement for charitable care across hospitals can attain only 12.3% of the surplus gains of the unconstrained policy. This large drop in policy efficiency reflects heterogeneity in demand, resources, and costs in healthcare markets and highlights a stark tradeoff between targeting and fiscal spending.

Our work bridges the literature on the private and public costs of uninsurance and that on the value of insurance. Existing research documents the effects of coverage on patients' utilization, health, and financial exposure (Finkelstein *et al.*, 2012; Miller *et al.*, 2021; Dobkin *et al.*, 2018), as well as the incidence of uninsured care on hospitals and other parties (Garthwaite *et al.*, 2018; Finkelstein *et al.*, 2018; Brevoort *et al.*, 2020). Our estimates provide complementary evidence on the costs of uninsurance, and our floor-and-trade policy analysis gives an explicit structure to the

role of nonprofit hospitals as insurers, as argued by Garthwaite *et al.* (2018). Our evidence of private patient crowdouts and our theory of resource allocations rationalize prior mixed findings regarding the effect of uncompensated care on insured patient access and contribute to a broader literature on the supply effects of insurance and spillovers across patient groups (Finkelstein, 2007; He *et al.*, 2015; McInerney *et al.*, 2017; Glied and Hong, 2018; Carey *et al.*, 2020). In particular, our theory shows that a higher uninsured burden might increase or decrease insured patient admissions, depending on the relative uninsured load and the comparable profitability of scheduled care. Our welfare analysis adds this resource-allocation channel to analyses of insurance value and incidence (Finkelstein *et al.*, 2019a,b; Hendren, 2021; Hendren and Sprung-Keyser, 2020). As in Finkelstein *et al.* (2019b), we find that beneficiaries' willingness to pay for coverage falls below the gross cost of providing it. Incorporating the benefits to non-recipients, however, changes the sign of the estimated welfare effect of expanding coverage. Similarly, our work complements concurrent work by Ghili *et al.* (2026) on the impact of uncompensated care on insurance take-up. While we focus on the externalities of such care for the provision of care to other patients, they focus on its effect on demand for formal insurance and its consequences for insurance markets. Both articles document externalities that provide an additional rationale for collective intervention in insurance markets with implicit coverage through uncompensated care (Coate, 1995; Mahoney, 2015), and connect to work showing how insurance externalities alter optimal policy (Chandra *et al.*, 2010; Cabral and Mahoney, 2019). Our analysis of how resource externalities affect optimal pseudo-insurance contracts also contributes to the literature on insurance design and the welfare consequences of insurance-market interventions (Zeckhauser, 1970; Einav *et al.*, 2010; Hackmann *et al.*, 2015; Marone and Sabety, 2022; Ghili *et al.*, 2024; Tebaldi, 2025; Vatter, 2025).

This article also contributes to the literature on hospital competition and healthcare production. Our analysis of capacity constraints and patient allocation builds on classical work on reserve capacity and uncertain demand by Joskow (1980) and Gaynor and Anderson (1995). Our model complements the literature on hospital pricing behavior (Gaynor and Vogt, 2003; Gowrisankaran *et al.*, 2015; Ho and Pakes, 2014; Ho and Lee, 2017; Handel and Ho, 2021) by having hospitals optimize over the supply of care rather than its price. The analysis accounts for capacity constraints and overflow costs, which are necessary to explain crowdouts and observed allocation. However, we abstract from potential effects of congestion on treatment quality (Hoe, 2022; Gowrisankaran *et al.*, 2026). Our focus is on how resource scarcity affects costs and allocation across patients and competing hospitals, and although waiting times motivate our interpretation of rationing, we leave them implicit because our data do not directly measure them. Our work is therefore complementary to that on how wait times affect choice, rationing, and patient outcomes (Sivey, 2012; Gruber *et al.*, 2023; Russo, 2024). Hospitals' discretion over admissions for scheduled charitable care also connects our analysis to the provider-selection literature (Newhouse, 1996; Ellis, 1998). Particularly close is Gandhi (2026), who studies selective admissions in nursing homes and the opportunity cost of committing beds to patients with long expected stays. His analysis emphasizes

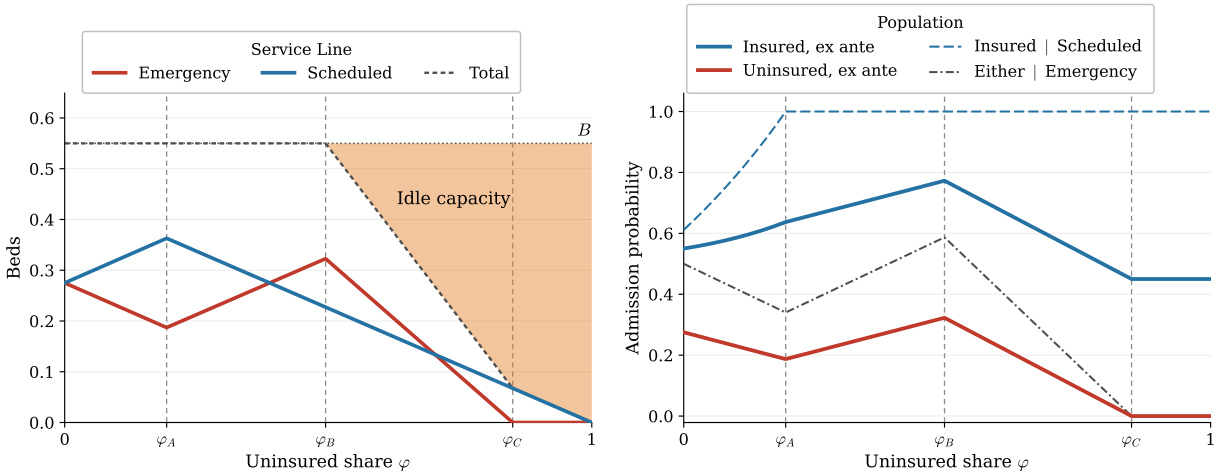
the possibility of displacing more profitable future admissions; ours emphasizes the opportunity cost of allocating shared resources across contemporaneous hospital services.

Finally, our work contributes to the literature on nonprofit hospital behavior and charity-care provision. A longstanding debate concerns whether hospital ownership changes providers' objectives and behavior (Newhouse, 1970; Weisbrod, 1988; Frank and Salkever, 1991; Duggan, 2000; Sloan *et al.*, 2001; Duggan, 2002; Horwitz and Nichols, 2009). Evidence that nonprofit hospitals respond to financial incentives much like for-profit hospitals has motivated a variety of models of firm behavior, some focusing exclusively on admission profits and others incorporating preferences for quality, volume, and social surplus (Gaynor and Vogt, 2003; Gowrisankaran *et al.*, 2015; Ho and Lee, 2017). Our estimates suggest nonprofit hospitals place positive weight on charity-care delivery consistent with models of (impure) altruism (Frank and Salkever, 1991). Our evidence, however, cannot distinguish altruism from an implicit quid pro quo with regulators (Fournier and Campbell, 1997), preservation of tax-exempt status, donor appeal, or political goodwill.

The remainder of the article is organized as follows. Section 2 presents a conceptual framework that highlights the ambiguous effects of resource externalities and how different comparative statics may arise from the same frictions. Section 3 provides further details on the setting and data used in this article. Section 4 presents our main reduced-form evidence, documenting the spillovers from Medicaid, the construction of the instrument for nearby hospitals' uninsured demand, and our key measurements of uninsured cost. Section 5 presents our empirical model, and Section 6 discusses its identification and estimation. Section 7 presents our key structural estimates including markups, resource allocation, and allocative efficiency. Section 8 presents the counterfactual policy analysis quantifying the resource externalities of charitable care and the potential impact of optimal pseudo-insurance and floor-and-trade policies. Finally, Section 9 concludes.

2 Conceptual Framework

We begin by illustrating the key economic forces behind our analysis, explaining how the insured and uninsured interact in the allocation and availability of hospital resources. For simplicity, we abstract here from the intricacies of competition and heterogeneity and consider a profit-maximizing monopolist hospital facing demand from a unit mass of homogeneous patients. A fraction λ of the population requires emergency care, and an independent fraction φ lacks insurance. Each admitted patient occupies one bed, costs the hospital c to serve, and, if insured, yields a revenue of p. Uninsured patients, by contrast, leave care uncompensated. The hospital has B beds available for service and must allocate them between emergency and scheduled care. Emergency care is non-discriminatory, and the hospital must accept patients as long as it has capacity. In contrast, scheduled care is discriminatory and thus excludes the uninsured from care. In this simple environment, the hospital must decide the number of patients to admit to each line subject to the total



(a) Resource Allocation **(b)** Admission Rates

Figure 1: Conceptual Framework – Resource Allocation Under Uninsurance Burden

Notes: These figures show how a monopolist hospital allocates resources between its emergency and scheduled care service lines as a share φ of its population is uninsured. The plots is based on the solution of Equation (1) with $B = 0.55$, $\gamma = 0.55$, $p = 1$, $c = 0.15$, $c^V = 0.55$. Capacity is set below demand to highlight all potentially constrained regions and to align with the observation that most hospitals have capacity below their total local demand.

available population and its capacity constraint:

$$\max_{\substack{0\leq q_E\leq\lambda \\ 0\leq q_S\leq(1-\lambda)(1-\varphi) \\ q_E+q_S\leq B}} \sum_{j\in\{E,S\}} (\pi_j q_j - c^V q_j^2) \tag{1}$$

where q_j is the number of admitted patients to line j, c^V is the additional cost arising from congestion, and π_j is the line's profit as determined by prices, cost, and uninsurance risk. This simple formulation allows analytic solutions and visual exploration, as shown in Figure 1.

Figure 1a illustrates how resources are allocated as the local uninsured share increases. As the rate increases from 0 to φ_A, expected profits from emergency admissions decline relative to those from scheduled care, leading the hospital to reallocate resources. This reallocation leads to fewer emergency admissions among all patients, but more admissions among the insured on the scheduled care line, as shown in Figure 1b. At φ_A, the hospital has exhausted its demand for scheduled care, but its emergency demand is sufficiently populated by the insured that it is still profitable to assign it resources. From φ_A to φ_B, emergency admission rates increase at the same rate as the demand for scheduled insured care decreases. At the same time, an increasing share $\varphi(1-\lambda)$ demands scheduled care without insurance and is left unserved. At φ_B, expected profits from emergency admissions become zero. The hospital gradually shuts down its emergency services, allocating capacity towards the scheduled line where no patient requires it. A growing share of its capacity sits idle until, at φ_C, the emergency department doors are closed, and the

hospital becomes a small clinic for the few insured left.

This illustration highlights the externalities of uninsurance on the insured. At low ($\leq \varphi_A$) and high ($\geq \varphi_B$) levels, increasing uninsurance reduces resource allocation to emergency care, reducing access for all patients. However, at low levels of uninsurance, insured access to scheduled care increases with uninsurance, implying that the externalities of uninsurance depend on the insured's relative need for emergency care and on their incremental value from obtaining emergency or scheduled care at the monopolist hospital relative to the outside option. In contrast, at high levels of uninsurance, insured welfare is systematically reduced with uninsured rates. The figure, however, also reveals a region, (φ_A, φ_B), where a higher uninsured rate leads to higher emergency care admission rates, resulting in positive externalities for the insured. Thus, the figure reveals that uninsurance externalities depend on which of the four regimes dominates market behavior, which in turn depends on the empirical primitives of each setting. Because each region is characterized by distinct responses in admission and resource allocation, empirical evidence on these outcomes can distinguish the governing regime, as we do toward the end of this article.

Ignoring the individual cost of insurance, having access to it is largely beneficial to the population. This is true not only for those who would otherwise go uninsured, but potentially so for those who are insured regardless. As illustrated here, patients inframarginal to changes in insurance coverage are affected by how these policies affect access. One of the central themes of our analysis is how the insured are affected not because we place additional value on their well-being, but rather because it affects the public accounting of the benefits of access to insurance. While this illustration abstracts from additional channels of externalities such as changes in competition or effects on patient sorting, it highlights how resource allocation might change with access to insurance, who might benefit from broader access, and who might prefer to keep hospital care exclusive. The forces governing the impact of insurance coverage depend on hospitals' costs, revenues, and patient preferences and needs, all of which we must quantify, as we do in the remainder of this article. As in this analysis, we abstract throughout from any additional value uninsured patients might derive from risk protection under counterfactual insurance and focus on how such insurance affects payments and access, framing the question as one of regulating payments and access to otherwise charitable care.

3 Data and Setting

We begin by describing the regulatory environment that interacts with the delivery of charitable care and non-profit hospital status, as well as the data we use for our analysis. We focus on Florida for state-specific examples, as it is the setting for our structural analysis. The description of tax benefits and public spending relies on the last available data when possible. However, because of gaps and inconsistent reporting, some numbers below span different years, which we make explicit. For a review of the history of access to insurance in the US, see Gruber (2008), and

for a summary of changes leading to and following the expansion of Medicaid access under the Affordable Care Act, see Gruber and Sommers (2019). Our analysis leverages private multistate systems, a feature of substantial consolidation in the American healthcare market, as reviewed by Gaynor *et al.* (2015) and Richards and Whaley (2026).

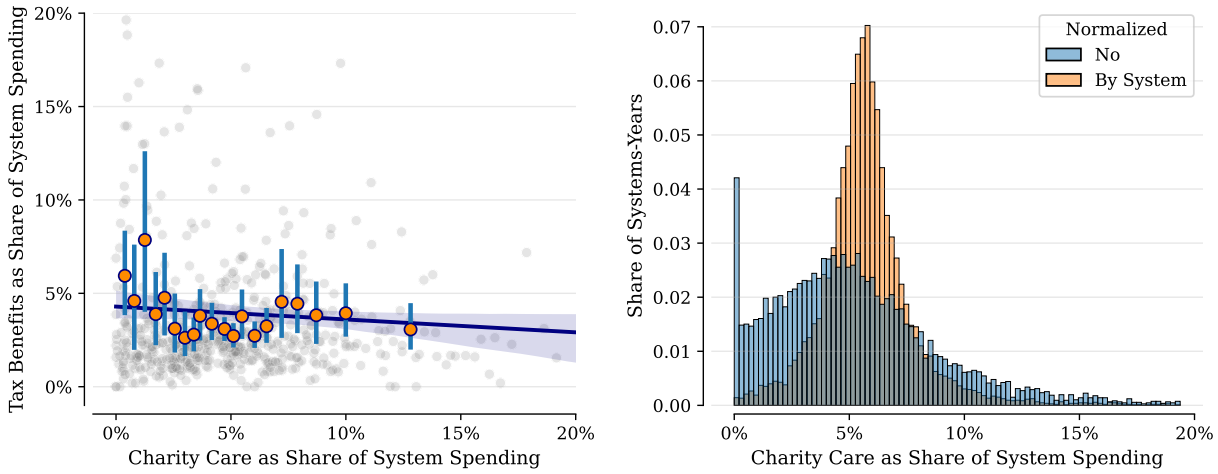
3.1 Setting

Charitable care refers to the provision of services free of charge or at heavily discounted rates to patients who are unable or have limited ability to pay. It constitutes a crucial avenue for care for the 23.5 million uninsured adults aged 18 to 64 in the US (Cohen and Briones, 2026) or the estimated 45 million underinsured, who are unlikely to afford their cost-sharing for any meaningful medical bill (Collins and Gupta, 2024). It is estimated that there are roughly 1.4 million charitable care admissions per year (Agency for Healthcare Research and Quality, 2025), with reported hospital expenditure on charity care being $31.3 billion in 2023 (Centers for Medicare & Medicaid Services, 2026a).[1] The geography of this care delivery is tied to that of insurance, with states that did not expand access to Medicaid following the Affordable Care Act (ACA) having higher uninsurance and charity care rates (Levinson *et al.*, 2025). For example, Florida did not expand Medicaid and in 2024 had an uninsured rate of 15.5% among adults aged 19 to 64, compared with 3.7% in Massachusetts, which did expand access (KFF, 2024). Florida's hospitals also spend 2.1 times more on uncompensated care per capita than those in Massachusetts (Centers for Medicare & Medicaid Services, 2026a). Florida also has 267,000 non-disabled non-elderly adults with incomes too low to qualify for subsidies on the ACA insurance exchanges but too high to qualify for the state's Medicaid pool. In contrast, Medicaid is broadly accessible for legal residents of Massachusetts who do not qualify for subsidies on the exchanges.[2]

Beyond the cost to hospitals, charitable care carries a high public cost through programs that incentivize and compensate hospitals for this care. In particular, hospitals that are given the federal 501(c)(3) non-profit status are expected to provide community benefits, which include the provision of charitable care. In return, these hospitals obtain a variety of benefits, including exemption from federal income taxes and, in most cases, from state income and sales tax as well as local sales and property taxes (Plummer *et al.*, 2024; Godwin *et al.*, 2023). In 2020, non-profit hospital tax benefits are estimated at $28.1 billion, exceeding that year's reported charitable care spending of

[1] Charitable care admission counts are not systematically collected. A standard practice—as we follow in this article—is to classify all patients who at discharge had a fully forgiven bill or registered as self-pay, as charitable (Cram *et al.*, 2010; Popescu *et al.*, 2019; Barrett *et al.*, 2021). We also note that bad debt is distinct from charitable care. Bad debt originates from patients the hospital attempted to collect from but could not. Charitable care, instead, is usually determined before any collection and sometimes before the patient is even admitted.

[2] As of 2026, Florida's income limit for Medicaid qualification for parents in a family of three was 26% of the federal poverty level, while other non-disabled adults had no income-based eligibility pathway. In Massachusetts, the effective income limit is 138% of the federal poverty line. Citizenship, immigration, residency, and other eligibility requirements also apply, so the absence of a statutory coverage gap does not imply universal take-up.



(a) Spending and Tax Benefits

(b) Distribution of Spending

Figure 2: Charitable Care Spending at Non-Profit Hospitals

Notes: These figures display properties of the distribution of charitable care spending for non-profit hospitals, at the system-year level. Figure 2a shows the distribution of tax benefits and charitable care spending for 2019. We computed the tax benefit using available public tax and accounting data, while we derived charitable care and overall system spending from tax returns. The gray dots correspond to the underlying data, while the orange dots correspond to ventiles of the x-axis and their mean estimate, and the line represents the overall sample regression line. Figure 2b shows the distribution of charitable care spending relative to systems' overall spending before and after controlling for a system fixed effect, restricting attention to systems observed for at least three years. Both figures restrict to systems with at least $100 in system-level spending in a year.

$16 billion (Godwin *et al.*, 2023). Non-profit hospitals are required to form clear and public plans for the provision of free and discounted care, have clear guidelines for who qualifies for this care, conduct a Community Health Needs Assessment at least once every three years to determine the community's needs for care, and follow strict procedures for evaluating whether patients qualify for assistance before any attempt to pursue extraordinary collection actions (Internal Revenue Service, 2026a,c). The IRS reviews hospitals' compliance with these rules, with non-compliant hospitals facing public disclosure of failures, a potential excise tax, income taxation, and, in very serious cases, revocation of federal exempt status (Internal Revenue Service, 2026b). Enforcement, however, is gradual and correction-oriented, with over three-quarters of financial assistance policy compliance checks resulting in corrective action (Internal Revenue Service, 2015; Treasury Inspector General for Tax Administration, 2025).[3]

However, viewed as a program to stimulate charitable care provision, the costs and benefits of non-profit status are often misaligned, as shown in Figure 2a. The correlation between a system's benefit from tax exemption and its spending on charitable care is negative, in part because hospitals

[3]Revocations are rare but do occur. For example, in 2017, the IRS revoked a hospital's 501(c)(3) recognition after finding serious failures in meeting program requirements (Internal Revenue Service, 2017). State enforcement also occurs, as in the case of Provena Covenant Medical Center v. Department of Revenue, in which the hospital lost its property-tax exemption, in part, because its charity-care program was inadequately publicized and patients were billed before being considered for assistance (Supreme Court of Illinois, 2010).

that face a high uninsurance burden also have very narrow profit margins. Thus, those who provide the most stand to benefit the least from the tax exemption. Notably, the IRS does not specify a mandatory level of charitable care or community benefits, and the U.S. Government Accountability Office found that between 2014 and 2016, roughly 60 nonprofit systems reported no charitable care at all and nearly 50 reported no community benefits spending (U.S. Government Accountability Office, 2023). However, it is hard to attribute these failures exclusively to hospital decisions. Hospitals are granted nonprofit status in part based on community benefit provision guidelines, but communities are not required to demand such benefits. Thus, some of the hospitals benefiting the most from tax exemption in the figure above are simply located in heavily insured areas.

Nevertheless, non-profit hospitals appear to budget for charitable care and operate around seemingly persistent spending targets. In particular, in their federal tax form filings, non-profit systems are asked whether they budgeted for free and discounted care for the tax year, whether their expenses exceeded the planned budget, and whether they denied care to eligible patients because they exceeded the budget. From 2011 to 2019, 93.3% reported budgeting, 51% reported exceeding it, but only 0.23% denied care because of this. In addition, although the budget is undisclosed, evidence suggests it may be persistent. In particular, 59.2% of the variance in charitable care in the data is explained by a system fixed effect, as shown in Figure 2b. Similar budgeting and target-setting behavior has been documented in hospital response to prior regulation on charitable care (Almond *et al.*, 2011). We record this persistence as a feature of the data and take no stand on its source. This behavior informs the analysis conducted in Section 4, and we provide further evidence supporting this at the end of that section.

Beyond tax exemption, several other state and federal programs compensate hospitals for charitable or uncompensated care. In 2024, Medicare spent $9.28 billion on Disproportionate Share Hospital (DSH) payments, with $5.94 billion directly compensating hospitals for uncompensated care (Medicare Payment Advisory Commission, 2026). In fiscal year 2024, Medicaid spent $12.6 billion on its DSH program, including federal and state funds, to compensate hospitals for Medicaid shortfalls and uninsured patient costs (Medicaid and CHIP Payment and Access Commission, 2026). However, these programs rarely affect hospitals' decisions on the margin because of lags in their effects, uncertainty about budget distribution, and other factors. For example, for Medicare payments, an additional dollar of current uncompensated care affects only the hospital's future share of a fixed national pool, after a reporting and audit lag, and the amount depends on uncompensated-care reporting by every other eligible hospital. Similarly, Medicaid DSH payments are assigned to the state to redistribute, resulting in allocations that do not reflect the burden of charitable care. For example, non-profit hospitals in Florida accounted for 70.9% of all uncompensated care in 2014 but obtained only 18.7% of that year's payments, with the vast majority going to public hospitals (U.S. Government Accountability Office, 2019). Other programs offer compensation to hospitals, but to our knowledge, none provided meaningful marginal com-

pensation during our period of analysis, 2011 to 2019. However, put together, public spending on charitable care is considerable and has often exceeded reported cost.

3.2 Data

Our analysis relies on several proprietary and public sources. The backbone is a discharge-level panel of inpatient admissions constructed from the State Inpatient Databases (SID) of the Healthcare Cost and Utilization Project (HCUP) for 2010–2019 (Agency for Healthcare Research and Quality, 2010–2019). The dataset covers 18 states and over 2,000 hospitals. Importantly, because the dataset is not built from insurance claims data—as is often used in analyses of hospital demand—it captures all patient types, including those without insurance. We follow standard practice by marking patients who, at discharge, had a fully forgiven bill or were registered as self-pay as charitable, which aligns with hospital coding practices for charitable care (Cram *et al.*, 2010; Popescu *et al.*, 2019; Barrett *et al.*, 2021). The SID lets us measure charitable care as realized admissions rather than reported expense, alleviating concerns that reported spending reflects accounting discretion. We require records to have non-missing hospital identifiers, primary payer type, diagnosis-related group (DRG) code, length of stay, and admission year and month or quarter depending on what the state reports. We collect patients' five-digit ZIP codes, admission types, and emergency department statuses. Finally, we retain only states that report both before and after 2014, as our empirical design leverages the expansion of Medicaid access following the Affordable Care Act.

We link these discharges to hospitals using the American Hospital Association (AHA) Annual Survey (American Hospital Association, 2010–2019), which provides hospital characteristics, bed capacity, ownership, and system affiliation, as well as each facility's CMS Certification Number (CCN). System affiliation defines our treatment taxonomy, and bed capacity constrains resources in our structural model. We link this data to the merger panel developed by Brot *et al.* (2024), which documents M&A activity during the time period and allows us to control for confounding changes in hospital ownership. In addition, we link the data to the Healthcare Cost Report Information System (HCRIS) (Centers for Medicare & Medicaid Services, 2010–2019), which provides hospital financial data including reported costs and revenue; because these items are self-reported and include accounting entries unrelated to patient care, our structural analysis estimates hospital costs rather than relying on them. We rely on the American Community Survey for local uninsurance rates (U.S. Census Bureau, 2012–2019).

Because non-profit regulation operates largely at the system level, we complement the discharge data with Schedule H of IRS Form 990 for the same period (Internal Revenue Service, 2019). Schedule H reports the gross and net cost of each community-benefit category, including financial assistance programs, Medicaid and other means-tested government programs, community health improvement, and health professional education. Reporting occurs at the system or state sub-

system level rather than at the facility level. We use the Community Benefit Insight (Ortiz *et al.*, 2022) Hospital Data API to obtain a linkage from each facility's CCN to its Employer Identification Number, which allows us to link the hospital data with the IRS data.[4]

While our descriptive analysis is national, our structural analysis focuses on Florida, a non-expansion state with a large uninsured population: between 2010 and 2019, its average county-level uninsured rate was approximately 23%, 8.2% of inpatient admissions were classified as charitable care, and 10.1% of emergency admissions were charitable care. Because the SID reports provider payments only as charges, we supplement the data with a few additional sources. First, for Medicare and Medicaid admissions, we use CMS's payment formulas, as specified in the Final Rule and Correction Notice files for each fiscal year, and apply Florida's specific Medicaid rules. These computations include payments hospitals receive from compensating programs such as DSH. For private insurers, we use the Florida Health Price Finder files (Kurowski, 2021), produced by the Florida Agency for Health Care Administration (AHCA), which report price percentiles for common procedures at each hospital. We follow longstanding evidence that hospitals negotiate private contracts as a multiple of Medicare rates or a percentage discount off charges, determine the best-fitting hypothesis for the published private rates, and use it to extend prices to unreported procedures. We use these payments only for structural analysis, which requires predicting average private payments for emergency and scheduled care. Thus, the lack of insurer identifiers does not limit our purposes. Finally, we rely on non-profit system tax returns to obtain the average direct payment hospitals receive from patients for charitable care, which averages to 5.68% of Medicaid payments or about $308 per charitable care patient. We use these rates to compute expected direct compensation by charitable care patients in the structural analysis. The supplemental Data Appendix provides further details on our sample construction and compares our imputed private prices with prices reported in the Truven MarketScan dataset.

4 Descriptive Analysis

This section presents our main analysis of how Medicaid expansion generated spillover effects in non-expansion states, which is our primary source of variation for identifying charitable care costs and externalities. Our strategy leverages a channel created by hospitals in non-expansion states that are part of systems also found in expansion states. Figure 3 shows our treatment taxonomy, which we use throughout to distinguish hospitals. In particular, we refer to hospitals in expansion states as *directly treated*. We refer to hospitals in non-expansion states with system siblings in expansion states as *spillover-treated* hospitals. Hospitals in non-expansion states that are found near spillover-treated hospitals and not part of systems with members in expansion states are referred to as *local untreated*. All other hospitals in non-expansion states are simply *untreated*. To

[4]Records that remain unmatched or misaligned, typically because of mergers and acquisitions, are verified and corrected manually using public sources.

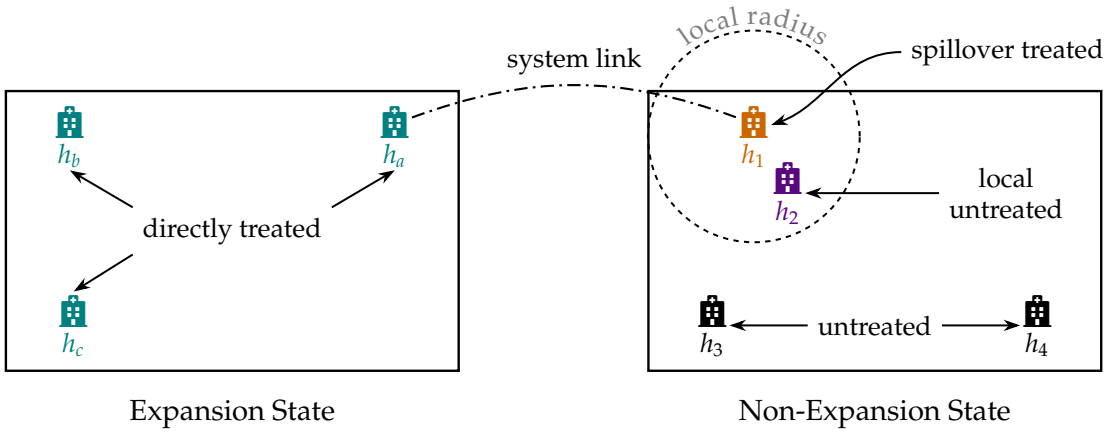


Figure 3: Treatment Taxonomy

Notes: Hospitals in expansion states are referred to as *directly treated*. Hospitals in non-expansion states belonging to systems with a directly treated hospital are *spillover treated*. Hospitals found within a local radius of a spillover-treated hospital are referred to as *local untreated*. We refer to other hospitals in non-expansion states simply as *untreated*.

avoid changes due to patient reallocation, throughout our analysis we consider only changes in demand for each hospital from patients living in the same state.

We begin by showing changes in demand for charitable care in direct- and spillover-treated hospitals. We document patterns of crowd-out of the insured and improved access by the uninsured. Using the spillover of Medicaid expansion into non-expansion states as an exogenous demand shifter for local untreated hospitals, we quantify the opportunity cost of a charitable care admission and the service-cost burden imposed by the uninsured population. We conclude this section with a brief analysis of the motives and mechanisms that might underlie our variation. Our model-guided evidence in Section 5 provides a more structured analysis of these incentives and allows us to quantify the externalities of uninsurance and the incidence of compensating programs.

4.1 Changes in Charitable Care

We begin by examining if and how Medicaid expansions changed the demand and supply of charitable care. We focus on general acute care hospitals with at least 50 beds that do not exit the market during our analysis period and are not acquired within a seven-year window centered on the expansion event. This last filter is particularly strict for for-profit hospitals in non-expansion states, dropping almost all of them from our sample. To keep our analysis consistent, we exclude all for-profit hospitals from the regressions presented in this section. We require treated hospitals—directly or spillover—to be continuously observable in our data for seven years centered at treatment. Directly treated hospitals are treated only once. For spillover hospitals, we focus on the first treatment period, which in most cases corresponds to 2014.[5]

[5]Hospitals found in states that expanded Medicaid access by 2020 are excluded from the potential set of spillover-treated, local-untreated, or untreated set.

We form stacked and balanced event panels for each directly or spillover-treated hospital by randomly matching it to control hospitals.[6] We require control hospitals to be continuously observed, located in non-expansion states, not be part of systems with members in expansion states, and not undergo acquisitions within the same windows as for treatment units. To further avoid potential confounding effects from mergers, we require control units not to be part of the same system as treated hospitals at any point in our sample. Control units are meant to capture trends in demand for, and supply of, charitable care services; however, the extent to which we can account for local changes differs between directly and spillover-treated hospitals. For spillover-treated hospitals, we rely on other not-for-profit hospitals within the same state to account for both national and local changes in demand for charitable care and incentives to provide it. In practice, we match spillover-treated hospitals to a random sample of not-for-profit system hospitals from the same state, which match the treated hospital's Medicare accreditation status and system size category and meet the requirements above.[7] To avoid mechanical inflation in estimated effects due to demand substitution, we exclude control hospitals that are located within 30 miles of treated hospitals. For directly treated hospitals, all local hospitals have experienced Medicaid expansion; thus, we rely on hospitals in other states. We match directly treated hospitals to other not-for-profit system-member hospitals with a bed size similar to the treated hospital and that meet the restrictions above.[8]

Using our stacked panel of events, we conduct regressions of the form:

$$y_{eht} = \alpha \mathbb{1}\{t \geq 0\}\mathbb{1}\{h = h_e^*\} + \gamma_{eh} + \delta_{et} + \varepsilon_{eht} \tag{2}$$

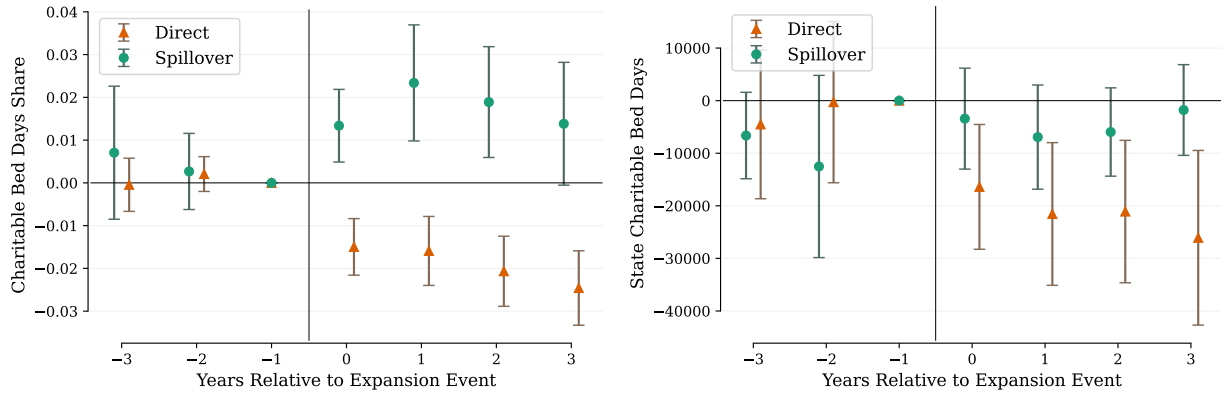
Where h indexes hospitals, t indexes years relative to treatment (i.e., the expansion event), and e indexes treatment events. The outcome of hospital h, t years after the treatment, is denoted y_{eht}, and γ_{eh} and δ_{ht} are event-hospital and event-time fixed effects. Each event has a unique treated hospital, identified by h_e^*, and a balanced set of control units. In addition, we use the standard event-study form, in which we estimate a separate α for each period relative to $t = -1$.

Figure 4a shows the estimated effect of Medicaid expansion on charitable care in directly and spillover-treated hospitals. To make changes comparable across hospitals, we first measure

[6]Our sample and regressions are designed to contend with known challenges in two-way fixed-effects analysis (Callaway and Sant'Anna, 2021; Baker *et al.*, 2022). Following Wing *et al.* (2024), we weight our regression such that each treated unit gets a weight of 1, and each control unit matched to an event e obtains a weight $1/N_e$, where N_e is the number of matched controls. To account for the stacked structure, we estimate all regressions using two-way clustered standard errors, clustering at the event and hospital levels.

[7]We define system size categories by the median of total system bed counts within each state in 2014.

[8]Size similarity is based on 50-tiles of bed size in 2014. We use this stricter match for direct treatment since the pool of potential controls is substantially larger than for spillover treatments, allowing us to match hospitals that are more similar to each other.



(a) Effects on the treated

(b) Aggregate effects on local demand

Figure 4: Direct and Spillover Effects From Medicaid Expansion on Charitable Care

Notes: Figure 4a shows the estimated effect of Medicaid expansion on hospitals in expansion states (triangles), and those in non-expansion states but that are linked to expansion states via system membership (circles). These estimates match the event-study variant of the stacked event-study panel regression (2), with standard errors clustered at the event and hospital levels. Figure 4b shows the estimated effect of expansion on the overall volume of charitable care in expansion states (triangles) and non-expansion states that are subject to spillover effects (circles). These estimates match the event-study variant of the state-year-level regression (3), with heteroskedasticity-robust standard errors. In both figures, we include only patients residing in the same state as the affected hospitals in the demand counts. Error bars represent 95% confidence intervals.

charitable care volume as a share of total bed-days served at each hospital.[9] The figure shows that directly treated hospitals experience a sharp decline following Medicaid expansion: a 1.9 percentage-point decrease in the relative volume of charitable care, or a 43.2% decrease relative to their pre-expansion baseline. In contrast, spillover-treated hospitals experience a 1.4 percentage-point increase in the relative volume of charitable care, amounting to a 25.5% increase relative to their pre-expansion baseline. Columns 1 and 5 of Table 1 show the supporting aggregate effects. A natural concern is that the effect stems not from the numerator (the number of charitable care admissions) but rather from the denominator (the total volume of admissions). Columns 2 and 6 of Table 1 show the effects measured in bed-day counts, with comparable magnitudes, and Appendix Figure A.1a shows the corresponding event-study plot. Both exhibit clear opposing changes between the two sets of treated hospitals.

Despite the opposing directions and comparable magnitudes reported in Figure 4a, the aggregate effects of the two treatments are strikingly different. First, as we report in column 3 of Table 1, directly treated hospitals experience a more-than-offsetting increase in demand from Medicaid patients, as expected given the expanded access to public low-income insurance. In contrast, column 7 of Table 1 shows no such effect for spillover hospitals. Appendix Figure A.1b shows the corresponding event-study plot to these observations. Second, consistent with uninsured individuals accessing Medicaid coverage in expansion states, aggregate charitable care admissions

[9]We weight admissions volume by the number of bed-days used, as we are interested in understanding how one patient crowds out resources for other patients. Because patients' length of stay varies with medical needs, which correlate with insurance status, comparing admission counts may misrepresent how constrained hospital resources are.

Table 1: Medicaid Expansion Effects on Charitable Care

	Panel A: Direct Treatment				Panel B: Spillover Treatment			
	Charitable Bed Days		Medicaid	State Charitable	Charitable Bed Days		Medicaid	State Charitable
	Share	Number	Bed Days	Bed Days	Share	Number	Bed Days	Bed Days
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Expansion	-0.019***	-1543.808***	3801.132***	-19681.283***	0.014***	1252.872***	-158.309	1859.326
Effect (α)	(0.003)	(316.991)	(836.559)	(3729.161)	(0.004)	(467.916)	(710.476)	(3414.858)
Mean Dep.	0.044	2748.671	14008.759	37229.152	0.055	3938.259	13709.186	77643.906
N	2081	2081	2081	959	3065	3065	3065	763
R^2	0.914	0.944	0.994	0.990	0.908	0.962	0.995	0.996

Notes: All measures include only patients residing in the hospital's state. "Charitable Bed Days Share" is the total length of stay of charitable care patients relative to the total length of stay of all admitted patients at the hospital. "Charitable Bed Days Number" presents the total length of stay of charitable care patients. "Medicaid Bed Days" corresponds to the total length of stay of Medicaid patients. "State Charitable Bed Days" corresponds to the total length of stay of all charitable care patients in the state. The first three columns of each panel are estimates of Equation (2), with two-way clustered standard errors at the event and hospital levels. The state-level outcomes are estimates from Equation (3), with heteroskedasticity-robust standard errors. In the latter, we keep only one observation per state-year. "Mean Dep." shows the mean of the dependent variable for treated units in the year prior to expansion. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

declined sharply in expansion states, but did not change in non-expansion states. We examine this by running a simple first-difference regression at the state-year level of the form:

$$y_{st} = \alpha \mathbb{1}\{t(s) \geq 0\} + \zeta_s + \varepsilon_{st} \tag{3}$$

where y_{st} is a state-year outcome, $t(s)$ is the year in which the state first experiences direct or spillover treatment, and ζ_s is a state fixed effect. Figure 4b shows the estimated effects per year, while columns 4 and 8 show the supporting aggregate estimate.

Appendix Table A.1 provides further evidence of opposing changes in admissions: a decline in charitable care in expansion states and its stability in non-expansion states. This last finding is particularly important, as it implies that the documented increase in charitable care at spillover-treated hospitals must stem from a reallocation of charitable care patients across hospitals within the state. This largely rules out the possibility that spillover-treated hospitals increase charitable care by offering uncompensated care to patients who would not have received it previously, or that the additional charitable care is for patients who would not have received it prior to the treatment. Any such change would necessarily increase the aggregate level of charitable care in spillover states. Thus, Medicaid expansion spillovers shift the uninsured demand curve faced by some of the untreated hospitals in non-expansion states. This demand shift is the key variation we use to identify the burden charitable care imposes on hospitals and the insured.

4.2 Access and Crowd-outs

Having established a spillover channel from expansion to non-expansion states, we now turn our attention to how changes in the supply of charitable care affected access and crowd-outs. For the remainder of the section, we focus on non-expansion states to understand how Medicaid spillovers affected care at spillover-treated hospitals and how the ensuing reallocation of charitable care demand affected other local untreated hospitals.

We first ask to what degree the expansion of charitable care services at spillover-treated hospitals crowded out other types of care. Following the regression design of Equation (2), Panel A of Table 2 shows that overall admissions (in bed-days) did not change materially at spillover-treated hospitals. While statistically insignificant, the estimates suggest a slight increase in emergency care admissions and a decrease in non-emergency care admissions.[10] Overall, these estimates suggest that the increase in charitable care admissions at spillover-treated hospitals was due to a substitution of patients rather than an expansion of services.

To further examine this substitution, Panel B of Table 2 studies the composition of insurance sources within the hospital. The first two columns show that the overall share of admissions from Medicare and Medicaid patients (government programs) did not change materially with treatment, whereas the share from privately insured patients was reduced. Column 3 shows that most additional charitable care admissions came from patients seeking emergency care, while columns 4 and 5 show a corresponding crowd-out of privately insured patients from emergency care and null effects on Medicare and Medicaid patients. Given that it is unlikely that Medicaid expansions affected demand among the insured in non-expansion states, and that the insured are uniformly more profitable to serve than charitable care patients, the substitution observed among patients seeking emergency care in spillover-treated hospitals suggests that hospitals are meaningfully resource-constrained.[11] This constraint is also consistent with the relative admitted volume being constant at treated hospitals and resources having to be transferred from non-emergency to emergency care to cover the additional charitable care demand. As further shown in columns 6 to 8 of panel B, non-emergency charitable care did not materially change, while private admissions fell with a partially offsetting increase in admissions by those on Medicare and Medicaid. The overall protection of admissions by Medicare and Medicaid patients is consistent with non-profit hospitals seeking to maintain good standing in their reported community benefits, which include these services.

Given the effects at spillover-treated hospitals, nearby hospitals should have experienced mean-

[10]The emergency status of an admission is classified by HCUP and determined by the patient's need for immediate care and includes more than a visit to the emergency department. Conceptually, we consider all patients in need of emergency care in hospitals with an emergency department (ED) to be protected by EMTALA, since even those not routed through the ED could re-enter the hospital via the ED to obtain regulatory protection.

[11]Otherwise, a hospital expanding charitable care would retain its profitable privately insured patients.

Table 2: Spillover Effects on the Treated: Crowd-outs

	Panel A: Spillover-Treated Hospitals Admissions (bed days)		
	Total Admitted (1)	Emergency Admissions (2)	Non-Emergency Admissions (3)
Expansion Effect (α)	486.549 (3654.607)	698.406 (2882.950)	-211.857 (2158.025)
N 3065	Mean Dep. 72637.6 R^2 0.996	Mean Dep. 44389.4 R^2 0.992	Mean Dep. 28248.2 R^2 0.987

	Panel B: Spillover-Treated Hospitals Insurance composition (share of bed days)							
	Total Admitted		Emergency Admissions			Non-Emergency Admissions		
	Gov. (1)	Private (2)	Charitable (3)	Private (4)	Gov. (5)	Charitable (6)	Private (7)	Gov. (8)
Expansion Effect (α)	0.007 (0.014)	-0.026*** (0.009)	0.013** (0.006)	-0.027** (0.012)	0.008 (0.016)	0.003 (0.004)	-0.032*** (0.008)	0.023* (0.013)
Mean Dep.	0.737	0.181	0.076	0.139	0.761	0.030	0.248	0.691
N	3065	3065	3060	3060	3060	3065	3065	3065
R^2	0.876	0.914	0.880	0.835	0.793	0.825	0.926	0.907

	Panel C: Local Untreated Hospitals Admissions (bed days)			
	Charitable (1)	Gov. (2)	Private (3)	Total (4)
Expansion Effect (α)	-1009.057** (424.829)	1173.584 (1604.996)	1681.869* (909.450)	2807.092 (1963.974)
N 1156	Mean Dep. 5418.0 R^2 0.944	Mean Dep. 50003.1 R^2 0.988	Mean Dep. 17910.6 R^2 0.974	Mean Dep. 75735.6 R^2 0.992

Notes: This table presents the effects of Medicaid expansion in non-expansion states. Panels A and B show effects on spillover-treated hospitals, as estimated from Equation (2) with standard errors clustered at the event and hospital levels (two-way). Panel C shows the effect on the local untreated, defined as hospitals within 15 miles of a spillover-treated hospital but more than 5 miles away. These are estimates of Equation (4), with standard errors clustered at the HRR level. Appendix Table A.2 shows a variant with Panel A in logs, Panel B measured in admitted bed days, and Panel C without the 5-mile local filter. "Gov." represents admissions from Medicare and Medicaid enrollees. All demand metrics incorporate only demand from patients residing in the same state as the affected hospitals. "Mean Dep." is the mean dependent variable measured the year before treatment. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

ingful changes in their patient composition. To study this effect, we examine the demand for care at untreated hospitals located within fifteen miles of a spillover-treated hospital. To be consistent with our analysis and to mitigate confounding effects from mergers, we adopt the same data requirements imposed on treated hospitals in our stacked event panel.[12] In addition, to avoid mechanical effects due to patient overflow and local transfer agreements, we exclude hospitals found within five miles of spillover-treated hospitals.[13] Using this sample, we estimate the model:

$$y_{ht} = \alpha T_{ht} + \gamma_h + \delta_{s(h)t} + \epsilon_{ht} \tag{4}$$

[12]Specifically, we require hospitals to be general care hospitals, have at least 50 beds, not undergo an acquisition within a seven-year window centered on treatment, be not-for-profit, and not exit by the end of our sample.

[13]See Appendix Table A.2 for the same analysis without the five-mile filter and with the additional requirement of hospitals being system hospitals, to increase comparability with spillover-treatment hospitals.

Where y_{ht} is the outcome, T_{ht} indicates whether the hospital was exposed locally to a spillover effect at a period prior to or equal to t, γ_h is a hospital fixed effect, and $\delta_{s(h)t}$ is a state-year fixed effect. This last fixed effect implies that α captures how the local untreated hospitals' demand changed relative to other untreated hospitals from the same state, following the taxonomy of Figure 3.

The estimates shown in Panel C of Table 2 indicate that local untreated hospitals experience changes opposite to those measured at spillover-treated hospitals. While the average spillover-treated hospital increases charitable care admissions by 1,253 bed-days, the average local untreated hospital decreases them by approximately 1,009. Similarly, while spillover-treated hospitals see a decrease in private admissions, local untreated hospitals see an increase. In total, the average local untreated hospital loses 18.6% of its charitable care burden, and increases its private admissions by approximately 9.4%.

A key feature of this change in patient composition is that it appears to be driven by changes in the supply of care at spillover-treated hospitals. To further support this hypothesis, in Appendix A.2 we document how these changes affected the quality of care patients accessed and the distance they traveled for care. There, we show that charitable care patients experience an overall decrease in the average distance traveled from home to hospital and that, relative to the crowded-out privately insured, they are admitted to hospitals with better performance on a variety of mortality outcomes. This suggests that charitable care patients might have preferred spillover-treated hospitals before treatment but were limited in their ability to be admitted there. Once spillover-treated hospitals changed their care supply and relaxed admission constraints, more patients were admitted to preferred hospitals, crowding out the insured.

In addition, Appendix Table A.3 documents how the composition of admitted medical conditions among charitable care patients changed at spillover-treated hospitals. The results show no meaningful changes in admissions for medical conditions that are hard to redirect, such as circulatory problems or pregnancies. Instead, the increases are driven by additional care for patients with HIV, cancer, and immune diseases, among others. This supports the hypothesis that these patients were denied access to spillover-treated hospital care due to capacity constraints before spillover treatments and were redirected to alternative local options.

4.3 Costs and Revenue

A central question in restructuring charitable care incentives is how uncompensated care affects hospitals' costs and revenues. Relevant costs include both the cost of serving the uninsured and the opportunity cost of any profit lost from crowding out insured patients. The evidence we have developed so far suggests that spillovers from Medicaid expansions shifted demand in non-expansion states, altering the demand for charitable care at local untreated hospitals in ways plausibly exogenous to changes in their costs or production technology. We can thus use these changes to estimate the impact of uninsured care on hospitals' bottom lines, using cost and revenue

data reported by hospitals in the HCRIS dataset.

We begin by examining how exposure to spillover treatment affected the local untreated, by following the design of Equation (4). Panel A of Table 3 shows that, consistent with the local untreated serving fewer charitable care and more privately insured patients, their profits increased by about 140 million, or 25.1% per admitted patient. The table shows that revenue and costs increased overall across these hospitals, but the reported costs of uncompensated and charitable care did not vary significantly. Two plausible reasons explain why uncompensated or charity care did not change despite the sharp decline in charity admissions: first, hospitals claim only a fraction of all care as charitable or uncompensated. Second, because self-reported revenues and costs are not entirely reliable, are subject to reporting errors, or include accounting cost items such as interest paid on debt. We can address the first issue by isolating the impact of charitable care admissions on reported costs, as we attempt in this section. The second is arguably irreparable. Because of this second concern, our structural analysis below will estimate hospital costs rather than rely on this data. Interestingly, as we show in Section 7.2, some of our key cost estimates from the structural work are remarkably close to the estimates we report here.

Given this evidence, we want to understand how charitable care changes hospitals' costs and revenue. Intuitively, this can be examined via the regression:

$$y_{ht} = \beta U_{ht} + \gamma_h + \epsilon_{ht} \tag{5}$$

where y_{ht} is a cost or revenue outcome, U_{ht} is the number of charity admissions the hospital has in period t, and γ_h is a hospital-fixed effect that accounts for the hospital's baseline. The main challenge is that changes in a hospital's uninsured demand likely correlate with changes in the local population, which may affect the hospital's revenue or costs. Panel B of Table 3 presents the OLS estimates of this equation, highlighting this endogeneity. For example, higher demand for charitable care increases inpatient revenue, most likely through a correlation between changes in demand for hospital care among privately insured and uninsured patients.

To address this endogeneity concern, we build an instrument based on the exposure of the local untreated to spillovers from Medicaid expansions, in the spirit of a shift-share instrument. In particular, we measure the extent to which each spillover-treated hospital is exposed to Medicaid expansions by its owning system's total spending on charitable care in expansion states in 2013. As we show in Appendix Table A.4, this measure is positively associated with changes in charitable care at the spillover-treated hospitals. Intuitively, systems that spent more on charitable care in expansion states before the policy lost more charitable care demand and thus faced a stronger push to increase spending in non-expansion states. We use this exposure as a *shift*, and distribute a *share* of it among all hospitals local to the spillover-treated hospital. To account for how distance affects substitution among hospitals, we assign each local untreated hospital a share proportional to the inverse of its distance to the spillover-treated hospital. Finally, the local exposure of a local

Table 3: Cost of Charitable Care at Local Untreated Hospitals

	Profit		Revenue	Costs		
	Total	Per admit (thousands)	Total	Total	Uncompensated Care	Charity Care
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Spillover						
Expansion Effect (α)	140.028***	3.612**	182.994***	42.965***	2.090	2.956
	(31.565)	(1.674)	(37.934)	(9.397)	(5.652)	(5.429)
Mean Dep.	334.159	14.387	623.769	289.610	32.780	21.127
N	3319	3319	3319	3319	3319	3319
R^2	0.947	0.936	0.965	0.982	0.852	0.815
Panel B: OLS						
Charitable Care Admissions (β)	0.141*	0.001	0.171**	0.031*	0.024**	0.011*
	(0.073)	(0.001)	(0.087)	(0.017)	(0.012)	(0.006)
Mean Dep.	334.159	14.387	623.769	289.610	32.780	21.127
N	981	981	981	981	981	981
R^2	0.045	0.006	0.046	0.029	0.033	0.008
Panel C: IV						
Charitable Care Admissions (β)	-0.262***	-0.010***	-0.295***	-0.033**	0.016**	0.014**
	(0.074)	(0.002)	(0.086)	(0.013)	(0.007)	(0.007)
Mean Dep.	334.159	14.387	623.769	289.610	32.780	21.127
N	981	981	981	981	981	981
R^2	-0.323	-0.331	-0.297	-0.094	0.029	0.008

Notes: This table presents the estimated cost of service for charitable care patients on the hospital's costs and revenue. Panel A shows the spillover effect, as in Equation (4), with standard errors clustered at the HRR. Panel B shows the estimates from Equation (5), estimated via OLS, while Panel C shows the two-stage least squares estimates, where the instrument corresponds to the hospital's exposure to Medicaid spillovers. Regressions in Panels B and C include a hospital fixed effect and use heteroskedasticity-robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

untreated hospital is defined as the sum of its share of each spillover exposure.[14]

Panel C of Table 3 shows the estimates when instrumenting U_{ht} with the local exposure metric.[15] We note that the instrument captures the effect of an additional charitable care admission at the local untreated hospitals, including any effect on the crowd-out of private insured patients. These estimates suggest that an additional charitable care admission reduces hospitals' profits by about $262,000, with $295,000 from lost revenue from displaced patients and an offsetting $33,000 decrease in costs. The large revenue decline is consistent with hospitals serving more charitable care patients and fewer patients overall. This suggests that while the effect on overall admissions

[14]Formally, local exposure is defined as $LE_{ht} = \sum_{\tau \leq t} \sum_{e \in \mathcal{E}_\tau} \mathbb{1}\{5 < dist(h, h(e)) \leq 15\} \frac{1/dist(', h(e))}{\sum_{h' \in \mathcal{H}} \mathbb{1}\{5 < dist(h', h(e)) \leq 15\} 1/dist(h', h(e))} E_{eh}$ where $\mathcal{E}_\tau$ is the set of expansion events occurring at period τ, $h(e)$ is the spillover-treated hospital associated with the event, and $dist(h, h(e))$ is the distance between hospital h and the spillover treated hospital.

[15]The first stage of the regression is shown in column four of Appendix Table A.4. In estimation, we also interact the instrument with indicators for whether the hospital is publicly or privately owned.

shown in column four of Panel C in Table 2 was marginally insignificant, its magnitude was large enough to generate a significant effect on hospitals' overall revenue. Similarly, the decrease in costs shown in column four of Table 3 is consistent with fewer patients being served and with privately insured patients being both more profitable and more complex.

The estimates suggest that for each additional charitable care patient, uncompensated care costs increase by $15,786.1 and charitable care costs by $13,819.6. Taking the first of these numbers and assuming that charitable care demand is driven by the uninsured, we can compute the average uncompensated cost per uninsured patient to have been $1,177.8 between 2011 and 2019.[16] For comparison, Garthwaite *et al.* (2018) estimate a cost of $830.5 per uninsured in 2011-adjusted dollars based on data from 1984 to 2011 and using hospital closures as a source of identifying variation. Converting to our 2023-adjusted dollars yields a remarkably similar estimate to ours: $1,125 dollars per uninsured.

The estimates highlight several important findings. First, consistent with Garthwaite *et al.* (2018), we find that hospitals appear unable to shift the meaningful burden of charitable care to the insured. Otherwise, we would expect insurance profit per admission to remain stable, with insured admission profit increasing to offset higher charitable care costs. Second, we find that the opportunity cost of charitable care is sizable and likely driven by a combination of lost profit from crowded-out insured patients and from an overall decrease in admitted volume. As we note in Section 5, it is natural to expect hospitals to shift resources away from emergency care and potentially leave them idle when uninsured demand increases, due to the risk of underpayment. If this is happening, it would imply that limited access to insurance can restrict insured individuals' overall access to care. Finally, the estimates suggest that the cost of uninsurance is sizable—with an estimated 29.3 million uninsured individuals in the US in 2019 (Keisler-Starkey and Bunch, 2020), the expected cost of their care is approximately $42.3 billion. For comparison, the realized total reported uncompensated care cost in 2019 was $40.5 billion (Medicaid and CHIP Payment and Access Commission, 2023), indicating that our variation fairly accurately predicts uncompensated care costs.

4.4 System Targets and Redistribution

The spillover patterns presented above raise the natural question of why and how systems achieve these changes. The evidence shows changes in admissions from patients residing in the same state as the affected hospitals; hence, the effects cannot be due to patient steering across state lines. Moreover, because our main evidence comes from discharge data and aggregate uninsured demand is stable at the non-expansion state level, the changes are unlikely to reflect relabeling

[16]The cost per uninsured is computed by taking the estimated cost of uncompensated care (which includes charity and bed debts) and multiplying it by the ratio of charitable care admissions to uninsured. Because our charitable care admissions do not cover all states, we use Florida as a representative large state and use American Community Survey data to count the uninsured population in the state.

of patients in either cost accounting or patient discharge. In addition, since the control group in our main analysis comprises other not-for-profit hospitals in non-expansion states, the effects are unlikely to be driven by local changes in hospitals' incentives to provide charitable care or by local policies unknown to us. Therefore, the main standing hypothesis is that treated hospitals simply expanded the resources available to charitable care. One plausible way hospitals can achieve this is by increasing the share of operating room and post-surgical capacity and resources reserved for emergency admissions, thereby limiting surgeons' ability to book these resources in advance for scheduled care. This mechanism requires unmet demand for charitable care at spillover-treated hospitals, allowing those hospitals to increase charitable care by increasing supply. The evidence on distance and quality discussed above supports this hypothesis.

With a potential theory for how spillover-treated hospitals achieve these changes, we now ask why. We begin by revisiting the hypothesis that hospitals report stable charitable care targets to the IRS. As noted in Section 3, spending on charitable care is highly stable over time within systems. If the expansions in care we observe in spillover hospitals are an attempt to retain these targets, then we should expect a few key patterns to hold. First, we would expect only systems whose spending is meaningfully affected and can stabilize it through supply expansions to change. Intuitively, a system whose hospitals are primarily located in expansion states has a limited margin to offset a reduction in charitable care demand by expanding services in non-expansion states. By contrast, a system found almost entirely in a non-expansion state experiences little impact from Medicaid expansions and thus little incentive to alter its operations. Panel A of Appendix Table A.5 shows that systems in the middle of the exposure distribution expand their services the most in non-expansion states. Second, we would expect spillover systems—which can respond to Medicaid expansions by expanding charitable care in non-expansion states—to retain more of their charitable care spending than systems found only in expansion states. Panel B of Appendix Table A.5 shows that this is the case.

Although the evidence above is consistent with an attempt to maintain stable targets, it does not preclude other motives. Notably, in Appendix Table A.6, we show that grants and donations respond positively to system-level charitable care spending. While these grants and donations are insufficient to cover the additional spending accrued by hospital systems, they may be especially valuable for hospitals at a moment of financial constraint and can capture overall public goodwill toward the system. In addition, the 340B program allows certain private nonprofit hospitals to access generous drug subsidies if they maintain a contract with a state or local government to provide care to low-income individuals who are otherwise ineligible for Medicare or Medicaid. These contracts, however, are not public, and it appears unlikely that they would allow hospitals to meet contract terms by providing more care in other states.

Thus, systems may be reallocating resources for a variety of reasons. Some might be doing so to maintain good standing with the IRS—although, to our knowledge, there has been almost no

enforcement of nonprofit status at the federal level—and some might be doing so to please donors or local governments. We note, however, that our main analysis is largely agnostic about the reason behind the change in operations at spillover-treated hospitals, as our primary identifying variation comes from the induced reallocation of demand to the local untreated hospitals.

4.5 Discussion

Our evidence indicates that non-profit hospitals spanning multiple states created a channel for Medicaid expansions to spill over to states that did not expand coverage. These spillover effects changed the behavior of the few hospitals forming this link, but created demand changes that rippled across broader regions. This shift produced variation that is particularly amenable to studying uninsured care costs and regulation. Unlike variation stemming from changes in coverage—such as the direct effects of Medicaid expansions—it speaks to the costs of uninsured care in areas where this care is still prevalent. Similarly, unlike variation produced by hospital closures or broader structural changes, it is more likely to reflect the costs hospitals incur on a regular basis. However, without further structure, our ability to leverage this variation to answer key questions about the externalities of uninsurance or the incidence of policies covering the uninsured is limited. The remainder of this article aims to close this gap by augmenting the variation with a theoretical framework that would afford us detailed quantification and counterfactual simulations, at the cost of stronger assumptions about hospital and patient behavior.

5 Model

In this section, we develop a framework to study the key mechanisms underlying the evidence presented above. We center our model on how hospitals allocate finite resources between the service lines that deliver the variety of procedures patients require. A service line, such as general emergency care or scheduled oncological surgery, is an organization of services offered by a hospital that shares resources and planning. We place particular attention to the finiteness of resources available to each hospital, the uncertainty of demand, and the potential for one patient to crowd out another. Our model is thus geared towards the questions we seek to answer and is parsimonious in various ways, which we discuss in detail at the end of this section. The model is not meant to explain why spillovers occur, but rather to capture hospitals' reactions to changing demand compositions and incentives, as experienced by the local-untreated in Section 4.

We model the interaction between hospitals and patients as the subgame-perfect Nash equilibrium of a two-stage game. First, hospitals determine how to allocate resources between their service lines and the rules governing patient admission. Second, patients fall ill stochastically and seek care at hospitals based on their preferences, insurance coverage, and each hospital's ability to admit them at the time of illness. These interactions occur at regular intervals (e.g., every quarter) and are independent across time and markets. Thus, we present the model below for a single

market and period, omitting time t and market m subscripts. We will reintroduce these when discussing identification and estimation in the following section. Throughout, we use bold lowercase letters to denote vectors and bold uppercase letters for matrices. We denote sets by calligraphic uppercase letters and their sizes by the same non-calligraphic letter.[17]

5.1 Demand

Each individual i in the market belongs to a population segment $s \in \mathcal{S}$, which indexes their demographic attributes and insurance status. Members of each segment fall ill with a diagnosis $d \in \mathcal{D}$ according to a Poisson process, which is independent across segments and diagnoses.[18] Conditional on falling ill, patient i seeks care among the set of hospitals available in her network $\mathcal{H}_i$. The patient's indirect utility of care at hospital $h \in \mathcal{H}_i$ is given by $v_{idh} + \epsilon_{idh}$, where v_{idh} is the systematic part of her utility and ϵ_{idh} is an idiosyncratic preference shock. In addition, the patient can seek care outside the market and obtain indirect utility of ϵ_{id0}. All idiosyncratic shocks are drawn independently from a type-1 extreme value distribution.

Independent of the patient's preferences, hospitals might reject her admission request due to capacity constraints. We denote by $\tilde{\boldsymbol{a}}_{id} \in \{0,1\}^{H_i}$ the realized acceptance vector indicating whether the patient could be admitted to each hospital at the time of her illness, and assume that the outside option always has excess capacity. Integrating over the distribution of idiosyncratic shocks and the acceptance vector, patient i is admitted to hospital $h \in \mathcal{H}_i$ conditional on illness with probability

$$\sigma_{idh} = \mathbb{E}\left[\frac{\tilde{a}_{idh} e^{v_{idh}}}{1 + \sum_{h' \in \mathcal{H}_i} \tilde{a}_{idh'} e^{v_{idh'}}}\right] = a_{idh} \mathbb{E}\left[\frac{e^{v_{idh}}}{1 + e^{v_{idh}} + \sum_{h' \in \mathcal{H}_i \setminus \{h\}} \tilde{a}_{idh'} e^{v_{idh'}}} \middle| \tilde{a}_{idh} = 1\right] \equiv a_{idh} R_{idh}, \tag{6}$$

where the first expectation is over the distribution of $\tilde{\boldsymbol{a}}_{id}$, while in the second we have expanded over the two possible realizations of $\tilde{a}_{idh}$, which takes a value of 1 with probability a_{idh}. We denote by R_{idh} the likelihood that the patient requests admission at hospital h. The choice probability of the outside option is similar, with the numerator in the expectation replaced with 1.

The structure above can be microfounded in a simple costless search model. In it, when the patient falls ill and draws her stochastic preferences, she and her doctor combine their current knowledge of hospital states and contact everyone about whom they lack information. Assuming that the cost of this search is negligible relative to differences in the value of care, the patient will obtain the current state of all admitting hospitals and choose her preferred one among those available. We can thus consider the patient to require care during the period in which she falls ill, and rejections to be the product of hospitals offering to schedule care at an unacceptable delay.

[17] For example, x is a scalar, $\boldsymbol{x}$ is a vector, and $\boldsymbol{X}$ is a matrix. If $\mathcal{X}$ is a set, then X is its cardinality.

[18] The independence assumption is ubiquitous in structural models of hospital demand, with exceptions in economic epidemiology and infectious disease models (Chan *et al.*, 2016; Greenwood *et al.*, 2019). Conditional on independence, the Poisson assumption is mild, since hospitalization is a rare event in a large population. See Appendix A.3.2 for a discussion of this assumption.

Thus, our model above makes two key assumptions: first, that patients cannot postpone their care beyond the current period; second, that preferences are independent of the state of the hospital, conditional on being ill in a given period.[19]

5.2 Supply

Each hospital h is characterized by the set of service lines it owns, $\mathcal{J}_h$, and its total resources, indexed by the total available bed-days B_h. The hospital's problem has two stages: resource allocation and admission planning. In the first stage, the hospital decides the resources to allocate among its lines to maximize its total profits:

$$\max_{\boldsymbol{b}_h \geq 0} \sum_{j \in \mathcal{J}_h} \mathbb{E}[\pi_{hj}(b_{hj})] \qquad \text{s.t.} \quad \sum_{j \in \mathcal{J}_h} b_{hj} \leq B_h \tag{7}$$

where $\mathbb{E}[\pi_{hj}(b_{hj})]$ is the subjective profit of line j given resources b_{hj} and integrating over its demand. Profits are subjective, as they may include preferences for consumer surplus.

The capacity allocation problem is well studied in hospital operations (Green, 2005; Hulshof *et al.*, 2012). In particular, hospitals must determine how much capacity to allow surgeons to book in advance and how much to reserve for patients arriving for emergency care (Gerchak *et al.*, 1996; Jung *et al.*, 2019). Total planning capacity at hospitals is often rigid, as adding permanent beds requires licensing, certifications, and substantial ancillary investment in bedside resources (e.g., monitors). Moreover, in some states, hospitals must justify the public need for additional capacity and obtain regulatory approval before adding beds (Virginia Department of Health, 2026; National Conference of State Legislatures, 2025). Therefore, in the short run, a hospital's response to changes in demand and profits across services and demographics is likely to begin with capacity reallocation rather than capacity adjustment.

Given the allocated resources, the hospital must determine whom to admit. In practice, hospitals manage lines via waitlists and priority queues (Green, 2006), which pose several challenges. First, waitlists rarely appear in administrative data and are imperfectly codified and discretionary (Dixon and Siciliani, 2009; Cullis *et al.*, 2000). Second, with waitlists, a hospital's queue affects demand at all other hospitals, creating an intractable interaction in demand across hospitals. This makes studying competition and estimation extremely challenging. Finally, even if one assumes away these issues by considering a single-firm problem or static demand, the data will largely reject any model of rigid capacity because hospitals sometimes operate above capacity (Sharma *et al.*, 2008; Kelen *et al.*, 2021). This is despite the rigidity of planning capacity, as hospitals can

[19]The independence assumption on preferences rules out elements such as realized wait times from appearing in preferences. However, expected wait times based on information available at the beginning of the period can be incorporated without loss. This contrasts with work on feedback between wait times and demand (Martin and Smith, 1999; Russo, 2024).

temporarily reallocate beds from one use to another, hire flexible staffing, and rent beds or equipment when demand is high (Hick *et al.*, 2004; ASPR TRACIE, 2025; Abir *et al.*, 2020).[20] Given these challenges, we adopt a simpler approach and model hospitals as choosing the rate at which they admit patients, subject to expected admissions not exceeding capacity. In Appendix A.3.1, we show that this model is closely related to one in which the hospital manages a queue and provides wait-time guarantees to patients, treating patients first-come, first-served. Specifically, admission-rate control affords hospitals payoffs similar to those of queue management while simplifying interactions among firms.

Formally, service line j's admission-planning problem consists of choosing a rate, a_{hj}, at which to accept incoming patients. Since each line is a set of procedures aimed at a collection of demographic groups and medical conditions, we can define $\mathcal{K}_j = \{(s_k, d_k)\}_{k=1}^{K_j}$, the index of population segments and diagnoses associated with the service line. We denote n_{hk} as the total number of admissions from group k at hospital h.[21] Upon admission, each patient i of type k draws a length of stay ℓ_i^k independently from a distribution that depends on their type and has strictly positive support. We denote the total admitted length of stay as $N_{hj}^L \equiv \sum_{k\in\mathcal{K}_j} \sum_{i=1}^{n_{hk}} \ell_i^k$. The hospital line solves

$$\pi_{hj}(b_{hj}) = \max_{a\in[0,1]} \mathbb{E}\left[\sum_{k\in\mathcal{K}_j} n_{hk}(a)p_{hk} - C_{hj}(a, b_{hj}) + \tau_{hj}S_{hj}^u(a)\right] \quad \text{s.t.} \quad \underbrace{\mathbb{E}[N_{hj}^L(a)] \leq b_{hj}}_{\text{ex-ante capacity constraint}} \tag{8}$$

where p_{hk} is the revenue per admission of type k, $C_{hj}(a, b_{hj})$ is the total realized cost of the service line given its admitted demand, $S_{hj}^u(a)$ is the realized surplus of the uninsured, and $\tau_{hj} \geq 0$ is the dollar-valued weight that the hospital puts on serving the uninsured. The ex-ante capacity constraint stipulates that the line manager cannot choose a rule that would lead the service line to overflow in expectation. This allows the model to capture the binding constraints hospitals face while incorporating the potential for excess demand.

While excess demand cannot occur in expectation, it can occur ex post. Hospitals with overflowing wards might face high costs from the additional burden on staff and resources, payments to flexible staff, and rented capital. In addition, hospital efficiency may suffer when demand exceeds the hospital's physical capacity. To incorporate these costs, we model the hospital's cost

[20] In practice, hospitals staff fewer beds than they are licensed to operate, which affords them flexibility. In our data, hospitals with more than 50 beds have on average 12.4% of their total licensed beds unstaffed, with a median of 5.7%. Regulation also allows hospitals to temporarily exceed licensed bed capacity subject to some restrictions (Centers for Disease Control and Prevention, 2026; California Department of Public Health, 2018; Maryland Department of Health, n.d.; Massachusetts Department of Public Health, 2022).

[21] The line subindex is omitted since each k is associated with a unique j. A patient ill with a given condition and from a particular demographic group can be served by a single service line at each hospital. For example, a pediatric orthopedic patient cannot be cared for by an adult oncological line.

function as:

$$C_{hj}(a,b) = \underbrace{\sum_{k\in\mathcal{K}_j} n_{hk}(a)c_{hk}}_{\text{surgical cost}} + \underbrace{c^V_{hj}N^2_{hj}(a)}_{\text{congestion cost}} + \underbrace{\exp\left(c^o_h(N^L_{hj}(a)-b)\right)}_{\text{overflow cost}} \tag{9}$$

where $N_{hj} \equiv \sum_{k\in\mathcal{K}_j} n_{hk}$ is the total number of admissions. Above, the first term corresponds to the surgical cost paid per admission, the second captures the cost of congestion, and the third, the cost of overflow.[22] We assume throughout that the cost parameters of the firm $(c_{hk}, c^V_{hj}, c^o_h)$ are positive and note that the overflow cost parameter is common across lines for the hospital, consistent with it emerging from a common cost of rental and staffing that grows increasingly expensive as a line overflows.

Given the non-discriminatory nature of admission rates and the simultaneity of hospital choices, we can simplify the firm's problem. In particular, since each individual-condition pair (i,d) is associated with a single consumer-type k, we can write the conditional choice probability in Equation (6) as $\sigma_{kh} = a_{hj(k)}R_{hk}$, where $j(k)$ is the service line associated with k. In addition, because choices are simultaneous, each admission rate is independent of the others, conditional on the information available at the beginning of the period. Hence, the request probability, R_{hk}, is not a function of hospital h's admission rate. Letting λ_k denote the rate at which type k patients fall ill, the Poisson thinning property implies that the number of type k patients admitted is Poisson-distributed with rate $a_{hj}\lambda_k R_{hk}$. Similarly, the total number of patients admitted into the line, N_{hj}, is Poisson-distributed with rate $a_{hj}\Lambda_{hj}$, where $\Lambda_{hj} \equiv \sum_{k\in\mathcal{K}_j} R_{hk}\lambda_k$ is the maximum admission rate. Finally, the total length of stay of admitted patients, N^L_{hj}, follows a compound Poisson distribution, with jumps drawn from a mixture of the length-of-stay distributions for each patient type. Given these properties, Equation (8) simplifies to (proof in Appendix A.3.3):

$$\begin{aligned} \pi_{hj}(b_{hj}) = \max_{a\in[0,1]} \quad & a\left(\sum_{k\in\mathcal{K}_j}\lambda_k R_{hk}(p_{hk}-c_{hk}) - c^V_{hj}\Lambda_{hj} + \tau_{hj}\Delta S^u_{hj}\right) - a^2 c^V_{hj}\Lambda^2_{hj} - \exp(a\chi_{hj} - c^o_h b_{hj}) \\ \text{s.t.} \quad & a\Lambda_{hj}\mu^L_{hj} \;\le\; b_{hj} \end{aligned} \tag{10}$$

Where μ^L_{hj} is the mean length of stay for patients admitted to line j, ΔS^u_{hj} is the expected contribution of line j to uninsured consumer surplus, and χ_{hj} is a known function of overflow cost parameter c^o_h and moments of the population length of stay.[23]

[22] The physical scale of the line is absorbed in the cost parameter c^V_{hj}. The overflow cost also introduces a slight under-utilization cost. The true, intractable, overflow cost of a hospital can be modeled as $\exp(c^o_h \max\{N^L_{hj}(a)-b,0\})-1$. The constant subtraction does not alter the firm's planning and is therefore ignored in the analysis. The lack of kink at zero is an important simplification but of minimal numerical relevance given the exponential decay.

[23] Formally, $\Delta S^u_{hj} = \sum_{k\in\mathcal{K}^u_j}\lambda_k \mathbb{E}_{i,\boldsymbol{a}_{-h}}[\ln(\frac{1+e^{v_{ih}}+\sum_{h'\in\mathcal{H}\setminus\{h\}}\tilde{a}_{h'i}e^{v_{ih'}}}{1+\sum_{h'\in\mathcal{H}\setminus\{h\}}\tilde{a}_{h'i}e^{v_{ih'}}})]$ where $\mathcal{K}^u_j$ is the subset of patients that are uninsured, and

Despite the capacity and ex-ante feasibility constraints, the firm's problem is well-behaved and carries four key properties that we leverage extensively in our analysis. In what follows, we denote $a^o_{hj}(b_{hj})$ and $a^*_{hj}(b_{hj})$ the interior and constrained solutions to Equation (10), respectively. We denote (a^*_{hj}, b^*_{hj}) as the solutions to the firm's overarching problem. The following proposition, proven in Appendix A.3.4, formalizes the key properties of the model.

Proposition 1. *If for any $j \in \mathcal{J}_h$, $\Lambda_{hj} > 0$ and for every $k \in \mathcal{K}_j$, $\mathbb{E}[e^{c^o_h \ell^k}] < \infty$. Then a^o_{hj} and $a^*_{hj}(b_{hj})$ have closed-form expressions and*

1. *The hospital's problem is jointly concave on $(\boldsymbol{a}, \boldsymbol{b})$ and has a unique solution.*
2. *$a^*_{hj}(b_{hj})$ is strictly increasing in b_{hj} whenever $a^*_{hj}(b_{hj}) \in (0, 1)$.*
3. *Let $j, j' \in \mathcal{J}_h$, such that $b^*_j, b^*_{j'} > 0$, then $\frac{\partial \pi_{hj}(b_{hj})}{\partial b_{hj}} = \frac{\partial \pi_{hj'}(b_{hj'})}{\partial b_{hj'}}$, whenever these derivatives are well-defined.*
4. *The game admits an equilibrium in pure strategies $(\boldsymbol{a}^*, \boldsymbol{b}^*)$*

The assumptions of the proposition simply state that expected demand for the line is positive and that the distribution of length of stay decays fast enough relative to overflow. Given the demand model developed above, a sufficient condition for positive demand is that at least one consumer type k has finite preferences and positive illness probabilities. The second part of the assumption imposes basic feasibility: otherwise, expected overflow tail risk would overwhelm the firm's objective, and the line should never be provided.

The first part of the proposition is fundamental, as it implies that each hospital's problem has a unique solution. The second establishes that admission rates can be inverted to recover resource allocations, which will prove crucial in identification and estimation. The third forms the core of the identification argument we develop later: because resources are transferable across service lines, at the optimum, the manager must be indifferent in its allocation of the marginal unit. This will allow us to identify differences in marginal costs across lines by the differences in their marginal revenues. The closed-form expression for $a^*_{hj}(b_{hj})$ significantly improves the tractability of the model's analysis and simulation, thereby reducing the firm's problem to the resource allocation stage. Finally, the fourth condition ensures that the game admits an equilibrium, which might not occur if the constraint on hospital actions generated discontinuities in the best-response mapping across firms. In the proof appendix, we also discuss conditions for equilibrium uniqueness, which require hospitals' reactions to rivals to be sufficiently well bounded. Intuitively, this would happen if hospitals are sufficiently differentiated. However, establishing economically meaningful global conditions for uniqueness in oligopolistic games is a common challenge in our framework. We discuss how we handle potential multiplicities when presenting our counterfactual policy analysis.

$\chi_{hj} = \Lambda_{hj}(M^L_{hj}(c^o_h) - 1)$, where $M^L_{hj}(t)$ is the moment generating function for the length-of-stay of line j patient. See Appendix A.3.3 for the derivation.

5.3 Discussion

The model above presents a tractable framework for studying resource allocation and crowd-out in hospital care. Service lines form the basis for product segmentation and production in the model, with demand and resources assigned to one line being isolated from those of others. While service lines are organizational structures used by hospitals in practice, our model assumes a few key simplifications. First, since admission rates are assumed non-discriminatory within each service line, any patient selection occurs only across lines. Thus, patients are selected by allocating fewer resources to less profitable lines, but patients sharing a service line must be treated equally.[24] Second, it is assumed that the binding resource is bed-days, which implicitly assumes that labor is flexible and that hospitals have sufficient fixed capital to treat patients across lines, or can rent it when needed. Third, because resources are line-specific, we assume any overflow resolved ex post by reallocating idle resources within the hospital carries the same opportunity cost as overflow resolved with external resources.[25]

Our model fits within the workhorse structural insurer-hospital framework (Town and Vistnes, 2001; Gowrisankaran *et al.*, 2015; Ho and Lee, 2017). In particular, as long as insurer-hospital contracts do not condition on resource allocations, the framework can easily accommodate a pricing stage. We do not model pricing here because our data are not amenable to studying hospital-insurer negotiations, and we expect changes in uninsured care to have minimal impact on these prices. Our demand model generalizes the standard logit demand model used in the literature but is more rigid than a general consideration-set model (Abaluck and Adams-Prassl, 2021; Barseghyan *et al.*, 2021; Agarwal and Somaini, 2026). In particular, we assume patients know all relevant hospitals, so admission events arise exclusively from preferences and availability. This entails a trade-off between generality and tractability: we show below that our demand model yields a simple choice-likelihood, whereas general consideration-set models are often hard to estimate and simulate. In addition, we assume that a patient's length of stay is beyond the hospital's control despite evidence that hospitals have a say over discharge timing (Doyle, 2005). If changes in the uninsured burden and the compensating policies we study do not meaningfully alter the speed at which hospitals discharge each consumer type, this assumption is likely innocuous. In exchange, our model provides a tractable framework for competition with binding capacity constraints that can rationalize crowdouts and resource reallocation, thereby explaining the patterns we documented in Section 4.

[24] This contrasts with work studying direct patient selection, such as Gandhi (2026).

[25] A sufficient condition is that hospitals be able to lease idle capacity at the same market rate at which they rent it.

6 Identification and Estimation

This section presents the identification arguments for the model primitives and discusses their estimation. We estimate the model using individual-level data from Florida for 2011 to 2019, with quarters as the period length. We adopt a parsimonious classification of service lines, splitting them into emergency care, scheduled insured care, and scheduled uninsured care. This classification allows hospitals in the model to meet emergency care regulations and to divide scheduled care along the insurance lines we are interested in. Scheduled insured surgery does not discriminate in admissions across different insurance types, but we allow hospitals to reject Medicaid or Medicare patients by excluding them from patients' choice sets.[26] Proofs, derivations, and technical details are relegated to Appendix A.4.

6.1 Demand

Our demand model and identifying variation are a special case of the general setting considered in Agarwal and Somaini (2026), and are almost identical in structure to the empirical application in that article. As established there, separately identifying patient preferences from acceptance probabilities requires a shifter of preferences excluded from hospitals' acceptance decisions, and a shifter of acceptance decisions excluded from patients' preferences. The first shifter is naturally the distance from a patient's home to the hospital, which affects patients' preferences for hospitals but—all else equal—is irrelevant for hospitals' admission-rate decisions. The second shifter is the expected congestion at the hospital, which shifts hospitals' chosen admission rates, but this is assumed to be excluded from patients' preferences. In particular, in our application, hospital congestion changes due to Medicaid spillover effects, disease seasonality, and the local population's exposure to these forces.

We estimate the model via a nested fixed-point maximum-likelihood procedure. In Appendix A.4.1 we show that patient i's conditional choice probabilities can be written as

$$\sigma_{idht} = a_{hj(i,d)t}\bar{\sigma}_{idht}\int_0^\infty e^{-u}\prod_{h'\in\mathcal{H}_i\setminus\{h\}}\left(a_{ih't} + (1-a_{ih't})e^{u\bar{\sigma}_{idh't}}\right)du \tag{11}$$

where $\bar{\sigma}_{idht} = e^{v_{idht}}/(1+\sum_{h'\in\mathcal{H}_{it}} e^{v_{idh't}})$ is the standard logit choice probability without admission constraints. We note that if all hospitals admit patient i with probability one ($\alpha_{hjt} = 1$), then the expression reduces to the standard logit form. In contrast, admission probabilities increase with the rejection rate of rivals, as governed by the last term in the integral. This expression allows us to efficiently evaluate the log-likelihood of the data given a parametrization of patients' preferences, v_{idh}, and a vector of acceptance probabilities, $\boldsymbol{a}$. In estimation, we recover acceptance probabilities

[26] In particular, we assume that hospitals without an emergency room are excluded from emergency patients' choice sets. For scheduled care patients, we assume that Medicaid patients can only access hospitals that have historically accepted Medicaid scheduled surgery, and similarly with Medicare patients.

from the fixed-point equation

$$a_{hjt} = \frac{N_{hj}^{L\star}}{\sum_{i \in \mathcal{I}_{jt}^*} R_{ih}(\boldsymbol{a}, \boldsymbol{v}) \ell_i^*} \tag{12}$$

where $N_{hj}^{L\star}$ is the observed total admitted length of stay at the hospital, $\mathcal{I}_{jt}^*$ is the observed set of individuals seeking admission in the period for line of service j, $R_{ih}(\boldsymbol{a}, \boldsymbol{v})$ is the model-implied request likelihood, and ℓ_i^* is the observed length of stay of the patient. We provide sufficient conditions for this expression to yield a contraction mapping in Appendix A.4.2.

We model patients' preferences as a linear function of distance and system preferences for hospital-diagnosis combinations. In particular, $v_{idh} = \alpha_{i,j(i,d)} dist_{ih} + \gamma_{j(i,d)dh}$, where $dist_{ih}$ is the distance from a patient's home to the hospital, and $\gamma_{j(i,d)dh}$ is a hospital line-by-diagnosis-group fixed effect. We allow patients' disutility of travel to vary by service line, market, gender, age group, and whether they live in a ZIP code with a mean income below the state median.[27] We define markets per service line by grouping contiguous Health Referral Regions (HRRs) such that at least 90% of the demand originating from patients residing in the market is served within the market.[28] Appendix Figure A.2 shows the resulting market classifications, highlighting how emergency markets are narrower given that patients usually seek care closer to home, while uninsured scheduled care markets are more extensive. In estimation, we normalize patients' distance to their in-network hospitals by their distance to their nearest out-of-market provider to account for differences in the cost of substituting local for out-of-market care. Additional technical details about the estimation procedure are provided in Appendix A.4.3.

6.2 Supply

Identifying hospitals' costs hinges on the optimality of their admission and resource allocation decisions. Recall that the hospital must allocate bed-days and set acceptance rates across its service lines to maximize its total (subjective) profits. The hospital is constrained by the requirement that the expected total length of stay cannot exceed the allocated capacity per line, and that the total allocated resources cannot exceed what the hospital has available. From the econometrician's perspective, we can consider the hospital's revenue, admissions, length of stay, and total available bed-days as data. Moreover, given the identified demand, we can treat the admission rates as

[27] Our age groups follow the definition in HCUP, which classifies patients as under 6, 6 to 17, 18 to 24, and then in ten-year increments until age 74. Ages 75 and over are a single category. We define diagnostic groups by selecting the 30 most common Diagnostic Related Group (DRG) codes for each age group and gender, and clustering the remaining DRGs into 10 bins based on their DRG weights.

[28] We adopt this broad market classification for two reasons. First, uninsured scheduled care is often delivered far from home, likely due to the scarcity of providers willing to offer those services. Hence, standard market classifications used in the literature are unlikely to be appropriate for this population. Second, to link the demand and supply model, we need to minimize the share of patients allocated to the abstract out-of-market provider to maintain accurate capacity utilization accounting in counterfactuals.

known. However, resource allocation across lines is unobserved. Proposition 1.2 provides the remedy: resource allocations may be inverted from admission rates.

Before stating the main identification result, we add two assumptions on the hospital's costs and present the key identifying conditions. First, we assume surgical costs are an affine function of the observed resource intensity measure (i.e., DRG weights), ω_k, such that $c_{hkt} = \omega_k c^S_{hjt} + c^F_{hjt}$.[29] Second, we assume that a hospital's cost parameters do not vary within a year, implying that we observe the hospital making four choices—one per quarter—under the same cost configuration, but potentially different demand. Collecting the linear cost and weight terms $\boldsymbol{\theta}_{hjt} = (c^F_{hjt}, c^S_{hjt}, c^V_{hjt}, \tau_{hjt})$, the first-order conditions for the optimal admission rate of Equation (10) and the optimal resource allocation of Equation (7) state that for any line with $a_{hjt} \in (0, 1)$ it holds that:

$$p_{hjt} = \overbrace{\boldsymbol{x}'_{hjt}\boldsymbol{\theta}_{hjt}}^{\text{marginal service cost}} + \overbrace{\frac{\chi_{hjt}(c^o_{ht})}{\Lambda_{hjt}} \exp(\chi_{hjt}(c^o_{ht})a_{hjt} - c^o_{ht}b_{hjt}) + \nu_{hjt}\mu^L_{hjt}}^{\text{marginal capacity cost of admission}} \tag{13}$$

$$\eta_{ht} = \underbrace{c^o_{ht} \exp(\chi_{hjt}(c^o_{ht})a_{hjt} - c^o_{ht}b_{hjt}) + \nu_{hjt}}_{\text{marginal value of allocated capacity}} \tag{14}$$

Where $p_{hjt} := \sum_{k \in \mathcal{K}_j} \frac{\lambda_{kt} R_{hkt}}{\Lambda_{hjt}} p_{hkt}$ is the expected revenue of the marginal admission to line and the term $\boldsymbol{x}'_{hjt}\boldsymbol{\theta}_{hjt}$ collects the surgical and congestion costs where $\boldsymbol{x}'_{hjt} := (1, \sum_{k \in \mathcal{K}_j} \frac{\lambda_{kt} R_{hkt}}{\Lambda_{hjt}} \omega_k, 1 + 2\Lambda_{hjt} a_{hjt}, -\frac{\Delta S^u_{hjt}}{\Lambda_{hjt}})$ are all simple aggregate elements: a constant, the expected resource intensity of an admission to the line, one plus twice the expected admissions, and the average effect of admission on uninsured surplus. The term labeled *marginal capacity cost of admissions* in Equation (13) corresponds to the effect of the marginal admit on overflow risk and the opportunity cost of their expected bed-day utilization. In particular, ν_{hjt} is the Lagrange multiplier associated with the admission-rate constraint and measures the within-line opportunity cost of an admission: it is zero if, in expectation, the line has idle capacity, and positive if the marginal patient crowds out a profitable one. Thus, Equation (13) simply states that the revenue of the marginal admission must equal its cost.

A tightly connected equality emerges from the optimality of resource allocation in Equation (14). This condition states that the opportunity cost of an allocated bed-day, η_{ht}, must equal its marginal value at its allocated line. This relates to Proposition 1.3, which ties the marginal profitability of resources across lines. In particular, it implies a unique internal price for a bed-day, recoverable from the optimality of the resource allocation. This optimality pins down the marginal capacity cost of an admission, leaving Equation (13) with its core linear structure. Thus, one might expect that the main additional assumption needed for identification is sufficient local variation in

[29] This structural assumption is commonplace in the literature on hospital competition (Gowrisankaran *et al.*, 2015; Ho and Lee, 2017; Liu, 2022; Cuesta *et al.*, 2024).

the components driving $\boldsymbol{x}_{hj}$, as in a linear model.

In Appendix A.4.4, we show that this is the case and prove the identification of $(\boldsymbol{\theta}, c^o_h)$, hospital-by-hospital, formally. The proof is extensive and somewhat technical because of the unobservability of the marginal capacity cost and the linear-exponential structure of Equation (13). Essentially, it proves that the internal cost of capacity is determined by Equation (14) up to the overflow cost parameter c^o_{ht}. This removes the unobserved allocation from the identifying equation and replaces it with a single unobserved price. This yields a structure similar to that of a linear equation, with an added exponential hospital-specific cost. A technical lemma then proves that sufficient local variation in expected marginal revenue and the components driving $\boldsymbol{x}_{hjt}$ is sufficient to identify the model. Crucially, this includes variation in demand among those seeking charitable care, as provided by the Medicaid spillovers we leverage. Without it, neither the cost of charitable care nor the hospital's preference for uninsured surplus is identified.[30]

We estimate the model using a maximum-likelihood approach that draws on key insights from our identification argument. The estimator contends with three challenges. First, it must manage the lack of observability of b_{hjt}. The identification argument resolved this by using Equation (14) and the complementarity conditions imposed by the admission rate problem. However, the estimator faces a second problem: complementarity problems are numerically challenging because they introduce significant non-convexity. Third, the estimator must leverage as much of the finite sample variation as possible, and thus cannot rely only on the cases where $a_{hjt} \in (0, 1)$. We address these three issues by estimating the shadow price of capacity, using a vanishing relaxation of the complementarity condition, and introducing a structural cost error with a distributional assumption to leverage the corner cases.

In particular, we introduce a parameter $\rho > 0$, and denote the slack of a service line as $s_{hjt} := b_{hjt} - \Lambda_{hjt}\mu^L_{hjt}a_{hjt}$. With this, we write the relaxed complementarity condition of line (h, j) in period t as $s_{hjt}\nu_{hjt} = \rho$. Clearly, as $\rho \to 0$, the condition converges to the Kuhn-Tucker complementarity condition, which holds at the optimum. Using these elements, we can write Equation (14) as

$$\eta_{ht} = c^o_{ht} \exp\left((\chi_{hjt}(c^o_{ht}) - c^o_{ht}\Lambda_{hjt}\mu^L_{hjt})a_{hjt} - c^o_{ht}s_{hjt}\right) + \frac{\rho}{s_{hjt}} \tag{15}$$

which has a unique solution for $s_{hjt}(\rho, \eta_{ht}, c^o_{ht})$.[31] From this and the total capacity constraint of the hospital, we have that η_{ht} is the unique value that satisfies $\sum_{j\in\mathcal{J}_h}\left(\Lambda_{hjt}\mu^L_{hjt}a_{hjt} + s_{hjt}(\rho, \eta_{ht}, c^o_{ht})\right) = B_{ht}$. This implies that for a given value of (ρ, c^o_h), resource slackness and the shadow cost of capacity are

[30] Additionally, the identification argument reveals that the econometrician need not know the hospital's payment per insurance source. This contrasts with models that identify hospital costs through hospital-insurer negotiations (e.g., Gowrisankaran *et al.* (2015)), in which the econometrician must assign negotiated prices to hospital-insurer pairs. This distinction matters because several prominent data sources—including ours—do not provide identifiers to distinguish among private insurers.

[31] See Appendix A.4.5 for a proof.

uniquely determined, as in the identification argument. Finally, we introduce a structural error by rewriting $c^F_{hjt} = \bar{c}^F_{hj} + \varepsilon_{hjt}$. Given that c^F_{hjt} enters equation (13) as an additive term, the structural error enters additively as well, yielding a non-linear regression structure:

$$p_{hjt} = \boldsymbol{x}'_{hjt}\tilde{\boldsymbol{\theta}}_{hjt} + \psi_{hjt}(c^o_{ht}; \rho) + \varepsilon_{hjt} \tag{16}$$

Where $\tilde{\boldsymbol{\theta}}_{hjt}$ is the set of cost parameters $\boldsymbol{\theta}$ with the decomposition c^F_{hjt} applied, and $\psi_{hjt}(c^o_{ht}; \rho)$ collects overflow and opportunity costs. We assume that structural cost shocks are normally distributed with a common variance and are drawn independently across periods and lines, slightly relaxing the assumption of yearly-homogeneous cost structures. This distributional assumption allows us to leverage the inequalities imposed by the optimality of corner solutions $a_{hjt} \in \{0, 1\}$. With this, we can write the log-likelihood of each hospital line in the data and solve[32]

$$\lim_{\rho \to 0} \max_{\boldsymbol{\theta}, \boldsymbol{c}^o} \sum_{h,j,t} \mathcal{L}_{hjt}(\boldsymbol{\theta}, c^o_{ht}; \rho) \quad . \tag{17}$$

We solve this problem over a decaying sequence of relaxation parameters. In estimation, we allow surgical costs (c^S_{hjt}) to vary by hospital-year, fixed-surgical cost (c^F_{hj}) to vary by service line type, congestion cost (c^V_{hjt}) to vary by hospital-line, and we allow the overflow cost parameter (c^o_{ht}) to vary by year and HRR, to capture local differences in rental and labor costs. Finally, we allow the hospitals' preference for uninsured surplus (τ^S_{hjt}) to vary by hospital system and to change in response to Medicaid spillover effects. Appendix A.4.5 presents further details on the estimator.

7 Results

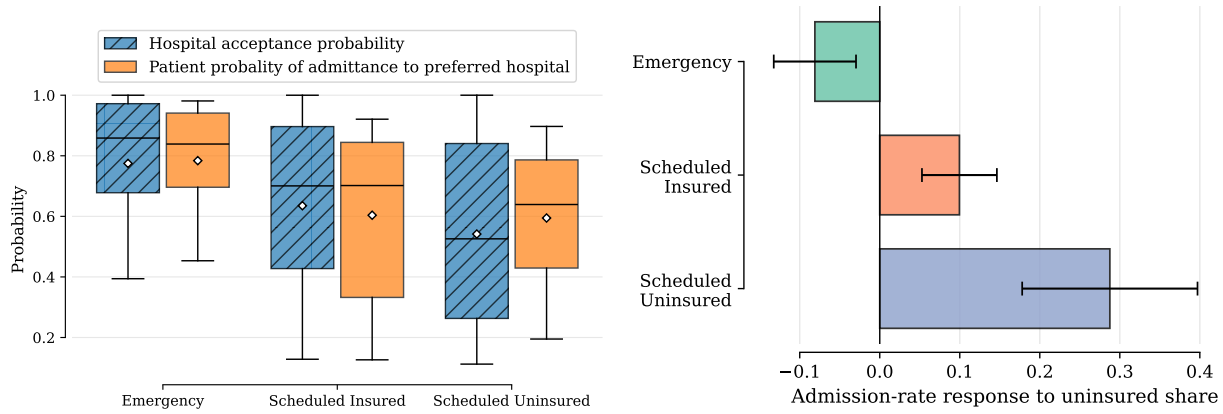
7.1 Demand

The demand estimates reveal substantial heterogeneity in acceptance rates, travel disutility, and hospital preferences.[33] Figure 5a shows the distribution of acceptance rates at hospitals in the data, marking how patients seeking different types of care are limited in their choices. As shown, the median hospital with an emergency room accepts over 85.9% of its incoming requests. In contrast, the median hospital accepts only about 70.1% of scheduled insured admission requests, while for scheduled uninsured care, the share is slightly over 52.6%.[34] To relate these acceptance rates to preferences, the figure also shows how often patients are admitted to their (stochastically) most preferred hospital. The difference between the two plotted distributions indicates how patients

[32] A similar limiting approach to estimation was used in Cuesta *et al.* (2024) to resolve the non-convexity induced in a combinatorial plan-design problem.

[33] Appendix Table A.7 display distance disutility estimates. Appendix Figure A.4 displays in- and out-of-sample fit.

[34] We note that these estimates are conditional on hospitals being within the associated patient choice set. Thus, the effective share of hospitals serving scheduled uninsured patients is significantly lower.



(a) Acceptance Rates

(b) Connection to Uninsurance

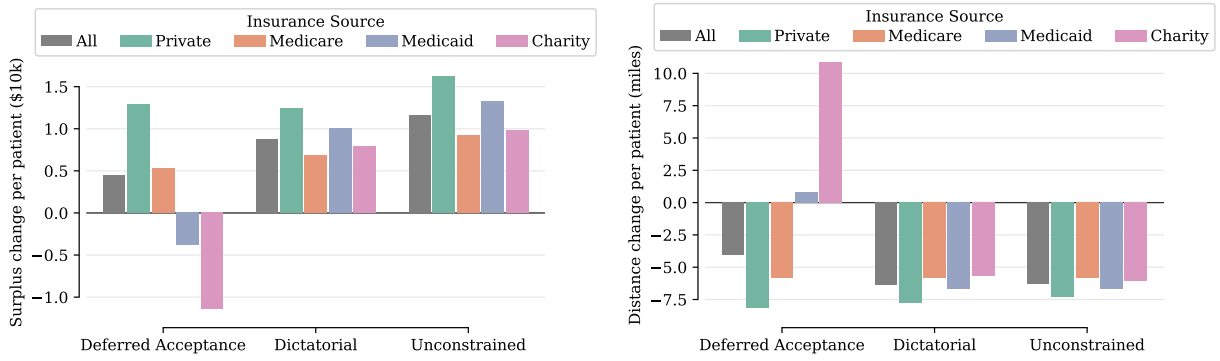
Figure 5: Acceptance Rates and Local Uninsurance

Notes: Figure 5a shows the estimated distribution of acceptance rates (a_{hjt}) from the demand model. The whiskers on the box plot cover the 10th to 90th percentiles, while the box covers the interquartile range. A horizontal line represents the median, and a white diamond represents the mean. The hatched blue boxes show the distribution across hospitals. The solid yellow boxes show the probability that patients are admitted to their most preferred hospital ($\sum_{h \in \mathcal{H}_i} a_{hj(i,d)t} \bar{\sigma}_{idht}$), including the outside option when applicable. Figure 5b shows how the probability of patients being admitted to their most preferred hospital changes with the local share of uninsured. The plot presents the estimate and 95% confidence interval for the coefficient from a regression of the ZIP Code Tabulation Area (ZCTA) average admission rate to preferred hospital on the local uninsurance rate, conditional on ZCTA-by-service-line fixed effects and a year fixed effect to account for state-level admission trends. Standard errors are clustered at the ZCTA level.

experience the skew in the capacity distribution (see Appendix Figure A.6). In particular, although overall acceptance rates for uninsured scheduled patients are low, patients tend to systematically favor larger public hospitals, resulting in lower rejection rates. Appendix Figure A.3 presents how the probability of being accepted to a preferred hospital varies across markets and how it is related to the mean local elasticity to travel distance.

The demand estimates reveal admission patterns consistent with our supply theory. Regressing the estimated probability with which a patient is admitted to her preferred hospital on the local uninsurance rate, and exploiting only variation within the local area in uninsurance rate relative to the rest of the state, we obtain the patterns shown in Figure 5b. As shown, as the local uninsurance rate grows, the probability that an emergency admission request is accepted at the patient's preferred hospital decreases, while the scheduled insured admission rate at preferred hospitals increases, consistent with the basic theory of Section 2. In addition, the scheduled uninsured admission rate also increases, which our richer model of Section 5 can rationalize via the hospital's weight on uninsured surplus.

7.1.1 Allocative Efficiency. To assess the importance of the estimated heterogeneity, we can use it to evaluate the efficiency of resource allocations across patients. In particular, conditional on insurance status and medical need, whether one patient is admitted while the other is referred to alternatives often depends more on the randomness of illness onset than on preferences. Patients cannot trade their allocations, and rigid hospital prices combined with insurance coverage make



(a) Change in Surplus

(b) Change in Distance Traveled

Figure 6: Allocative Efficiency Holding Supply Fixed

Notes: These figures display the estimated allocative efficiency of allocation in 2019 relative to three alternative benchmarks: patient-proposed deferred acceptance matching, dictatorial surplus-maximizing assignment, and unconstrained surplus-maximizing assignment. The first two assignments are constrained to resort patients across hospital lines and quarters using no more of each hospital line's resources than were used in the status quo. All changes are relative to the status quo. All dollar values are in 2023 adjusted dollars, with surplus transformed at $1,033 dollars per mile, following Prager (2020) estimates.

prices a negligible sorting force in hospital markets. Therefore, efficient trade might exist between patients, depending on how their heterogeneity interacts. To measure this, we conduct a three-step counterfactual assignment exercise: first, we introduce a market allocation mechanism; second, we consider patient-optimal dictatorial assignment; and finally, we measure the surplus that would result if each patient obtained access to their most preferred hospital. The first step thus reveals the loss the market incurs from lacking a coordination mechanism. The difference between this and the dictatorial assignment reveals the extent of misallocation that a simple market mechanism cannot remedy. The final step reveals the extent to which capacity is the binding constraint for surplus.

Figure 6 displays the estimates for 2019, measuring changes in per-patient surplus and the effect on miles traveled for care. Given that our model yields patient surplus, measured as the compensating variation in miles traveled for care, we dollarize surplus throughout using the estimates of Prager (2020).[35] First, to approximate what would occur under a market-based allocation mechanism, we use a patient-proposed Deferred Acceptance (DA) matching mechanism.[36] Sim-

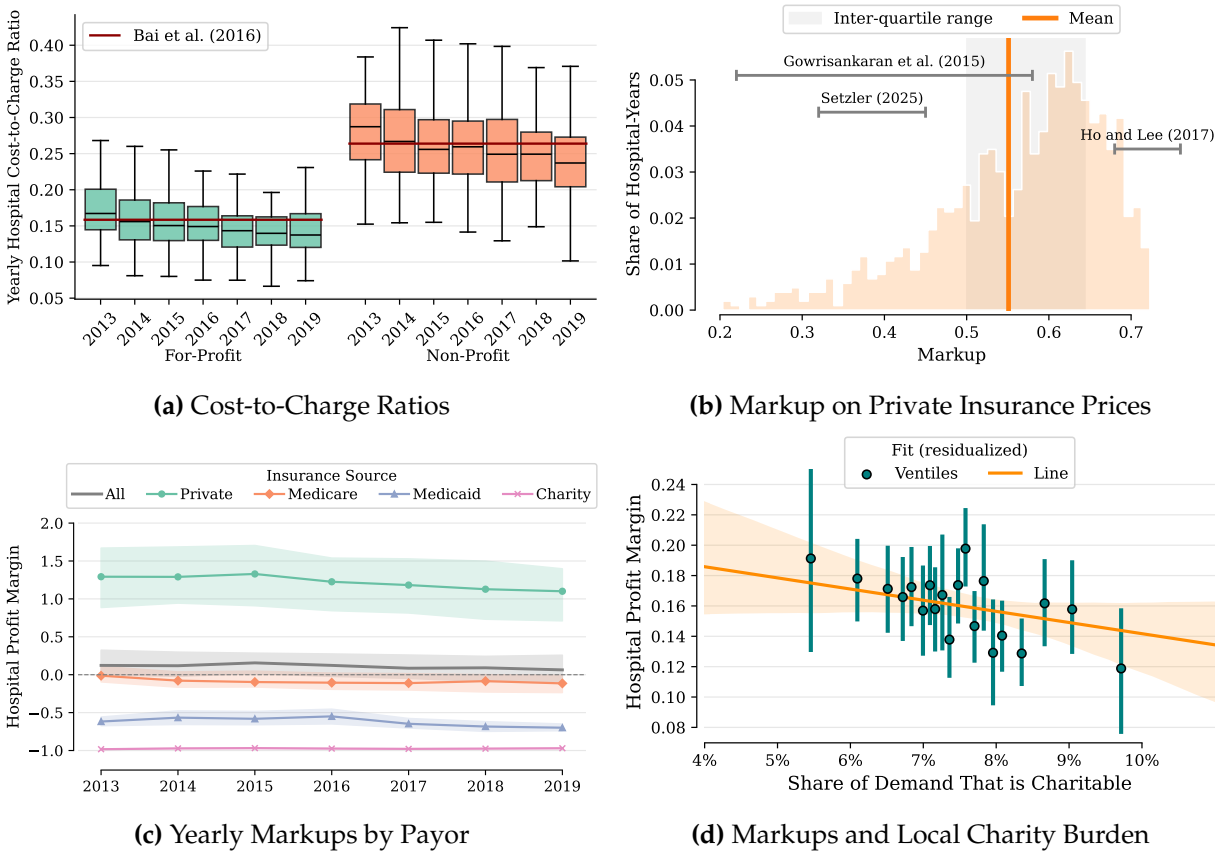
[35]Prager (2020) estimates patients' dollar-valued disutility from travel in a setting where patients face simple, deterministic, and known copays for hospital care. This observability and simplicity make these estimates particularly appealing to our setting. To be conservative, we rely on the lowest estimate in the article, $1,033 per patient-mile (adjusted to 2023 dollars using CPI). We discuss alternative measures in Appendix A.5.1. The outcome of the analyses in this section is independent of the conversion scale used.

[36]Ideally, we would conduct a counterfactual where hospitals could price care directly to consumers to determine a standard market-based allocation outcome. However, patient elasticity to prices is known to be modest because factors that might not hold if patients had to pay for care directly. Nevertheless, an extensive theoretical literature draws tight connections between stable matching mechanisms and competitive prices, establishing criteria under which our matching counterfactual can be viewed as approximating a price equilibrium (Hatfield and Milgrom, 2005; Bulow and Levin, 2006; Hatfield *et al.*, 2013; Azevedo and Leshno, 2016).

ilar mechanisms are often used to allocate scarce resources among heterogeneous players when prices are unavailable or undesirable, such as in labor markets (Roth and Peranson, 1999), education (Arteaga *et al.*, 2022), the military (Greenberg *et al.*, 2024; Davis *et al.*, 2026), and refugee settlement (Delacrétaz *et al.*, 2023), among others. Intuitively, a similar system could be implemented in hospital care by establishing a common scheduling system across hospitals in the state. These mechanisms consider patient and hospital preferences, are pairwise stable, yield the unique patient-optimal stable match, and make truth-telling optimal for patients. Moreover, by the Rural Hospital Theorem, every stable match assigns the same set of patients to the same set of hospitals, and every hospital with idle resources is assigned the same set of patients in every stable match (Roth, 2008). The estimates in Figure 6, reveal clear winners and losers from coordination, or alternatively, from the lack thereof in the current markets. In particular, Figure 6a shows that privately insured patients lose \$12,914.5 per patient per year due to a lack of market coordination, Medicare patients lose \$5,256.9, while Medicaid and charitable care patients gain \$3,757.6 and \$11,33.2. This shows how the current market assignment system protects what is likely the most vulnerable segment of the population. Figure 6b shows that this protection leads to significant additional travel for care for privately insured and Medicare patients, while decreasing it for Medicaid and, most importantly, for charitable care patients. As expected, lack of coordination benefits those outside of formal network arrangements the most, and establishing such coordination would push a large share of emergency charitable care demand outside their markets and often out of the state (see Appendix Figure A.7).

The dictatorial assignment in Figure 6 reveals the extent of misallocation in the status quo. As shown, directly allocating patients across hospitals could, on average, increase the surplus for all patient types by \$8,721.6 per patient per year. The surplus gain for private and Medicare patients, as well as their travel distance, changes little under dictatorial assignment compared to under DA. The dictatorial assignment thus reveals three things: first, that the misallocation produced by the current resources could be improved significantly under a more heavy-handed approach, which would likely entail establishing transfers across hospitals to account for differences in revenue; second, that some of these optimal assignments would be pairwise unstable, implying that some patients and hospitals would prefer to deviate and match out of the system if allowed; third, and more subtly, that many private and Medicare patients are nearly indifferent between some of their local alternatives, which generates space for large total surplus gains by resorting patients while moving near the Pareto frontier.

Finally, the unconstrained solution shows the extent to which status quo resource allocation limits surplus. In particular, we estimate that average patient surplus would increase by \$11,546.2 per year if capacity constraints were removed. This exercise is meaningful when compared to the dictatorial assignment—it reveals that 24.4% of allocatively efficient gains are infeasible due to capacity constraints, with the insured being the most harmed by these constraints in relative terms. The comparison in travel distance between dictatorial and unconstrained reveals that most of this



(a) Cost-to-Charge Ratios

(b) Markup on Private Insurance Prices

(c) Yearly Markups by Payor

(d) Markups and Local Charity Burden

Figure 7: Estimated Markups

Notes: These figures display markup statistics implied by the structural cost estimates. Figure 7b shows the estimated cost-to-charge ratio for each hospital, where costs are estimated and charges are as reported by hospitals. Figure 7b shows the distribution of estimated hospital markups over private insurance prices. The shaded gray area represents the interquartile range of our estimates, and the solid orange line represents the mean. Figure 7c shows the estimated profit margin of hospitals, as a percent of their total cost. The lines show the median, and the shaded areas represent the interquartile range of estimates in each line and year. Figure 7d shows the estimates from regressing a hospital's profit margin on the share of their demand that obtains charitable care, conditional on a hospital fixed effect. The green dots represent ventiles of the data after being residualized on the hospital fixed effect. The orange line shows the fit over the residualized data. The estimated slope is −0.735 with a standard error of 0.274.

loss comes from worse matches between patients and hospitals according to perceived quality, rather than distance. Thus, adjusting for their relative probability of illness, we estimate that the average privately insured Floridian loses $790 per year from lack of market coordination — 1.09 times the average marketplace premium in Florida at the time — and up to $991 from limited capacity and resources. In contrast, the average uninsured Floridian gains $1,081 from the lack of market coordination, and would benefit by $932 if they could be admitted to their most preferred hospital without capacity or insurance constraints.

7.2 Supply

The supply estimates show hospital costs and stark differences in markups across payors.[37] Given that our model's structure and identification argument are different from those in the literature we build upon, we begin by benchmarking our estimates to prior ones. Figure 7a shows how the distribution of cost-to-charge—the ratio of estimated hospital yearly costs to their reported total charges—relates to the estimates of Bai and Anderson (2016). While the reference computes these ratios directly from 2013 financial data, our estimator infers them from optimal resource allocation using no supplemental cost data. Yet, both yield similar estimates, with costs at 16.7% of charges for for-profit hospitals (compared to 15.8%) and 28.7% for non-profit hospitals (compared to 26.4%). Figure 7b instead compares our estimated markup (over prices) from private insurers to a collection of reference papers. Our estimates lie within the range of markups implied by the structural models of Gowrisankaran *et al.* (2015), Ho and Lee (2017), and Setzler (2025). However, each reference estimate is based on a different sample of hospitals, insurers, and years.[38]

Figure 7c shows how hospital markups differ across payment sources and have varied over time. A few notable patterns emerge. First, Medicaid pays hospitals far below their costs, while Medicare pays very close to hospital costs.[39] In particular, we find that median hospitals' profit margin from Medicare patients has been declining, with an estimate of -1.3% in 2013, decreasing to -11.4% by 2019. Similarly, the Medicare Payment Advisory Commission's (MedPAC) analysis of hospitals' financial data found that short-term acute-care hospitals experienced declining profit margins on Medicare services (including, but not limited to, inpatient hospitalizations), reaching -8.7% in 2019 (Medicare Payment Advisory Commission, 2015, 2021). Second, in sharp contrast, the privately insured pay an extraordinary but declining markup for hospital care, with a median profit margin from 129.3% over cost to 110.1% over cost. This is consistent with evidence that private payment rates for inpatient hospital services range across studies from 150% to 222% of Medicare rates (Lopez *et al.*, 2020). Third, despite this markup, hospitals' overall profit margin is limited and declining, with the median margin at 12.3% in 2013 and falling to 6.3% by 2019. For comparison, MedPAC estimated a 7.6% all-payer margin for inpatient hospitals for 2019. These declining trends in overall hospital markups contribute to an extensive debate about the evolution of markups in the US economy (Basu, 2019; Berry *et al.*, 2019; De Loecker *et al.*, 2020; Grieco *et al.*, 2024; Shapiro and Yurukoglu, 2026). Notably, we estimate that the bottom quartile of hospital

[37] See Appendix Figure A.8 for the estimated hospital cost components. See Appendix Figure A.9 for fit.

[38] Gowrisankaran *et al.* (2015) uses a sample of 11 hospitals from Northern Virginia between 2003 and 2006 with data from four large private insurers. Ho and Lee (2017) study the California market in 2004 with data from large private insurance. We note that the article does not report hospital markups directly, so we infer the range from reported estimates. Setzler (2025) studies a large collection of hospitals from 1996 to 2022, infers private prices from HCRIS reports, and reports markups by hospital size categories.

[39] We note that our measurement of hospital revenue includes all payments received by hospitals from Medicare and Medicaid that compensate them for serving low-income or uninsured patients, as well as an approximate recovery rate from charitable care patients, which we obtain from hospitals' tax returns.

margins has been negative since 2016.

Against the backdrop of hospitals' limited profit margins and the large markups on private payors, it may appear that private insurers subsidize care for other enrollees. A large body of literature debates the existence of such cost-shifting (or cross-subsidizing), with mixed evidence (Cutler, 1998; Frakt, 2011; White, 2013; Clemens and Gottlieb, 2017; Dranove *et al.*, 2017). A natural test for this hypothesis is whether overall hospital markups are affected by their local exposure to uninsured demand. If hospitals can shift costs to the privately insured, we would expect markups to be stable in local exposure. Figure 7d shows the slope resulting from regressing hospital profit margins on the share of their demand that is admitted for charitable care, conditional on a hospital fixed effect. The result implies that for every 1% increase in a hospital's charitable care load, its profit margin declines by 4.59%. This decline suggests cost-shifting, if present, is very limited. Thus, a natural reaction to the inability to cost-shift is for hospitals to constrain their exposure to the uninsured, which is consistent with our reduced-form evidence in Section 4 and our demand estimates from the previous section. These are also consistent with evidence on the financial distress and service closure at rural hospitals, particularly in states that did not expand Medicaid access (Hulver *et al.*, 2025; Dranove *et al.*, 2025).

Overall, our structural cost estimates replicate key patterns in financial data when such comparisons are available. Our estimated hospital cost per uninsured admission is $12,406.8, averaged across hospitals for 2014 to 2019. When restricted to hospitals experiencing spillover, this average increases to $12,901.4, which resembles the estimated charitable care cost in Section 4. The estimates shown in Appendix Figure A.10a reveal that the median hospital values its impact on uninsured surplus at about 46% of its average revenue from insured care. We find that private non-profit hospitals place a weight on uninsured surplus that is, on average, 27.5% higher than for-profit hospitals, while public hospitals place a weight that is 39.6% higher. In addition, Figure A.10b shows that spillovers from Medicaid increased this weight by 8.9% among spillover-treated hospitals, rationalizing affected hospital behavior. In addition, the estimates allow us to recover hospitals' implied resource allocation across lines, shown in Appendix Figure A.11. As expected, emergency care accounts for the bulk (62%) of resource allocation, while uninsured scheduled care receives the least (11%). Weighted by acceptance rate, only 3% of capacity is available to uninsured scheduled care, with public hospitals offering the largest share of their capacity to the uninsured. Importantly, 12% of hospitals have at least one line at capacity, implying that admission rates cannot be increased without violating the line's resource constraint. However, fewer than 5% of hospitals have all lines at capacity, suggesting that many hospitals limit acceptance rates to control costs. Therefore, regulation that changes hospital compensation might affect aggregate capacity supply, benefiting both insured and uninsured patients.

8 Counterfactual Policy and Externalities

Using our estimated model, we explore the impact of alternative supply-side policies. First, we study policies that compensate hospitals directly for each uninsured admission, following a structure that resembles existing public insurance programs. Because we focus only on the payment aspect of these policies, not their risk-protection value, we call them pseudo-insurance policies. Second, we explore policies that strengthen nonprofit hospitals' obligation to provide charity care by establishing an optimal minimum target and a market in tradable uninsured-care credits.

This policy analysis has three goals. First, the pseudo-insurance policies reveal how payment differences across patients generate resource externalities, connecting directly to the economics of Section 2. This quantifies the importance of these mechanisms and establishes which uninsurance regime the market operates under. Second, the credit-trading policies reveal where charitable care resources are in deficit and where tax-exempt benefits are in excess. They provide a simple metric to quantify the heterogeneity underlying the tax-to-benefit dispersion shown in Section 3.1. Finally, both policy exercises allow us to benchmark the externalities of uninsurance relative to plausible alternatives, such as restructuring DSH payment schemes, expanding Medicaid coverage, or establishing minimum requirements for hospitals, all of which have been implemented in one form or another.

The following analyses cover the last year and quarter in our sample (2019, Q4). Each policy exercise involves solving counterfactual equilibria, which entails finding a new set of acceptance rates, resource allocations, and patient request probabilities that satisfy each agent's best-response criterion. While the existence of an equilibrium is easy to guarantee, the sufficient conditions for uniqueness developed in Appendix A.3.4 cannot be verified globally. Thus, we refine the equilibria by choosing those nearest in best-response iterations from the status quo, as in Cuesta *et al.* (2024). Finally, throughout, we assume public hospitals are unresponsive to these policy changes and exclude them, meaning their resource allocations and admissions remain at the status quo. This assumption captures public hospital behavior as governed by public authorities, which may respond more slowly, pursue different objectives, or exclude these hospitals from the binding requirements of these policies for political reasons. We present robustness to this assumption in Appendix A.5.4, showing that relaxing public-hospital participation yields nearly identical key findings to those reported below. Additional implementation details are given in Appendix A.5.5.

8.1 Pseudo-Insurance Programs

Pseudo-insurance programs compensate hospitals directly for uninsured care, without forming a formal insurance system for the uninsured or incorporating them into existing structures. By adjusting the payment, the regulator can influence hospitals' acceptance rate and resource alloca-

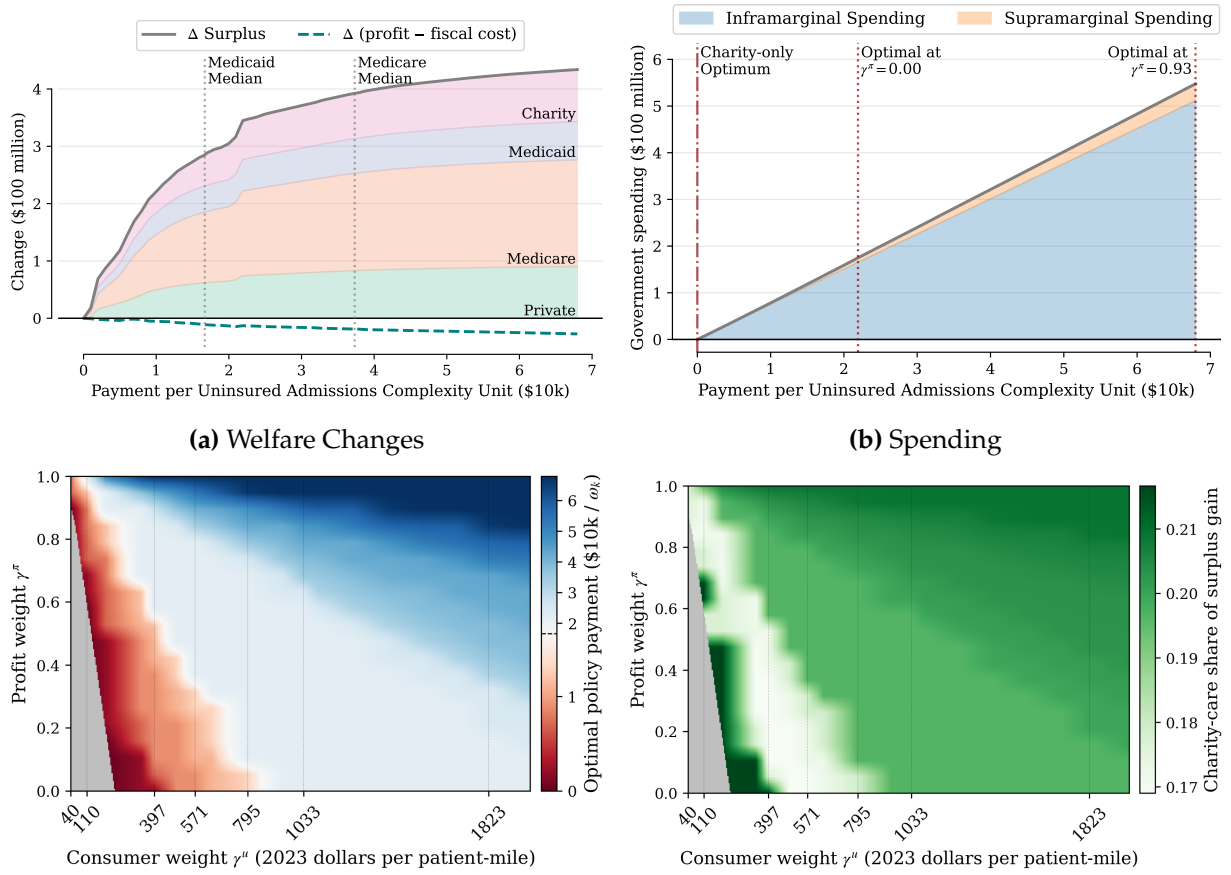
tion decisions, and thus alter all patients' access to care. We consider the case of a regulator that chooses a payment τ per *complexity* unit (DRG weight) of admitted uninsured demand served, and seeks to maximize a weighted welfare criterion:

$$W(\tau;\gamma^u,\gamma^\pi) = \underbrace{\gamma^u S_t(\boldsymbol{a}^*(\tau))}_{\text{patient surplus}} + \underbrace{\gamma^\pi \sum_{h,j} \pi_{hjt}(\boldsymbol{a}^*(\tau),\boldsymbol{b}^*(\tau))}_{\text{hospital profits}} - \underbrace{\left(\sum_{h,j}\sum_{k\in\mathcal{K}_j^u} \lambda_k R_{hkt} a^*_{hjt}(\boldsymbol{\tau})\omega_k \tau\right)}_{\text{fiscal cost}}. \tag{18}$$

In this objective, patient surplus $S_t(\cdot)$ is measured in miles and is dollarized and weighted relative to fiscal cost by the parameter γ^u. Similarly, hospital profits are weighted relative to fiscal costs by the parameter γ^π. The third and final term is the policy's expected fiscal cost. By paying per unit of complexity, the counterfactual policy aligns with Medicare payment structures and links payments to hospitals' surgical costs. In Appendix A.5.2, we present results for an alternative policy that pays hospitals a flat fee per uninsured admission, resembling existing compensating programs, such as DSH. This alternative policy yields a welfare analysis of uninsurance and optimal policies very similar to the ones presented here.

We begin by exploring the policy's impact and optimality when surplus is dollarized and weighted at \$1,033 per equivalent mile of travel, as in Section 7.1.1. Figure 8a shows how consumer surplus and hospital profits net of fiscal costs change as policy payments increase. Net profits are negative and decline as government transfers to hospitals cancel out, leaving only the incremental cost of service. In contrast, as the policy's generosity increases, consumer surplus rises, benefiting both insured and uninsured patients. Given their large numbers, Medicare beneficiaries capture the bulk of the increase in surplus. However, as shown in Appendix Table A.8, per patient, the uninsured benefit most from the policy. For example, at a Medicaid-equivalent rate, the policy increases uninsured patient surplus by \$1,235 per patient, \$617 for Medicaid beneficiaries, \$467 for Medicare beneficiaries, and \$486 for privately insured patients.

These changes, however, come at high fiscal costs, as shown in Figure 8b. For example, at the Medicaid-equivalent level, the program would spend an additional \$132 million per quarter, which, over the year, would amount to 1.47 times the state's total Medicaid DSH budget for 2019 or a 5.3% increase to the state's Medicaid budget. Moreover, at any meaningful payment level, government spending is between 2 and 7 times the surplus it delivers to the uninsured. The source of this wedge becomes clear when separating spending into inframarginal care—the volume of uninsured care delivered without the policy—and the incremental supramarginal care captured from the outside option. As shown, most spending compensates hospitals for care that would have been delivered regardless. Thus, the surplus gains perceived by the uninsured in Figure 8a come from improved sorting rather than from extensive-margin gains. An immediate consequence of this wedge is that, if we ignore other factors, the optimal policy is to provide hospitals with no



(a) Welfare Changes

(b) Spending

(c) Optimal Policy Across Weights

(d) Charity Share of Surplus

Figure 8: Pay-per-Complexity Policies

Notes: These figures display the equilibrium welfare effects associated with compensating hospitals per uninsured admission when γ^u = 1,033 dollars per mile. Policy payments are measured in tens of thousands of dollars per DRG-weight of an uninsured admission. The vertical bars in Figure 8a mark the rate that would match the median payment given by Medicaid and Medicare to hospitals per DRG-weight unit. In Figure 8b, the vertical bars show the smallest γ^π weight on hospital profits that would make that payment level optimal. Supramarginal spending equals the total policy payments for additional uninsured complexity units admitted to within-market hospitals. All remaining spending is inframarginal. Figure 8c shows the optimal policy at a range of γ^π and γ^u. The color scheme is anchored at the median Medicaid rate: blue colors are at or above the Medicaid level, and red colors are below. The gray area marks where the optimal policy is null. The figure marks estimates from the literature on the disutility per dollar traveled by patients, converted to 2023 dollars, as discussed further in Appendix A.5.1. The markers at $1,033 and $1,823 correspond to the adjusted range of estimates from Prager (2020), where the first is our preferred value. The markers at $397 and $795 are based on Gowrisankaran *et al.* (2015), while those at $110 and $571 are from Ho and Pakes (2014). However, we caveat that in this last article, the authors estimate patient-provider responsiveness to insurance rather than patient prices. Finally, Figure 8d shows the share of patient surplus going to charitable care patients in each solution.

additional compensation. This finding aligns with previous research showing that Medicaid is broadly valued by its recipients at levels below its cost and points to the inframarginality of the care delivered as a source of this wedge (Finkelstein *et al.*, 2018, 2019a,b).

This perspective, however, is overturned once the benefits to non-recipients are incorporated into the analysis. Figure 8b and Table 4 show that the surplus-optimal policy level—ignoring any benefit to hospitals—pays hospitals 31.4% over Medicaid, or about 58.7% of Medicare payments. Total patient surplus for the quarter-year increases by 345.0 million, with only 19.6% of that benefit

Table 4: Counterfactual Policy Welfare Effects

	Regulatory	Δ **Welfare** ($100M)				Δ **Available Capacity** (1k bed-days)			
	Price	Patients		Hospitals	Gov.	Emergency	Scheduled		Idle
	τ	All	Charity				Insured	Uninsured	(p.p.)
Panel A: Pay-per-Complexity									
Medicaid-equivalent	1.668	2.849	0.545	1.214	1.320	155.423	-5.358	103.931	-7.460
Medicare-equivalent	3.732	3.923	0.789	2.807	2.995	213.085	-4.997	175.472	-11.265
Optimum at $\gamma^{\pi} = 0$	2.192	3.450	0.678	1.619	1.749	191.112	-2.463	145.155	-9.803
Optimum at $\gamma^{\pi} = 0.93$	6.797	4.339	0.908	5.203	5.476	230.867	-20.289	185.055	-11.619
Panel B: Floor-and-Trade (10% tax-at-risk)									
Unconstrained	11.686	2.516	0.545	-0.172	0.000	119.247	-19.013	110.235	-6.181
Homogeneous	0.251	0.310	0.141	-0.002	0.000	13.831	-5.085	25.268	-0.999

Notes: Welfare is measured as the change relative to the baseline, in hundreds of millions. The Patients columns report the gains for all patients and charitable-care patients, respectively, with "Gov." representing the change in fiscal cost. Available capacity is the sum across hospitals of the product of each line's allocated bed-days and its acceptance rate, measured in thousands. Idle capacity ratio is defined as $(\sum_h B_h - \sum_{h,j} a_{hj} b_{hj}) / \sum_h B_h$ and the change is presented in percentage points. The first column reports the regulatory payment or credit price in $10,000 units. All values are quarterly.

going to the uninsured. At that level, hospital profits increase by 161.9 million, and government spending by 174.9 million. Moreover, any payment level from this level to the median private rate (the upper bound in the figure) can be rationalized with a weight on firm profits between 0 and 94% of the fiscal cost. Relative to the surplus-optimal level, the upper bound on the policy pays hospitals rates equivalent to private insurers, tripling government spending and hospital profits while increasing patient surplus by only 25.7%. Payment at Medicare levels increases government spending by 71.2% relative to the surplus-optimal, increasing surplus by 13.7%.

Table 4 also shows the effect of these policies on available hospital capacity, which equals allocated resources times each line's acceptance rate. We find that any reported level of pseudo-insuring the uninsured increases emergency and scheduled available capacity. In addition, idle capacity decreases. Thus, the evidence is consistent with the market being in the high-uninsurance regime illustrated by the simple theory in Section 2, where uninsurance negatively affects resource allocation and leaves capacity idle. However, a subtlety absent in the homogeneous setting of Section 2 is that the increase in scheduled capacity following the expansion of insurance benefits the uninsured at the expense of the insured, on average.

The optimality of the policies above depends on the relative value the regulator places on patient surplus. However, three findings hold across a wide range of regulatory preferences. First, the optimal policy appears to be to compensate hospitals for uninsured care at rates comparable to Medicaid. Figure 8c shows the optimal policy as we vary the regulator's preference and conversion rate across a range of estimates from the literature. In 4.6% of the graph (gray area), the optimal policy is inaction, while in 16.9% (red area), the optimal policy is to pay hospitals at rates below Medicaid rates. For the remaining 78.5% of the configuration space (blue), the optimal policy is at

or above Medicaid levels, making it the dominant outcome in the analysis. Second, as illustrated in Figure 8d, in any such policy, the majority of the surplus gains from expanding insurance accrues to non-beneficiaries. In particular, the uninsured obtain only 19.7% of the average patient surplus, and never more than a quarter of the total. Finally, a policy analysis centered exclusively on the value of insurance to the uninsured will recommend inaction or very minimal payments (shown in Appendix Figure A.13). Thus, it is the externalities of insurance that make the policy worth its cost.

These pseudo-insurance simulations vary only hospital compensation, which differs from what a true expansion of insurance would entail. For example, even if limited to hospital care, Medicaid expansion would change beneficiaries' networks and the rate at which they demand emergency care relative to scheduled care. In addition, true Medicaid rates are subject to a variety of local adjustments that are more complex than simply a multiple of DRG weights, as simulated here. Finally, from the perspective of the model and hospital operations, such patients would likely no longer be assigned to a separate discriminatory service line. Thus, to evaluate the true predicted impact of a Medicaid expansion subject to the resource externalities we have quantified, we need to extend our counterfactual along various complex, but clear dimensions.

Appendix Table A.10 shows how each of these changes would affect patient surplus. First, actual Medicaid rates would increase patient surplus by an additional 9.4 million relative to the Medicaid-equivalent pseudo-insurance, with 4.6 million of that increase going to charity-care patients. Second, holding patient risk constant but assuming that the relative rate of emergency-to-scheduled need among the formerly uninsured would transition to match that of Medicaid beneficiaries of similar age and gender, patient surplus would nearly double to 542.8 million. Importantly, at this point the uninsured would receive 59.4% of the total surplus gain associated with the policy—enough to justify the fiscal cost even ignoring the benefit to the insured. From this point, restricting patients to Medicaid networks and moving them to the same service line as other patients slightly decreases their surplus by about 1.3%, resulting in a net increase of 535.4 million per quarter for Florida, with 59.8% of the gains falling on the formerly uninsured. However, the largest leap in surplus comes from the rebalancing of emergency risk, which should be considered with caution given two conflicting pieces of evidence: while emergency care rates are lower among Medicaid patients than among the uninsured, evidence from randomized expansions of Medicaid showed an increase in emergency care usage, at least in the short run (Finkelstein *et al.*, 2019a). Nevertheless, these results confirm that—caveat changes in risk—a true Medicaid expansion would not overturn the results of the analysis above.

8.2 Floor-and-Trade Policies

We now turn to policies that leverage hospitals' non-profit tax exemption to incentivize charitable care. Congress and health-policy circles have proposed minimum requirements on nonprofit hos-

pitals' charity-care or community-benefit obligations (U.S. Senate Committee on Finance, Minority Staff, 2007; U.S. Senate Committee on Health, Education, Labor, and Pensions, Majority Staff, 2023; Wofford *et al.*, 2024). The topic has had bipartisan support, with similar concerns originating from both Republican Senator Charles E. Grassley and Democratic Senator Bernie Sanders. These proposals recommend variants of tying hospital tax incentives or non-profit status to their performance, with evidence on past performance of such policies coming from a similar policy in Texas (Kennedy *et al.*, 2010).[40] However, a natural concern is that under binding minimum standards, many hospitals would be unable to provide the required care because demand is not present or because their financial situation is such that doing so would force the hospitals to shut down. An alternative that alleviates some, but not all, of this concern is to implement a floor-and-trade system, as proposed in academic circles by Dranove *et al.* (2015). Under this system, non-profit hospitals must provide a minimum level of charitable care, and a market is created for hospitals to buy and sell charitable-care credits, allowing those with little uninsured demand to finance those with a significant uninsured burden. Such policies have subsequently been discussed in congressional testimony and state legislative deliberations (Levinson, 2023; Georgia House of Representatives, 2019). Importantly, these policies require no additional spending by state or federal governments; instead, they demand larger market interventions.

We focus on two variants of floor-and-trade policies with varying degrees of regulatory discretion. First, we consider an *unconstrained* policy where the regulator can freely allocate obligations and credits across hospitals, allowing it to adapt to each hospital's setting and financial situation. Second, we consider a *homogeneous* policy that must give each hospital the same obligation as a share of its profits. To present a conservative perspective on the potential of these policies and to recognize that non-profit hospitals might provide additional value to the community beyond charitable care, we require our floor-and-trade policies to be incentive-compatible with all existing non-profit hospitals retaining their status. For this reason, we also tie the homogeneous policy floor to profits rather than revenue. In our sample, any meaningful requirement on non-profit supply relative to revenue or expenses would cause some of the largest suppliers of charitable care to drop their non-profit status when the credit price is too low, highlighting the challenges created by a minimum-requirement policy with no trade. We present our main analysis assuming a moderate 10% tax at risk for hospitals, and hold γ^u to our preferred value of \$1,033 per mile and $\gamma^\pi = 1$ to minimize the policy's negative effect on hospitals.

We begin by stating the unconstrained floor-and-trade policy and firm problem. The policy has two elements: a vector of endowments $\boldsymbol{u}$, whose components are each hospital's credit or

[40]Texas imposes minimum requirements on the community benefits provided by non-profit hospitals to qualify for state and local tax exemptions (State of Texas, 2026a; Texas Department of State Health Services, 2025; State of Texas, 2026b). Hospitals must either provide charity and government-sponsored indigent care at a cost equal to the hospital's tax-exempt benefits or provide community benefits of at least 5 percent of net patient revenue, with at least 4 percent for charity and indigent care. Hospitals, however, may also appeal to a needs-based standard to justify that their provision of charitable and indigent care is reasonable relative to their community needs and the hospital's available resources.

obligation in charity-care complexity units; and z^p, the equilibrium trading price. Hence, the expected revenue from credit market trades of non-profit hospital $h \in \mathcal{H}^{np}$ is

$$\mathbb{E}[\pi_h^u(\boldsymbol{a}, \boldsymbol{z})] := z^p \left(\sum_{j \in \mathcal{J}_h} \sum_{k \in \mathcal{K}_j^u} \omega_k a_{hj} \lambda_k R_{hk}(\boldsymbol{a}) - u_h\right)$$

Therefore, all else equal, non-profit hospitals in the floor-and-trade system behave as in the pseudo-insurance policy, with the price per uninsured complexity unit being the credit trading price.

In addition to satisfying the firm's optimal behavior, an equilibrium outcome under the policy must be such that the trading price clears the market and such that endowments and prices make it incentive-compatible for non-profit hospitals to retain their status. These two conditions are:

$$\sum_{h \in \mathcal{H}^{np}} \left(\sum_{j \in \mathcal{J}_h} \sum_{k \in \mathcal{K}_j^c} \omega_k a_{hj} \lambda_k R_{hk}(\boldsymbol{a}) - u_h\right) = 0 \tag{19}$$

$$(1 - \gamma^{tax}) \sum_{j \in \mathcal{J}} \mathbb{E}[\pi_{hj}(a_{hj}^{fp}, b_{hj}^{fp}; \boldsymbol{p}^c)] \leq \sum_{j \in \mathcal{J}} \mathbb{E}[\pi_{hj}(a_{hj}, b_{hj}; \boldsymbol{p}^c + z^p)] - z^p u_h \quad \forall h \in \mathcal{H}^{np} \tag{20}$$

Where $(a_{hj}^{fp}, b_{hj}^{fp})$ is the optimal strategy of the hospital if it were to become a for-profit hospital, with no access to the charity care trading system and had to pay a share γ^{tax} of profit as taxes.[41] Here we denote $\pi_{hj}(a_{hj}^{fp}, b_{hj}^{fp}; \boldsymbol{p}^c + z^p)$ as the hospital-line profit function, taking an extra argument for the payment per uninsured admission complexity unit, where $\boldsymbol{p}^c$ is the status-quo payment rate of charitable care admissions. The following Proposition, proven in Appendix A.5.3, provides an important property of this policy.

Proposition 2. *The unconstrained floor-and-trade market clearing and incentive compatibility conditions hold if and only if total participation profit exceeds total credit price:*

$$\sum_{h \in \mathcal{H}^{np}} \sum_{j \in \mathcal{J}} \left(\mathbb{E}[\pi_{hj}(a_{hj}, b_{hj}; \boldsymbol{p}^c + z^p)] - (1 - \gamma^{tax}) \mathbb{E}[\pi_{hj}(a_{hj}^{fp}, b_{hj}^{fp}; \boldsymbol{p}^c)]\right) \geq z^p \left(\sum_{h \in \mathcal{H}^{np}} \sum_{j \in \mathcal{J}} \sum_{k \in \mathcal{K}_j^c} a_{hj} \lambda_k R_{hk} \omega_k\right) \tag{21}$$

The value of this proposition is twofold. First, it lets us replace $H^{np} + 1$ constraints in the planner's problem with a single scalar constraint. Second, it allows us to switch the planner's control variable from the endowment vector to the trading price: for any credit price satisfying the condition of Equation (21) and any vector of hospital actions that are individually optimal under this price, there is an endowment vector that makes the outcome an equilibrium. Thus, the planner problem is to maximize $W(z; \gamma^u, \gamma^\pi)$ as in Equation (18), subject to the aggregate constraint.

[41] In this analysis, we assume that the hospital's preferences for uninsured care are not a feature of their non-profit status, but rather organizational goals. Relaxing this assumption makes for-profit behavior worse for hospitals, which relaxes the incentive compatibility constraint, and strengthens the policy. This is because most charity care is delivered through the emergency line, which for-profit hospitals cannot avoid without also avoiding profitable emergency patients.

The first row of Panel B in Table 4 shows the impact of the optimal homogeneous floor-and-trade policy. The policy delivers 88.3% of the total surplus value of the Medicaid-equivalent pseudo-insurance policy, but an almost identical surplus to the uninsured. Thus, floor-and-trade is comparatively more targeted towards the uninsured, with 21.6% rather than 19.1% of the consumer surplus gain accruing to the uninsured. The policy also reduces hospital profits by 17.2 million, increases available emergency care capacity slightly less than Medicaid-equivalent pseudo-insurance does, and skews scheduled care slightly more in favor of the uninsured. Idle capacity also decreases in comparable terms, with a 6.2 percentage-point gain in total accessible capacity.

This unconstrained trade policy allows us to visualize where charitable care supply and demand are mismatched. As shown in Appendix Figure A.14, there is substantial geographic dispersion in hospitals' trading positions. For example, the largest buyer of credits in this analysis is AdventHealth of Orlando, buying 23.0% of all traded credits for a total of $2.72 million, while Holy Cross Hospital in Fort Lauderdale is the largest seller, supplying 22.5% of all credits at $2.66 million. These two hospitals also highlight the policy's benefit, as it enables credits to flow between hospitals located 182 miles apart—a distance very few patients are likely to cross for care.

The mismatch between the supply and demand of charitable care is not explained simply by hospitals' location and their local uninsured demand. For example, the second-largest buyer of credits is in Palm Beach County, which has a higher uninsured rate than Broward County, where Fort Lauderdale and the largest seller of credits are located. In particular, the correlation between a hospital's optimal net credit position and its local uninsured rate is -0.19, indicating that higher-burden hospitals sell more credits, as expected. Similarly, it is -0.056 with the number of charitable care patients living within 30 miles of the hospital. However, these correlations are modest, and the economics of the unconstrained floor-and-trade policy mean the largest credit buyers are those with the largest taxable benefits and the lowest uninsured burden. The optimal credit obligations are hard to predict based solely on revenue, the local uninsured rate, or hospital size. Instead, they depend on patients' preferences for each hospital, local competition among hospitals, the availability and capacity of other charitable care providers, and the local uninsured burden. The flexibility of the unconstrained policy allows the regulator to assign credit obligations based on how slack the incentive compatibility constraint is for each hospital to maximize valuable trade. The outcome of the homogeneous policy illustrates this value best, as we discuss next.

Unlike the flexible unconstrained policy, the homogeneous floor-and-trade policy has to find a common target that works for all hospitals. Because of this homogeneity, the result of Proposition 2 breaks, forcing the planner to solve the fully constrained program implied by the set of incentive compatibility constraints and market-clearing conditions. The policy's homogeneity keeps it a scalar rate, with the equilibrium trading price determined as the fixed point of the market-clearing condition. The outcome of this optimal policy is shown in the second row of Table 4. While

the policy improves patient surplus overall, it achieves only 12.3% of the surplus gain of the unconstrained variant. This stark drop highlights the cost of trying to fit the vast heterogeneity in patient demand, competition, and hospital resources with a single target based on hospitals' finances. Appendix Figure A.14 shows that trading positions under this homogeneous policy differ from those in the unconstrained case and have only a mild positive correlation of 0.24. For example, while the largest seller of credits in the unconstrained regime remains a seller under the homogeneous policy, the largest buyer of credits becomes a net seller. This indicates that the homogeneous rule cannot extract more from hospitals with excess supply without harming others or violating incentive compatibility.

Finally, Appendix Table A.9 shows how both floor-and-trade policy outcomes change with different tax-at-risk levels. While the unconstrained policy barely changes with additional regulatory power, the homogeneous policy is tightly linked to it. Doubling the tax-at-risk rate to 20% of profit nearly doubles the consumer surplus gain generated by this policy. However, this outcome remains considerably lower than the unconstrained policy outcome, showing that increasing the tax penalty adds little value relative to better targeting obligations. This also speaks indirectly to the limited evidence from the Texas minimum benefit policy, which documents modest effects from a homogeneous target (Kennedy *et al.*, 2010). The evidence here further states that adding a trading market does not improve much with limited policy targeting.

8.3 Discussion

The analysis of counterfactual policies shows how limited access to insurance for some affects all who share the same resources, regardless of insurance. Hospitals facing higher uninsured demand limit their exposure to payment risk by reallocating resources from emergency to scheduled care. This reallocation forces patients to seek care in less preferred locations, reduces access to local emergency care, and leaves valuable capacity idle. These effects occur despite hospitals placing considerable value on uninsured care; thus, they are not an outcome of pure profit-seeking behavior. These negative effects can be mitigated by compensating hospitals for uninsured care, but the analysis shows clearly that the uninsured on their own would not value this policy more than it would cost to run. However, the externalities produced by resource reallocation would benefit the insured substantially, providing a stronger rationale for such policies. Moreover, the analysis of floor-and-trade policies shows that a large fraction of the consumer surplus gains from direct compensation policies can be achieved at negligible public cost by targeting precisely where non-profit tax benefits exceed local demand for charitable care and where they fall short. However, the analysis also cautions that this policy requires targeting, as the underlying heterogeneity driving the misalignment between non-profit benefits and burden cannot be easily corrected by a uniform policy.

9 Conclusion

This article studies the private and public costs of charitable care, the externalities it generates for insured patients, and the design of regulations that compensate hospitals for providing it. Our empirical strategy exploits a channel through which Medicaid expansion traveled across state lines. Non-profit systems operating in both expansion and non-expansion states expanded charitable care at their non-expansion hospitals, raising the share of bed-days devoted to such care by 1.4 percentage points, while the aggregate volume of charitable care in those states was unchanged. The additional care was therefore drawn from patients who would otherwise have been served elsewhere, producing a contraction in charitable care demand at neighboring hospitals. We use this contraction to measure the burden that uninsured care places on hospitals and find that a marginal charitable care admission significantly reduces hospital profits, almost entirely through foregone revenue from displaced insured patients rather than through the cost of service itself. We then embed this variation in a model of hospitals allocating finite resources across service lines under stochastic demand, which we estimate for Florida. The model recovers a cost of $12,406.8 per uninsured admission and markups over private prices that decline with a hospital's local charitable care burden, indicating that hospitals do not shift that burden onto the privately insured and instead ration access by reallocating resources away from the lines the uninsured can reach. Simulating policies that pay hospitals for uninsured care, we find that the value of those payments to the uninsured falls well short of their fiscal cost, because most of the spending compensates care that would have been delivered regardless. Once we account for effects on non-recipients, however, the conclusion reverses: at Medicaid-equivalent rates, 80.9% of the surplus gain accrues to patients who are already insured, through additional emergency capacity and shorter travel for care. A floor-and-trade market in charitable care credit obligations delivers 88.3% of that surplus gain at no fiscal cost, but only when obligations are unconstrainedly targeted to each hospital. In contrast, a uniform obligation in the floor-and-trade market tied to hospital finances achieves only 12.3% as much.

Several avenues for future research follow. First, we hold hospitals' total capacity fixed within a period and abstract from entry and exit. This is appropriate for the medium-run responses our variation identifies, but it leaves our counterfactuals silent on the long-run margin along which the pressures we document have been most visible, namely the disappearance of emergency and obstetric services in high-uninsured areas. Second, our analysis remains agnostic about why non-profit systems hold their charitable care spending near persistent internal benchmarks. That agnosticism is a strength for identification because our variation requires only that the reallocation occurred, not an account of its motive. However, a model of the non-profit objective that rationalizes this pattern would allow researchers to endogeneize behavior we attribute to firm preferences. Finally, our framework abstracts from dynamic reallocation of patients and the impacts of congestion of quality and wait times, all of which have been found to be relevant independently

in the literature. A comprehensive framework would likely require significant additional data and variation, but it could further our understanding of the externalities of uninsurance and the potential need to expand hospital resources. We view each of these extensions as complementary to, rather than in tension with, the article's central finding that the costs of charitable care are borne well beyond the uninsured, and that policies evaluated only by what they deliver to their intended beneficiaries will systematically undervalue them.

References

Abaluck, J. and Adams-Prassl, A. (2021). What do consumers consider before they choose? Identification from asymmetric demand responses. *The Quarterly Journal of Economics*, **136** (3), 1611–1663.

Abir, M., Nelson, C. D., Chan, E. W., Al-Ibrahim, H., Cutter, C., Patel, K. V. and Bogart, A. (2020). *Critical Care Surge Capacity in U.S. Hospitals: Strategies for Responding to the COVID-19 Pandemic*. Tech. Rep. RB-A164-1, RAND Corporation.

Agarwal, N. and Somaini, P. (2026). Demand analysis under latent choice constraints. *Review of Economic Studies*, **93** (4), 2215–2249.

Agency for Healthcare Research and Quality (2010–2019). HCUP State Inpatient Databases (SID). Healthcare Cost and Utilization Project (HCUP), `https://hcup-us.ahrq.gov/sidoverview.jsp`, state-level inpatient discharge files for data years 2010–2019.

Agency for Healthcare Research and Quality (2025). HCUP weighted summary statistics report: NIS 2023 core file. Healthcare Cost and Utilization Project (HCUP), National Inpatient Sample, `https://hcup-us.ahrq.gov/db/nation/nis/tools/stats/MaskedStats_NIS_2023_Core_Weighted.PDF`.

Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, **84** (3), 488–500.

Almond, D., Currie, J. and Simeonova, E. (2011). Public vs. private provision of charity care? Evidence from the expiration of Hill–Burton requirements in Florida. *Journal of Health Economics*, **30** (1), 189–199.

American Hospital Association (2010–2019). AHA Annual Survey Database. `https://www.ahadata.com/aha-annual-survey-database`, annual survey files for 2010–2019.

Arteaga, F., Kapor, A. J., Neilson, C. A. and Zimmerman, S. D. (2022). Smart matching platforms and heterogeneous beliefs in centralized school choice. *The Quarterly Journal of Economics*, **137** (3), 1791–1848.

ASPR TRACIE (2025). Surge concepts. U.S. Department of Health and Human Services, Administration for Strategic Preparedness and Response Technical Resources, Assistance Center, and Information Exchange, `https://asprtracie.hhs.gov/technical-resources/resource/12877/surge-concepts`, accessed July 7, 2026.

Azevedo, E. M. and Leshno, J. D. (2016). A supply and demand framework for two-sided matching markets. *Journal of Political Economy*, **124** (5), 1235–1268.

Bai, G. and Anderson, G. F. (2016). US hospitals are still using chargemaster markups to maximize revenues. *Health Affairs*, **35** (9), 1658–1664.

Baker, A. C., Larcker, D. F. and Wang, C. C. Y. (2022). How much should we trust staggered difference-in-differences estimates? *Journal of Financial Economics*, **144** (2), 370–395.

Barrett, M. L., Welch, J. and Jiang, J. (2021). *User Guide: An Examination of Expected Payer Coding in HCUP Databases (Updated for 2019 HCUP Data)*. Tech. Rep. 2021-01, Agency for Healthcare Research and Quality, Rockville, MD.

Barseghyan, L., Coughlin, M., Molinari, F. and Teitelbaum, J. C. (2021). Heterogeneous choice sets and preferences. *Econometrica*, **89** (5), 2015–2048.

Basu, S. (2019). Are price-cost markups rising in the United States? A discussion of the evidence. *Journal of Economic Perspectives*, **33** (3), 3–22.

Berry, S., Gaynor, M. and Scott Morton, F. (2019). Do increasing markups matter? Lessons from empirical industrial organization. *Journal of Economic Perspectives*, **33** (3), 44–68.

Bidder, R. M., Krainer, J. R. and Shapiro, A. H. (2021). De-leveraging or de-risking? How banks cope with loss. *Review of Economic Dynamics*, **39**, 100–127.

Bound, J., Braga, B., Khanna, G. and Turner, S. (2020). A passage to America: University funding and international students. *American Economic Journal: Economic Policy*, **12** (1), 97–126.

Brevoort, K., Grodzicki, D. and Hackmann, M. B. (2020). The credit consequences of unpaid medical bills. *Journal of Public Economics*, **187**, 104203.

Brot, Z., Cooper, Z., Craig, S. V. and Klarnet, L. (2024). Is there too little antitrust enforcement in the US hospital sector? *American Economic Review: Insights*, **6** (4), 526–42.

Bulow, J. and Levin, J. (2006). Matching and price competition. *American Economic Review*, **96** (3), 652–668.

Cabral, M. and Mahoney, N. (2019). Externalities and taxation of supplemental insurance: A study of Medicare and Medigap. *American Economic Journal: Applied Economics*, **11** (2), 37–73.

California Department of Public Health (2018). AFL 18-09: Requesting Increased Patient Accommodations Including Medical Surge Tent Use. All Facilities Letter, `https://www.cdph.ca.gov/Programs/CHCQ/LCP/Pages/AFL-18-09.aspx`, authorities: Government Code section 8558 and Title 22, California Code of Regulations, sections 70129, 70217, 70805, and 70809; accessed July 7, 2026.

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

Carey, C. M., Miller, S. and Wherry, L. R. (2020). The impact of insurance expansions on the already insured: The Affordable Care Act and Medicare. *American Economic Journal: Applied Economics*, **12** (4), 288–318.

Centers for Disease Control and Prevention (2026). NHSN connectivity initiative: Hospital bed capacity project instruction book. National Healthcare Safety Network guidance, `https://www.cdc.gov/nhsn/bed-capacity/Connectivity-Initiative_Hospital-Bed-Capacity-Project-Instruction-Book.pdf`, accessed July 7, 2026.

Centers for Medicare & Medicaid Services (2010–2019). Healthcare cost report information system (HCRIS): Hospital cost reports. https://www.cms.gov/data-research/statistics-trends-and-reports/cost-reports, hospital cost-report forms CMS-2552-96 and CMS-2552-10; fiscal-year files for 2010–2019.

Centers for Medicare & Medicaid Services (2026a). Hospital provider cost report. Healthcare Cost Report Information System (HCRIS), https://data.cms.gov/provider-compliance/cost-reports/hospital-provider-cost-report, 2023 final data, file CostReport_2023_Final.csv; accessed August 7, 2026.

Centers for Medicare & Medicaid Services (2026b). National Health Expenditure Accounts: Historical Data. Office of the Actuary, National Health Statistics Group, https://www.cms.gov/data-research/statistics-trends-and-reports/national-health-expenditure-data/historical, tables 1 and 7; data through calendar year 2024.

Chakraborty, I., Goldstein, I. and MacKinlay, A. (2018). Housing price booms and crowding-out effects in bank lending. *The Review of Financial Studies*, **31** (7), 2806–2853.

Chan, T. Y., Hamilton, B. H. and Papageorge, N. W. (2016). Health, risky behaviour and the value of medical innovation for infectious disease. *The Review of Economic Studies*, **83** (4), 1465–1510.

Chandra, A., Gruber, J. and McKnight, R. (2010). Patient cost-sharing and hospitalization offsets in the elderly. *American Economic Review*, **100** (1), 193–213.

Clemens, J. and Gottlieb, J. D. (2017). In the shadow of a giant: Medicare's influence on private physician payments. *Journal of Political Economy*, **125** (1), 1–39.

Coate, S. (1995). Altruism, the Samaritan's dilemma, and government transfer policy. *American Economic Review*, **85** (1), 46–57.

Cohen, R. A. and Briones, E. M. (2026). *Health Insurance Coverage: Early Release of Estimates From the National Health Interview Survey, 2025*. Tech. rep., National Center for Health Statistics, Hyattsville, MD.

Collins, S. R. and Gupta, A. (2024). *The State of Health Insurance Coverage in the U.S.: Findings from the Commonwealth Fund 2024 Biennial Health Insurance Survey*. Tech. rep., The Commonwealth Fund.

Cram, P., Bayman, L., Popescu, I., Vaughan-Sarrazin, M. S., Cai, X. and Rosenthal, G. E. (2010). Uncompensated care provided by for-profit, not-for-profit, and government owned hospitals. *BMC Health Services Research*, **10**, 90.

Cuesta, J. I., Noton, C. E. and Vatter, B. (2024). *Vertical Integration and Plan Design in Healthcare Markets*. Working Paper 32833, National Bureau of Economic Research.

Cullis, J. G., Jones, P. R. and Propper, C. (2000). Waiting lists and medical treatment: Analysis and policies. In A. J. Culyer and J. P. Newhouse (eds.), *Handbook of Health Economics*, vol. 1B, chapter *23*, Amsterdam: Elsevier, pp. 1201–1249.

Cutler, D. M. (1998). Cost shifting or cost cutting? The incidence of reductions in Medicare payments. In J. M. Poterba (ed.), *Tax Policy and the Economy, Volume 12*, MIT Press, pp. 1–28.

Davis, J. M. V., Greenberg, K. and Jones, D. (2026). An experimental evaluation of deferred acceptance: Evidence from over 100 Army officer labor markets. *Econometrica*, **94** (2), 641–662.

De Loecker, J., Eeckhout, J. and Unger, G. (2020). The rise of market power and the macroeconomic implications. *Quarterly Journal of Economics*, **135** (2), 561–644.

Delacrétaz, D., Kominers, S. D. and Teytelboym, A. (2023). Matching mechanisms for refugee resettlement. *American Economic Review*, **113** (10), 2689–2717.

Diamond, R., McQuade, T. and Qian, F. (2019). The effects of rent control expansion on tenants, landlords, and inequality: Evidence from San Francisco. *American Economic Review*, **109** (9), 3365–94.

Dixon, H. and Siciliani, L. (2009). Waiting-time targets in the healthcare sector: How long are we waiting? *Journal of Health Economics*, **28** (6), 1081–1098.

Dobkin, C., Finkelstein, A., Kluender, R. and Notowidigdo, M. J. (2018). The economic consequences of hospital admissions. *American Economic Review*, **108** (2), 308–352.

Doyle, J. J., Jr. (2005). Health insurance, treatment and outcomes: Using auto accidents as health shocks. *Review of Economics and Statistics*, **87** (2), 256–270.

Dranove, D., Garthwaite, C. and Ody, C. (2015). *A floor-and-trade proposal to improve the delivery of charity-care services by US nonprofit hospitals*. Discussion Paper 2015-07, The Hamilton Project, Washington, DC.

—, — and — (2017). How do nonprofits respond to negative wealth shocks? The impact of the 2008 stock market collapse on hospitals. *RAND Journal of Economics*, **48** (2), 485–525.

—, Gaynor, M. and Geddes, E. (2025). *Expecting harm? The impact of rural hospital acquisitions on maternal health care*. NBER Working Paper 34159, National Bureau of Economic Research.

Duggan, M. (2002). Hospital market structure and the behavior of not-for-profit hospitals. *RAND Journal of Economics*, **33** (3), 433–446.

Duggan, M. G. (2000). Hospital ownership and public medical spending. *Quarterly Journal of Economics*, **115** (4), 1343–1373.

Einav, L., Finkelstein, A. and Cullen, M. R. (2010). Estimating welfare in insurance markets using variation in prices. *Quarterly Journal of Economics*, **125** (3), 877–921.

Ellis, R. P. (1998). Creaming, skimping and dumping: Provider competition on the intensive and extensive margins. *Journal of Health Economics*, **17** (5), 537–555.

Finkelstein, A. (2007). The aggregate effects of health insurance: Evidence from the introduction of Medicare. *The Quarterly Journal of Economics*, **122** (1), 1–37.

—, Hendren, N. and Luttmer, E. F. P. (2019a). The value of Medicaid: Interpreting results from the Oregon Health Insurance Experiment. *Journal of Political Economy*, **127** (6), 2836–2874.

—, — and Shepard, M. (2019b). Subsidizing health insurance for low-income adults: Evidence from Massachusetts. *American Economic Review*, **109** (4), 1530–1567.

—, Mahoney, N. and Notowidigdo, M. J. (2018). What does (formal) health insurance do, and for whom? *Annual Review of Economics*, **10** (1), 261–286.

—, Taubman, S., Wright, B., Bernstein, M., Gruber, J., Newhouse, J. P., Allen, H., Baicker, K. and Oregon Health Study Group (2012). The Oregon Health Insurance Experiment: Evidence from the first year. *The Quarterly Journal of Economics*, **127** (3), 1057–1106.

Fournier, G. M. and Campbell, E. S. (1997). Indigent care as quid pro quo in hospital regulation. *The Review of Economics and Statistics*, **79** (4), 669–673.

Frakt, A. B. (2011). How much do hospitals cost shift? A review of the evidence. *Milbank Quarterly*, **89** (1), 90–130.

Frank, R. G. and Salkever, D. S. (1991). The supply of charity services by nonprofit hospitals: Motives and market structure. *The RAND Journal of Economics*, **22** (3), 430–445.

Gandhi, A. (2026). Picking your patients: Selective admissions in the nursing home industry, working paper, June 7, 2026 revision.

Garthwaite, C., Gross, T. and Notowidigdo, M. J. (2018). Hospitals as insurers of last resort. *American Economic Journal: Applied Economics*, **10** (1), 1–39.

Gaynor, M. and Anderson, G. F. (1995). Uncertain demand, the structure of hospital costs, and the cost of empty hospital beds. *Journal of Health Economics*, **14** (3), 291–317.

—, Ho, K. and Town, R. J. (2015). The industrial organization of health-care markets. *Journal of Economic Literature*, **53** (2), 235–84.

— and Vogt, W. B. (2003). *Competition among hospitals*. NBER Working Paper 9471, National Bureau of Economic Research.

Georgia House of Representatives (2019). HR 584: House Study Committee on Exploring a Floor and Trade Charity Care System. House Resolution, 2019–2020 Regular Session; summary in the House End of Session Report, `https://www.legis.ga.gov/api/document/docs/default-source/house-budget-and-research-office-document-library/session-reports/2019_end_of_session_report_by_committee_with_vetoes.pdf?sfvrsn=210de3be_2`, lead sponsor: Representative Todd Jones; created a legislative study committee.

Gerchak, Y., Gupta, D. and Henig, M. (1996). Reservation planning for elective surgery under uncertain demand for emergency surgery. *Management Science*, **42** (3), 321–334.

Ghili, S., Handel, B., Hendel, I. and Whinston, M. D. (2024). Optimal long-term health insurance contracts: Characterization, computation, and welfare effects. *Review of Economic Studies*, **91** (2), 1085–1121.

—, Hendel, I., Sullivan, M. and Whinston, M. D. (2026). What prevents health insurance take-up?: Quantifying frictions and policy responses, working paper.

Glied, S. and Hong, K. (2018). Health care in a multi-payer system: Spillovers of health care service demand among adults under 65 on utilization and outcomes in Medicare. *Journal of Health Economics*, **60**, 165–176.

Godwin, J., Levinson, Z. and Hulver, S. (2023). The estimated value of tax exemption for nonprofit hospitals was about $28 billion in 2020. KFF, https://www.kff.org/health-costs/the-estimated-value-of-tax-exemption-for-nonprofit-hospitals-was-about-28-billion-in-2020/, accessed: 2026-08-07.

Gowrisankaran, G., Marquardt, K. R. and Town, R. (2026). *Quality, Selection, and Congestion: Equilibrium Hospital Outcomes During COVID-19*. Working Paper 35426, National Bureau of Economic Research.

—, Nevo, A. and Town, R. (2015). Mergers when prices are negotiated: Evidence from the hospital industry. *The American Economic Review*, **105** (1), 172–203.

Green, L. (2006). Queueing analysis in healthcare. In *Patient flow: reducing delay in healthcare delivery*, Springer, pp. 281–307.

Green, L. V. (2005). Capacity planning and management in hospitals. In M. L. Brandeau, F. Sainfort and W. P. Pierskalla (eds.), *Operations Research and Health Care: A Handbook of Methods and Applications*, *International Series in Operations Research & Management Science*, vol. 70, Boston, MA: Springer, pp. 15–41.

Greenberg, K., Pathak, P. A. and Sönmez, T. (2024). Redesigning the US Army's branching process: A case study in minimalist market design. *American Economic Review*, **114** (4), 1070–1106.

Greenwood, J., Kircher, P., Santos, C. and Tertilt, M. (2019). An equilibrium model of the African HIV/AIDS epidemic. *Econometrica*, **87** (4), 1081–1113.

Grieco, P. L. E., Murry, C. and Yurukoglu, A. (2024). The evolution of market power in the U.S. automobile industry. *The Quarterly Journal of Economics*, **139** (2), 1201–1253.

Gruber, J. (2008). Covering the uninsured in the United States. *Journal of Economic Literature*, **46** (3), 571–606.

—, Hoe, T. P. and Stoye, G. (2023). Saving lives by tying hands: The unexpected effects of constraining health care providers. *Review of Economics and Statistics*, **105** (1), 1–19.

— and Sommers, B. D. (2019). The Affordable Care Act's effects on patients, providers, and the economy: What we've learned so far. *Journal of Policy Analysis and Management*, **38** (4), 1028–1052.

Hackmann, M. B., Kolstad, J. T. and Kowalski, A. E. (2015). Adverse selection and an individual mandate: When theory meets practice. *American Economic Review*, **105** (3), 1030–1066.

Handel, B. and Ho, K. (2021). The industrial organization of health care markets. In K. Ho, A. Hortaçsu and A. Lizzeri (eds.), *Handbook of Industrial Organization*, vol. 5, chapter *16*, Elsevier, pp. 521–614.

Hatfield, J. W., Kominers, S. D., Nichifor, A., Ostrovsky, M. and Westkamp, A. (2013). Stability and competitive equilibrium in trading networks. *Journal of Political Economy*, **121** (5), 966–1005.

— and Milgrom, P. R. (2005). Matching with contracts. *American Economic Review*, **95** (4), 913–935.

He, D., McInerney, M. and Mellor, J. (2015). Physician responses to rising local unemployment rates: Healthcare provision to Medicare and privately insured patients. *Journal of Health Economics*, **40**, 97–108.

Hendren, N. (2021). Measuring ex ante welfare in insurance markets. *The Review of Economic Studies*, **88** (3), 1193–1223.

— and Sprung-Keyser, B. (2020). A unified welfare analysis of government policies. *The Quarterly Journal of Economics*, **135** (3), 1209–1318.

Hick, J. L., Hanfling, D., Burstein, J. L., DeAtley, C., Barbisch, D., Bogdan, G. M. and Cantrill, S. (2004). Health care facility and community strategies for patient care surge capacity. *Annals of Emergency Medicine*, **44** (3), 253–261.

Ho, K. and Lee, R. S. (2017). Insurer competition in health care markets. *Econometrica*, **85** (2), 379–417.

— and Pakes, A. (2014). Hospital choices, hospital prices, and financial incentives to physicians. *American Economic Review*, **104** (12), 3841–3884.

Hoe, T. P. (2022). Does hospital crowding matter? Evidence from trauma and orthopedics in England. *American Economic Journal: Economic Policy*, **14** (2), 231–262.

Horwitz, J. R. and Nichols, A. (2009). Hospital ownership and medical services: Market mix, spillover effects, and nonprofit objectives. *Journal of Health Economics*, **28** (5), 924–937.

Hulshof, P. J. H., Kortbeek, N., Boucherie, R. J., Hans, E. W. and Bakker, P. J. M. (2012). Taxonomic classification of planning decisions in health care: A structured review of the state of the art in OR/MS. *Health Systems*, **1** (2), 129–175.

Hulver, S., Levinson, Z., Godwin, J. and Neuman, T. (2025). *10 Things to Know About Rural Hospitals*. Tech. rep., KFF.

Internal Revenue Service (2015). Revenue procedure 2015-21: Correction and disclosure procedures under section 501(r). 2015-13 Internal Revenue Bulletin 817, https://www.irs.gov/irb/2015-13_IRB, accessed August 7, 2026.

Internal Revenue Service (2017). Final adverse determination letter, written determination 201731014. https://www.irs.gov/pub/irs-wd/201731014.pdf, letter dated February 14, 2017; released August 4, 2017. Revocation involving failures under Section 501(r)(3).

Internal Revenue Service (2019). Schedule H (Form 990): Hospitals. https://www.irs.gov/pub/irs-prior/f990sh--2019.pdf, representative terminal-year form for the 2010–2019 study period; annual forms were used for the corresponding tax years.

Internal Revenue Service (2026a). Charitable hospitals—general requirements for tax-exemption under section 501(c)(3). https://www.irs.gov/charities-non-profits/charitable-hospitals-general-requirements-for-tax-exemption-under-section-501c3, accessed August 7, 2026.

Internal Revenue Service (2026b). Consequence of non-compliance with section 501(r). https://www.irs.gov/charities-non-profits/consequence-of-non-compliance-with-section-501r, accessed August 7, 2026.

Internal Revenue Service (2026c). Requirements for 501(c)(3) hospitals under the Affordable Care Act—section 501(r). https://www.irs.gov/charities-non-profits/charitable-organizations/requirements-for-501c3-hospitals-under-the-affordable-care-act-section-501r, accessed August 7, 2026.

Joskow, P. L. (1980). The effects of competition and regulation on hospital bed supply and the reservation quality of the hospital. *Bell Journal of Economics*, **11** (2), 421–447.

Jung, K. S., Pinedo, M., Sriskandarajah, C. and Tiwari, V. (2019). Scheduling elective surgeries with emergency patients at shared operating rooms. *Production and Operations Management*, **28** (6), 1407–1430.

Keisler-Starkey, K. and Bunch, L. N. (2020). *Health Insurance Coverage in the United States: 2019*. Current Population Reports P60-271, U.S. Census Bureau, Washington, DC.

Kelen, G. D., Wolfe, R., D'Onofrio, G., Mills, A. M., Diercks, D., Stern, S. A., Wadman, M. C. and Sokolove, P. E. (2021). Emergency department crowding: The canary in the health care system. *NEJM Catalyst Innovations in Care Delivery*, **2** (5).

Kennedy, F. A., Burney, L. L., Troyer, J. L. and Stroup, J. C. (2010). Do non-profit hospitals provide more charity care when faced with a mandatory minimum standard? Evidence from Texas. *Journal of Accounting and Public Policy*, **29** (3), 242–258.

KFF (2024). Health insurance coverage of adults ages 19–64. KFF State Health Facts, `https://www.kff.org/state-health-policy-data/state-indicator/adults-19-64/`, 2024 estimates based on the American Community Survey.

Kurowski, D. (2021). Florida Health Price Finder: Results summary, plan year 2018/2019 data update. Presentation to the State Consumer Health Information and Policy Advisory Council, `https://ahca.myflorida.com/content/download/10499/file/SCHIP_MeetingPacket061621.pdf`, presentation begins on page 5 of the meeting packet.

Levinson, Z. (2023). *Oversight of Nonprofit Hospital Tax-Exempt Status: Background and Key Considerations*. Prepared testimony, KFF, Washington, DC, prepared for the Subcommittee on Oversight, Committee on Ways and Means, U.S. House of Representatives, April 26, 2023.

—, Hulver, S., Godwin, J. and Neuman, T. (2025). Key facts about hospitals. KFF, `https://www.kff.org/health-costs/key-facts-about-hospitals/`.

Liu, T. (2022). *Bargaining with private equity: Implications for hospital prices and patient welfare*. Working paper, MIT Sloan School of Management.

Lopez, E., Neuman, T., Jacobson, G. and Levitt, L. (2020). *How Much More Than Medicare Do Private Insurers Pay? A Review of the Literature*. Tech. rep., KFF.

Mahoney, N. (2015). Bankruptcy as implicit health insurance. *American Economic Review*, **105** (2), 710–746.

Marone, V. R. and Sabety, A. (2022). When should there be vertical choice in health insurance markets? *American Economic Review*, **112** (1), 304–342.

Martin, S. and Smith, P. C. (1999). Rationing by waiting lists: An empirical investigation. *Journal of Public Economics*, **71** (1), 141–164.

Maryland Department of Health (n.d.). COMAR 10.07.01.06-1: Licensed Bed Capacity. Code of Maryland Regulations, `https://regs.maryland.gov/us/md/exec/comar/10.07.01.06-1`, permanently adopted effective October 30, 2000; Accessed July 7, 2026.

Massachusetts Department of Public Health (2022). 105 CMR 130.122: Beds Out of Service and Discontinuation of Service. Code of Massachusetts Regulations, https://www.law.cornell.edu/regulations/massachusetts/105-CMR-130-122, as amended June 24, 2022; accessed July 7, 2026.

McInerney, M., Mellor, J. M. and Sabik, L. M. (2017). The effects of state Medicaid expansions for working-age adults on senior Medicare beneficiaries. *American Economic Journal: Economic Policy*, **9** (3), 408–438.

Medicaid and CHIP Payment and Access Commission (2023). Annual analysis of Medicaid disproportionate share hospital allotments to states. In *Report to Congress on Medicaid and CHIP*, chapter 4, Washington, DC: Medicaid and CHIP Payment and Access Commission, pp. 81–140.

Medicaid and CHIP Payment and Access Commission (2026). *Exhibit 24: Medicaid Supplemental Payments to Hospital Providers by State, FY 2024 (millions)*. MACStats: Medicaid and CHIP data book, Medicaid and CHIP Payment and Access Commission, amounts include federal and state funds and are based on CMS-64 expenditure data; accessed August 7, 2026.

Medicare Payment Advisory Commission (2015). *Report to the Congress: Medicare Payment Policy*. Tech. rep., Medicare Payment Advisory Commission, Washington, DC, see Chapter 3, Figure 3-5, p. 62, and Figure 3-6, p. 64.

Medicare Payment Advisory Commission (2021). *Report to the Congress: Medicare Payment Policy*. Tech. rep., Medicare Payment Advisory Commission, Washington, DC, see Chapter 3, Figure 3-4, p. 69, and Figure 3-8, p. 76.

Medicare Payment Advisory Commission (2026). *Acute Inpatient Services*. Section 6 of *A Data Book: Health Care Spending and the Medicare Program*, Medicare Payment Advisory Commission, Washington, DC, chart 6-9 reports that DSH adjustments accounted for 3.2 percent of FY2024 IPPS payments excluding uncompensated-care payments; Chart 6-10 reports $5.9 billion in FY2024 uncompensated-care payments; and Chart 6-11 reports $105 billion in aggregate FY2024 IPPS payments excluding uncompensated-care payments. Accessed August 8, 2026.

Miller, S., Johnson, N. and Wherry, L. R. (2021). Medicaid and mortality: New evidence from linked survey and administrative data. *The Quarterly Journal of Economics*, **136** (3), 1783–1829.

Moran, K. (2025). Hospital charity care & the Hill–Burton Act, working paper, University of Miami.

Naor, P. (1969). The regulation of queue size by levying tolls. *Econometrica*, **37** (1), 15–24.

National Conference of State Legislatures (2025). Certificate of Need State Laws. National Conference of State Legislatures, https://www.ncsl.org/health/certificate-of-need-state-laws, updated April 2025; accessed July 6, 2026.

Newhouse, J. P. (1970). Toward a theory of nonprofit institutions: An economic model of a hospital. *American Economic Review*, **60** (1), 64–74.

— (1996). Reimbursing health plans and health providers: Efficiency in production versus selection. *Journal of Economic Literature*, **34** (3), 1236–1263.

Ortiz, A., Quattrone, W., Underwood, M., Zmuda, M., Goode, L. S. A., Saur, C., Frkovich, J. and Marcial, L. (2022). *The Development and Management of Community Benefit Insight: A Web-Based Resource That Aggregates US-Based Nonprofit Hospital Community Benefit Spending Data*. Tech. Rep. OP-0076-2207, RTI Press, Research Triangle Park, NC.

Plummer, E., Socal, M. P. and Bai, G. (2024). Estimation of tax benefit of US nonprofit hospitals. *JAMA*, **332** (20), 1732–1740.

Popescu, I., Fingar, K. R., Cutler, E., Guo, J. and Jiang, H. J. (2019). Comparison of 3 safety-net hospital definitions and association with hospital characteristics. *JAMA Network Open*, **2** (8), e198577.

Prager, E. (2020). Healthcare demand under simple prices: Evidence from tiered hospital networks. *American Economic Journal: Applied Economics*, **12** (4), 196–223.

Richards, M. R. and Whaley, C. M. (2026). The current era of health care consolidation. *Journal of Economic Perspectives*, **40** (2), 93–116.

Roth, A. E. (2008). Deferred acceptance algorithms: History, theory, practice, and open questions. *International Journal of Game Theory*, **36** (3–4), 537–569.

— and Peranson, E. (1999). The redesign of the matching market for American physicians: Some engineering aspects of economic design. *American Economic Review*, **89** (4), 748–780.

Russo, A. (2024). Waiting or paying for healthcare: Evidence from the Veterans Health Administration, working paper, Harvard University, August 9, 2024 draft.

Setzler, B. (2025). *Labor and Product Market Power, Endogenous Quality, and the Consolidation of the US Hospital Industry*. Working Paper w34180, National Bureau of Economic Research.

Shapiro, C. and Yurukoglu, A. (2026). Trends in competition in the United States: What does the evidence show? *Journal of Political Economy Microeconomics*, **4** (1), 316–370.

Sharma, R., Stano, M. and Gehring, R. (2008). Short-term fluctuations in hospital demand: Implications for admission, discharge, and discriminatory behavior. *The RAND Journal of Economics*, **39** (2), 586–606.

Sivey, P. (2012). The effect of waiting time and distance on hospital choice for English cataract patients. *Health Economics*, **21** (4), 444–456.

Sloan, F. A., Picone, G. A., Taylor, D. H., Jr. and Chou, S.-Y. (2001). Hospital ownership and cost and quality of care: Is there a dime's worth of difference? *Journal of Health Economics*, **20** (1), 1–21.

State of Texas (2026a). Texas Health and Safety Code, Section 311.045: Community Benefits and Charity Care Requirements. Tex. Health & Safety Code Ann. § 311.045, `https://tcss.legis.texas.gov/resources/HS/htm/HS.311.htm#311.045`, current codification accessed August 10, 2026; originally enacted by Acts 1993, 73rd Leg., ch. 360.

State of Texas (2026b). Texas Tax Code, Section 11.1801: Charity Care and Community Benefits Requirements for Charitable Hospital. Tex. Tax Code Ann. § 11.1801, `https://tcss.legis.texas.gov/resources/TX/htm/TX.11.htm#11.1801`, current codification accessed August 10, 2026.

Supreme Court of Illinois (2010). Provena Covenant Medical Center v. Department of Revenue. 236 Ill. 2d 368, 925 N.E.2d 1131, https://www.illinoiscourts.gov/Resources/3c9f4ff7-1013-4ad8-9cbe-c80165118ee6/107328.pdf, docket No. 107328; decided March 18, 2010.

Tebaldi, P. (2025). Estimating equilibrium in health insurance exchanges: Price competition and subsidy design under the ACA. *Review of Economic Studies*, **92** (1), 586–620.

Texas Department of State Health Services (2025). Annual Statement of Community Benefits Standard—2024: Texas Nonprofit Hospitals. Center for Health Statistics, Hospital Data Collection and Reporting, Form F25-11047, https://www.dshs.texas.gov/sites/default/files/chs/hosp/Forms/ASCBS24NP-revised-7-22-2025.pdf, reporting form for Texas nonprofit hospitals.

Town, R. and Vistnes, G. (2001). Hospital competition in HMO networks. *Journal of Health Economics*, **20** (5), 733–753.

Treasury Inspector General for Tax Administration (2025). *Vague and Outdated Guidance Creates Challenges for Tax-Exempt Hospital Oversight*. Tech. Rep. 2025-100-019, U.S. Department of the Treasury, final Audit Report; accessed August 7, 2026.

U.S. Census Bureau (2012–2019). American Community Survey five-year estimates, table S2701. https://data.census.gov/table?q=S2701, rolling five-year county estimates; health-insurance five-year estimates begin with the 2008–2012 release. Table title: "Health Insurance Coverage Status" for 2012–2014; "Selected Characteristics of Health Insurance Coverage in the United States" for 2015–2019.

U.S. Government Accountability Office (2019). *Medicaid: States' Use and Distribution of Supplemental Payments to Hospitals*. Tech. Rep. GAO-19-603, U.S. Government Accountability Office, accessed August 7, 2026.

U.S. Government Accountability Office (2023). *Tax Administration: IRS Oversight of Hospitals' Tax-Exempt Status*. Testimony GAO-23-106777, United States Government Accountability Office.

U.S. Senate Committee on Finance, Minority Staff (2007). *Tax-Exempt Hospitals: Discussion Draft*. Staff discussion draft, U.S. Senate Committee on Finance, Washington, DC, released by Senator Charles E. Grassley; explicitly not proposed legislation.

U.S. Senate Committee on Health, Education, Labor, and Pensions, Majority Staff (2023). *Executive Charity: Major Non-Profit Hospitals Take Advantage of Tax Breaks and Prioritize CEO Pay Over Helping Patients Afford Medical Care*. Tech. rep., U.S. Senate Committee on Health, Education, Labor, and Pensions, Washington, DC, Bernard Sanders, Chair.

Vatter, B. (2025). Quality disclosure and regulation: Scoring design in Medicare Advantage. *Econometrica*, **93** (3), 959–1001.

Virginia Department of Health (2026). Certificate of Public Need Program. Virginia Department of Health, Office of Licensure and Certification, https://www.vdh.virginia.gov/licensure-and-certification/the-certificate-of-public-need-program/, last updated July 2, 2026; accessed July 6, 2026.

Weisbrod, B. (1988). *The nonprofit economy*. Harvard University Press.

White, C. (2013). Contrary to cost-shift theory, lower Medicare hospital payment rates for inpatient care lead to lower private payment rates. *Health Affairs*, **32** (5), 935–943.

Wing, C., Freedman, S. M. and Hollingsworth, A. (2024). *Stacked Difference-in-Differences*. Working Paper 32054, National Bureau of Economic Research.

Wofford, D., Murdock, K. and Kendall, D. (2024). *A National Charity Care Law to Improve Nonprofit Hospitals*. Tech. rep., Third Way.

Zeckhauser, R. (1970). Medical insurance: A case study of the tradeoff between risk spreading and appropriate incentives. *Journal of Economic Theory*, **2** (1), 10–26.

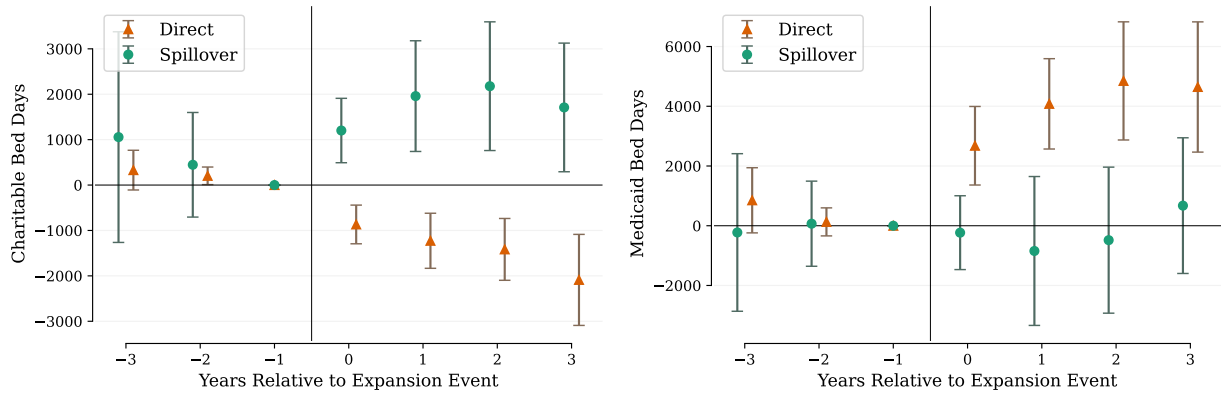
Online Appendix

The Cost of Charitable Care: Resource Externalities and the Value of Insurance

Xuelin Li, Tong Liu, and Benjamin Vatter

A Appendix

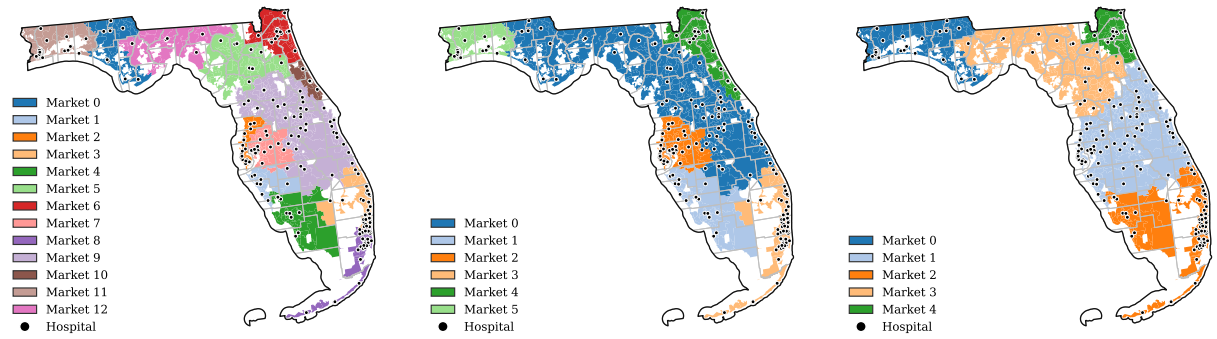
A.1 Additional Figures and Table



(a) Effects on Charity Care Admissions **(b)** Effects on Medicaid Admissions

Figure A.1: Medicaid Expansion Effects on Charity Care and Medicaid Admissions

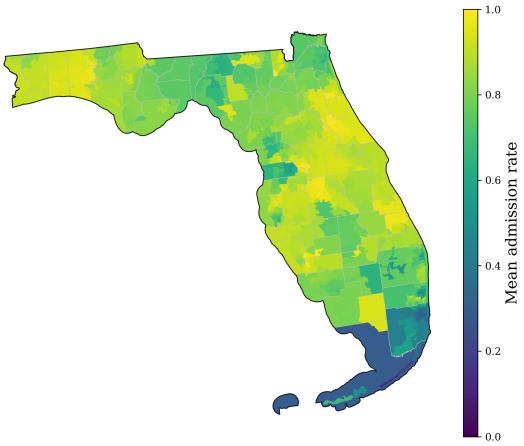
Notes: The two figures show the estimated effects of Medicaid expansion on hospitals in expansion states (triangles), and those in non-expansion states but who are linked to expansion states via system membership (circles). These estimates match the event-study variant of the stacked event-study panel regression (2) and support the evidence in the main text Section 4.1



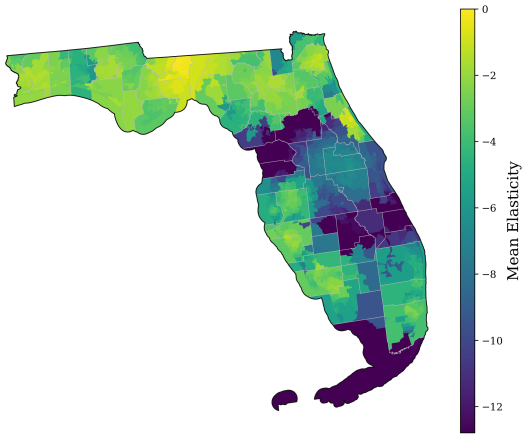
(a) Emergency Care **(b)** Scheduled Insured **(c)** Scheduled Uninsured

Figure A.2: Market Definitions by Service Line

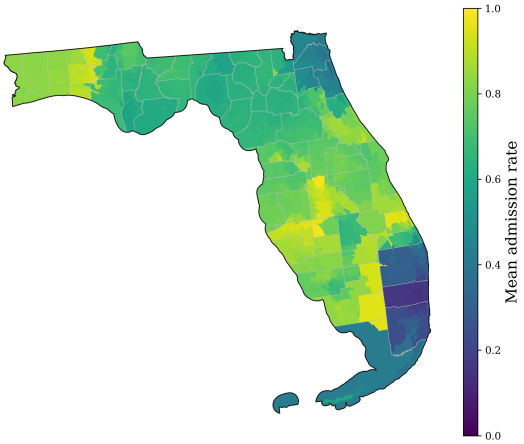
Notes: These figures show how HRRs are grouped by market in our structural analysis. Our algorithm clusters contiguous HRRs such that at least 90% of demand originating within the market is retained within the market. White regions in the interior of Florida correspond to zip codes where we do not observe sufficient consumers in our data to meet masking thresholds.



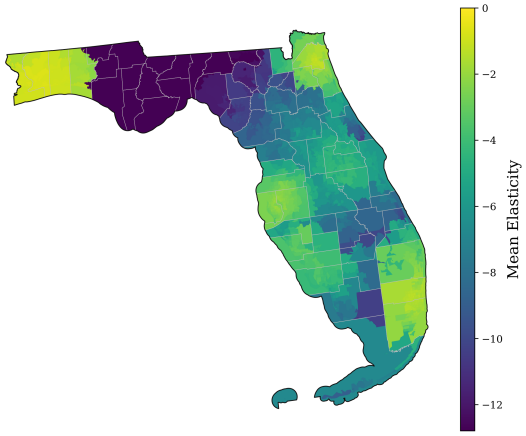
(a) Acceptance Prob– Emergency



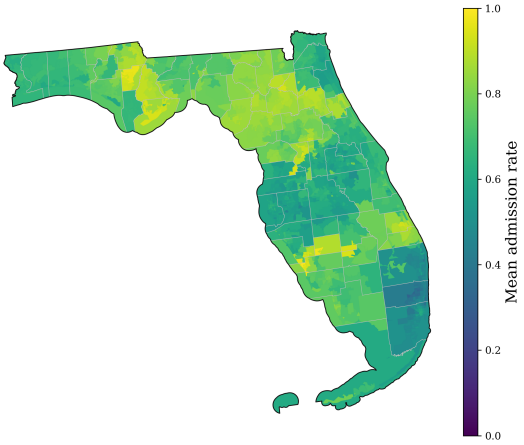
(b) Demand Elasticity – Emergency



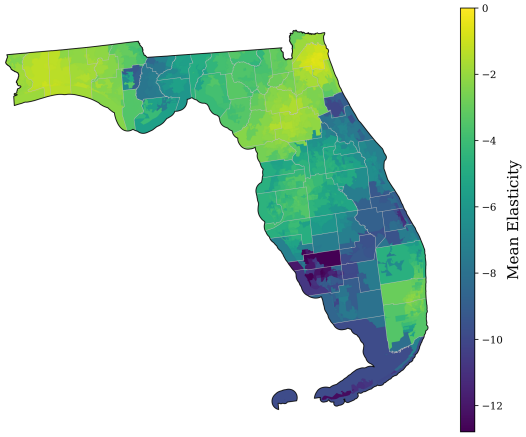
(c) Acceptance Prob – Scheduled Insured



(d) Demand Elasticity – Scheduled Insured



(e) Acceptance Prob. – Scheduled Uninsured



(f) Demand Elasticity – Scheduled Uninsured

Figure A.3: Preferred Admission Probabilities and Distance Elasticities

Notes: The acceptance probability figures on the left show the mean across ZCTAs within the state of the estimated probability that patients are admitted to their most preferred hospitals. The figures on the right show the mean estimated elasticity of travel distance in each ZCTA.

Table A.1: Medicaid Expansion Effects on Charitable Care

	Panel A: Direct Treatment				**Panel B: Spillover Treatment**			
	Charitable		State Charitable		Charitable		State Charitable	
	Admissions	DRG	Admissions	Spending	Admissions	DRG	Admissions	Spending
Expansion	-348.113***	-446.040***	-4681.280***	-108.198***	256.221***	287.658**	-35.697	33.693
Effect (α)	(72.171)	(82.103)	(853.197)	(13.583)	(91.020)	(142.274)	(809.488)	(21.040)
Mean Dep.	678.005	809.760	8767.576	285.124	1120.333	1267.883	19225.818	547.119
N	2081	2081	959	771	3065	3065	763	613
R^2	0.948	0.951	0.991	0.961	0.978	0.967	0.996	0.944

Notes: Admission and DRG-based measures include only admissions from patients residing in the same state as the associated hospital. "Charitable Admissions " corresponds to the count of charitable care admissions. "Charitable DRG" corresponds to the sum of DRG weights of charitable care patients. "State Charitable Admissions" corresponds to the total number of charitable care admissions in the state. "State Charitable Spending" corresponds to the total reported expenditure on charitable care by non-profit systems in each state, as reported in their tax returns. The first two columns of each panel are estimates from Equation (2) with two-way cluster-robust standard errors at the event and hospital levels. The state-level outcomes are estimated using Equation (3), with heteroskedasticity-robust standard errors. For state-level regressions, only one observation is kept per state-year. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.2: Spillover Effects on the Treated: Crowd-outs

	Panel A: Spillover-Treated Hospitals Admissions (log bed days)		
	Total Admitted (1)	Emergency Admissions (2)	Non-Emergency Admissions (3)
Expansion Effect (α)	0.016 (0.046)	0.198 (0.137)	-0.092 (0.071)
N 3065	Mean Dep. 10.643 R^2 0.956	Mean Dep. 9.873 R^2 0.948	Mean Dep. 9.610 R^2 0.939

	Panel B: Spillover-Treated Hospitals Insurance Composition (bed days)							
	Total Admitted		Emergency Admissions			Non-Emergency Admissions		
	Gov. (1)	Private (2)	Charitable (3)	Private (4)	Gov. (5)	Charitable (6)	Private (7)	Gov. (8)
Expansion Effect (α)	345.021 (2316.662)	-1473.660 (1453.212)	845.229* (486.538)	-861.456 (725.756)	411.406 (2015.353)	407.643* (229.791)	-612.204 (884.010)	-66.385 (1224.427)
Mean Dep.	50885.332	15247.444	3163.407	6831.593	32872.371	774.852	8415.852	18012.963
N	3065	3065	3065	3065	3065	3065	3065	3065
R^2	0.996	0.989	0.952	0.986	0.992	0.913	0.986	0.984

	Panel C: Local Untreated Hospital Admissions (bed days)			
	Charitable (1)	Gov. (2)	Private (3)	Total (4)
Expansion Effect (α)	-1116.019* (592.085)	1146.643 (2200.531)	2392.507** (1090.029)	2911.283 (2791.728)
N 724.000	Mean Dep. 6123.541 R^2 0.948	Mean Dep. 53604.418 R^2 0.987	Mean Dep. 21779.393 R^2 0.977	Mean Dep. 83970.578 R^2 0.992

Notes: This table presents the estimated effect of Medicaid expansion on hospitals in non-expansion states. It is a companion to the main-text Table 2. Panel A shows the total change in admissions, measured in log bed-days. Panel B shows admissions (measured in bed-days) by insurance source, with "Gov." denoting the sum of Medicare and Medicaid. Panels A and B are estimates of Equation (2) with standard errors clustered at the event and hospital levels (two-way). Panel C shows the effect on local untreated hospitals, defined as all system hospitals located within 15 miles of a spillover-treated hospital. The estimates are based on Equation (4), with standard errors clustered at the HRR level. All admissions are from patients residing within the same state as the associated hospital. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.3: Medicaid Expansion Effects on Charitable Care by MDC

	Nervous	Eye	ENT	Respiratory	Circulatory
Expansion Effect (α)	52.959 (45.835)	3.547 (2.259)	18.137*** (6.978)	152.857*** (42.371)	154.113 (99.162)
Mean Dep.	261.259	9.593	61.222	405.370	533.148
N	3065	3065	3065	3065	3065
R^2	0.930	0.871	0.949	0.941	0.939
	Digestive	Hep./Pancreas	MSK	Skin/Breast	Endocrine
Expansion Effect (α)	147.007*** (53.025)	88.712*** (31.990)	22.634 (37.997)	38.772 (26.075)	41.352 (31.156)
Mean Dep.	550.333	286.037	227.370	245.704	171.519
N	3065	3065	3065	3065	3065
R^2	0.968	0.959	0.964	0.969	0.936
	Kidney/Urinary	Male Repro.	Female Repro.	Pregnancy	Newborns
Expansion Effect (α)	66.676** (30.672)	0.339 (4.151)	36.169** (16.879)	9.083 (26.156)	91.677 (88.304)
Mean Dep.	173.741	22.185	61.852	71.481	145.815
N	3065	3065	3065	3065	3065
R^2	0.934	0.783	0.898	0.938	0.726
	Blood/Immune	Neoplasms	Infectious	Mental Health	Substance Use
Expansion Effect (α)	30.130** (12.152)	29.676** (13.631)	110.223 (84.794)	23.492 (77.038)	26.505 (23.768)
Mean Dep.	57.741	37.852	249.889	101.889	93.111
N	3065	3065	3065	3065	3065
R^2	0.888	0.878	0.887	0.858	0.853
	Injuries	Burns	Health Factors	Trauma	HIV
Expansion Effect (α)	15.686 (15.105)	1.833 (2.764)	65.434 (44.866)	-33.692 (27.464)	38.224** (19.172)
Mean Dep.	101.000	1.926	26.630	12.889	28.704
N	3065	3065	3065	3065	3065
R^2	0.953	0.964	0.665	0.878	0.815

Notes: This table presents the estimated effect of Medicaid expansion on spillover-treated hospitals. Each cell represents an estimate from an independent regression, where the dependent variable is the total length of stay for charitable care admissions associated with the given MDC. ENT refers to Ears, Nose, and Throat. MSK refers to the musculoskeletal system. Hep refers to the hepatobiliary system. All regressions include admissions only from patients residing in the same state as the associated hospital. Standard errors are clustered at the hospital and event levels (two-way). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.4: Spillover Exposure Effects

	Panel A: Charitable Care at Spillover-Treated			**Panel B: Admissions at Local-Untreated** (bed days)			
	Admissions	Bed Days	Bed Day Share	Charitable	Gov.	Private	Total
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Expansion	40.262	-231.764	-0.002				
Effect (α)	(66.117)	(277.356)	(0.005)				
$\hookrightarrow \times$ Exposure	11.214**	77.091***	0.001***				
	(5.473)	(27.392)	(0.000)				
Local				-1025.090**	845.144	1845.665**	2038.912
Exposure (α^e)				(515.306)	(2044.127)	(876.103)	(2370.865)
Mean Exposure	18.240	18.240	18.240	3.605	3.605	3.605	3.605
Std Exposure	10.228	10.228	10.228	8.137	8.137	8.137	8.137
Mean Dep.	1120.333	3938.259	0.055	5905.180	52316.129	21043.859	81643.383
N	3065	3065	3065	764	764	764	764
R^2	0.978	0.962	0.910	0.949	0.988	0.977	0.992

Notes: This table is a companion to main text Table 3. Panel A presents how exposure, measured as millions of dollars spent by systems in Medicaid expansion states in 2013, is associated with charitable care at spillover-treated hospitals. The estimates follow regression Equation (2), with standard errors clustered at the event and hospital level. Panel B presents the effect of the instrument—Local Exposure—as defined in the main text, on admissions. These correspond to estimates of Equation (4) with HRR-clustered standard errors. Column 4 is thus the first stage of our IV regression in Equation (5). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.5: System Targets and Reaction to Expansion Effects

	Panel A: Charitable Care Admissions at Spillover-Treated Hospitals		**Panel B: Expansion Effect on System Charity Spending**	
	Exposure Metric:		Treated Definition:	
	Share of Charitable Expenditure in Expansion States	Change in Medicaid Admissions in Expansion States	Spillover	Direct
Exposure Tercile 1	99.227	73.400		
	(84.486)	(72.399)		
Exposure Tercile 2	330.872**	320.427**		
	(134.904)	(132.761)		
Exposure Tercile 3	150.091	254.215*		
	(133.504)	(130.412)		
Expansion Effect (α)			-1.581	-16.956***
			(9.854)	(2.478)
N	3065	3065	920	11193
R^2	0.978	0.978	0.966	0.918
Mean Dep.	1120.333	1120.333	119.510	26.957
Mean Exposure	0.370	3608.114		

Notes: This table shows evidence on the response of systems to Medicaid expansions. Panel A shows how charitable care admissions at spillover-treated hospitals change in response to spillover treatment, where we have allowed for heterogeneity across terciles of two different exposure metrics. The share of charitable care expenditure in expansion states was computed in 2013. The change in Medicaid admissions in the expansion corresponds to the difference between the average in 2012–2013 and 2014–2015. Both exposure metrics are defined at the system level. Panel B presents the effect of Medicaid expansion on spillover-treated and directly-treated hospitals, where the outcome is the system's total spending on charitable care, as reported to the IRS. All regressions are based on the design in Equation (2) and include standard errors clustered at the system and event (two-way). Control units are weighted to sum to one in each event, following Wing et al. (2024). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.6: Grant Response to Charitable Care Spending

	Panel A: System State				**Panel B: System**			
	Gov. Grants		All Grants		Gov. Grants		All Grants	
	Contemp.	Lag.	Contemp.	Lag.	Contemp.	Lag.	Contemp.	Lag.
spending	0.144**	0.027	0.293***	0.112*	0.270***	0.086*	0.824***	0.245
	(0.072)	(0.039)	(0.099)	(0.064)	(0.074)	(0.045)	(0.249)	(0.152)
N	11922	10217	11922	10217	11386	11386	11386	11386
R^2	0.932	0.943	0.971	0.971	0.881	0.878	0.813	0.807
Mean Dep.	5.512	5.512	12.866	12.866	5.886	5.886	13.738	13.738
Mean Spending	10.971	11.489	10.971	11.489	11.472	10.339	11.472	10.339

Notes: This table shows the relationship between grants given to hospitals and their charitable care spending. Panel A are estimates of α from the regression:

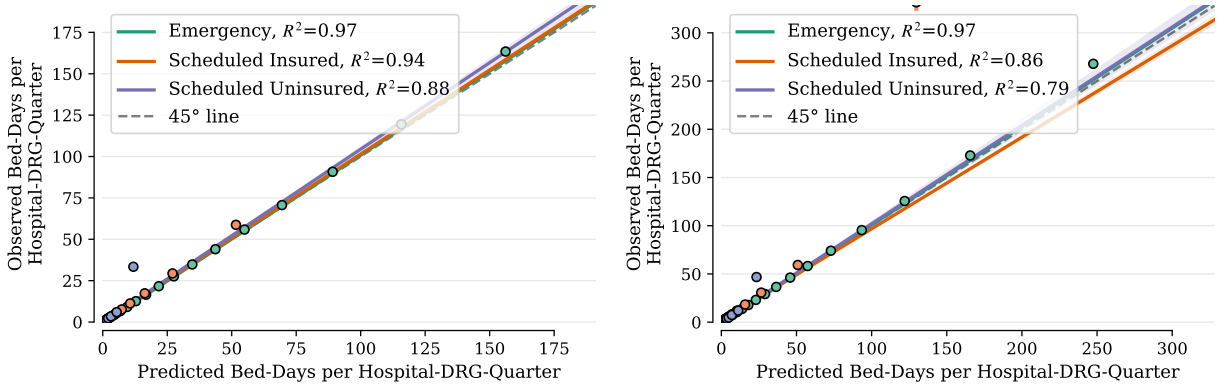
$$y_{smt} = \alpha c_{smt} + \gamma_{sm} + \delta_t + \epsilon_{smt}$$

where s indexes a system, c_{smt} is the system's spending in state m in year t, γ_{sm} are system-state fixed effects, δ_t are year fixed effects, and errors are clustered at the system level. Panel B is the associated estimate when aggregating grants and spending at the system level and adjusting the regression to have system and year fixed effects. Standard errors in Panel B are heteroskedasticity-robust. "Contemp." means comptemporaneous charitable care spending is used in c_{smt}, while "Lag." means last year's spending is used. Government grants exclude payments for services rendered, such as DSH payments. All Grants include direct donations and exclude any payments for services rendered. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.7: Demand Model Estimates of Disutility of Travel Parameters

	Emergency Care													Scheduled Insured						Scheduled Uninsured				
Market	0	1	2	3	4	5	6	7	8	9	10	11	12	0	1	2	3	4	5	0	1	2	3	4
Age: 6–17	−0.219	2.455	−2.652	1.894	1.060	0.540	−0.336	0.854	−0.235	0.662	−0.819	−0.226	0.174	0.372	0.209	−0.030	−0.016	−0.599	−0.260	−0.782	0.089	−4.253	0.841	3.852
	(0.129)	(0.352)	(0.349)	(0.116)	(0.115)	(0.186)	(0.094)	(0.192)	(0.231)	(0.101)	(0.160)	(0.040)	(0.074)	(0.039)	(0.067)	(0.061)	(0.059)	(0.093)	(0.020)	(0.350)	(0.358)	(0.244)	(0.677)	(0.939)
Age: 18–24	−1.616	1.582	−14.309	−3.356	0.507	−0.966	−3.098	−3.121	1.272	−3.219	−3.240	−3.116	−1.062	−3.448	−2.371	−7.025	−1.082	−3.717	−2.308	−1.055	−0.520	−3.074	−1.261	4.598
	(0.135)	(0.299)	(0.330)	(0.107)	(0.100)	(0.147)	(0.077)	(0.169)	(0.176)	(0.089)	(0.161)	(0.038)	(0.073)	(0.038)	(0.067)	(0.058)	(0.050)	(0.087)	(0.023)	(0.275)	(0.311)	(0.179)	(0.603)	(0.673)
Age: 25–34	−2.431	−0.527	−14.802	−4.385	−0.519	−1.450	−5.346	−3.475	3.466	−4.346	−3.724	−4.599	−1.407	−3.494	−2.614	−7.420	−3.161	−4.832	−2.842	−1.076	−0.934	−3.522	−0.681	5.065
	(0.118)	(0.293)	(0.316)	(0.100)	(0.097)	(0.140)	(0.074)	(0.157)	(0.163)	(0.084)	(0.139)	(0.035)	(0.065)	(0.036)	(0.063)	(0.055)	(0.049)	(0.081)	(0.022)	(0.259)	(0.308)	(0.174)	(0.589)	(0.660)
Age: 35–44	−2.572	−2.226	−16.938	−5.735	−1.450	−2.053	−6.345	−4.097	2.445	−5.703	−4.679	−5.279	−1.931	−2.008	−1.476	−4.161	−0.010	−3.422	−1.358	−1.162	0.322	−3.271	−0.566	4.311
	(0.109)	(0.291)	(0.314)	(0.100)	(0.095)	(0.140)	(0.074)	(0.154)	(0.161)	(0.083)	(0.141)	(0.035)	(0.063)	(0.036)	(0.062)	(0.054)	(0.049)	(0.081)	(0.019)	(0.262)	(0.308)	(0.176)	(0.592)	(0.662)
Age: 45–54	−3.075	−4.716	−19.396	−7.642	−2.668	−2.983	−7.862	−6.334	1.154	−7.169	−6.364	−5.544	−2.681	−2.048	−1.790	−3.171	1.943	−3.544	−1.405	−0.801	−0.121	−3.211	−1.065	4.082
	(0.103)	(0.287)	(0.311)	(0.096)	(0.093)	(0.136)	(0.071)	(0.151)	(0.159)	(0.081)	(0.134)	(0.032)	(0.062)	(0.034)	(0.060)	(0.051)	(0.048)	(0.077)	(0.018)	(0.258)	(0.307)	(0.174)	(0.591)	(0.657)
Age: 55–64	−3.526	−6.572	−21.969	−9.240	−3.637	−3.424	−10.298	−7.974	0.865	−8.557	−8.096	−5.797	−3.043	−2.577	−2.788	−5.228	2.036	−4.086	−1.633	−1.381	−0.617	−3.503	−1.413	2.795
	(0.099)	(0.285)	(0.311)	(0.096)	(0.092)	(0.135)	(0.072)	(0.151)	(0.158)	(0.081)	(0.133)	(0.030)	(0.062)	(0.033)	(0.058)	(0.051)	(0.048)	(0.074)	(0.017)	(0.263)	(0.308)	(0.175)	(0.591)	(0.675)
Age: 65–74	−4.260	−9.028	−25.092	−11.900	−5.386	−4.043	−12.745	−10.123	−2.112	−10.036	−10.537	−6.968	−3.488	−3.421	−4.204	−7.388	0.968	−5.265	−2.163	−4.344	0.682	−2.087	0.845	−1.211
	(0.100)	(0.284)	(0.311)	(0.096)	(0.092)	(0.136)	(0.073)	(0.152)	(0.158)	(0.081)	(0.135)	(0.031)	(0.064)	(0.033)	(0.058)	(0.050)	(0.048)	(0.074)	(0.018)	(0.326)	(0.321)	(0.204)	(0.651)	(0.718)
Age: 75+	−5.692	−14.216	−30.158	−19.240	−7.557	−5.179	−15.558	−13.112	−5.905	−12.618	−14.948	−10.354	−4.584	−5.133	−6.771	−9.757	−2.156	−7.111	−4.944	−3.735	0.951	−2.524	0.147	−23.579
	(0.097)	(0.282)	(0.311)	(0.093)	(0.091)	(0.135)	(0.073)	(0.150)	(0.157)	(0.080)	(0.134)	(0.032)	(0.063)	(0.034)	(0.059)	(0.051)	(0.048)	(0.077)	(0.023)	(0.385)	(0.324)	(0.206)	(0.651)	(1.455)
Male	−2.964	−8.675	−10.858	−11.902	−6.868	−3.142	−4.099	−12.081	−22.906	−7.126	−6.596	−1.196	−0.903	−3.531	−6.271	−4.815	−5.489	−5.105	−1.017	−2.458	−6.684	−4.483	−3.414	−10.717
	(0.096)	(0.281)	(0.310)	(0.088)	(0.091)	(0.131)	(0.069)	(0.144)	(0.156)	(0.078)	(0.132)	(0.027)	(0.056)	(0.032)	(0.059)	(0.049)	(0.047)	(0.073)	(0.018)	(0.257)	(0.306)	(0.173)	(0.588)	(0.667)
Female	−3.490	−9.771	−12.319	−13.450	−7.647	−4.532	−6.480	−13.372	−24.327	−8.187	−7.147	−2.534	−1.233	−4.751	−7.498	−7.367	−8.088	−6.439	−1.704	−2.693	−7.574	−6.980	−4.130	−11.742
	(0.097)	(0.281)	(0.311)	(0.090)	(0.091)	(0.132)	(0.070)	(0.145)	(0.155)	(0.079)	(0.132)	(0.029)	(0.057)	(0.032)	(0.059)	(0.049)	(0.047)	(0.073)	(0.017)	(0.261)	(0.306)	(0.170)	(0.589)	(0.649)
Low income	−0.920	−4.786	0.588	−2.548	−3.582	−2.021	−2.719	−3.293	−7.247	−0.934	1.593	−1.483	−2.062	−1.092	−1.703	−0.429	−1.830	0.299	−0.381	−0.578	−1.619	−2.935	−0.765	−1.671
	(0.038)	(0.053)	(0.051)	(0.035)	(0.024)	(0.036)	(0.024)	(0.047)	(0.043)	(0.024)	(0.049)	(0.013)	(0.028)	(0.012)	(0.017)	(0.016)	(0.009)	(0.027)	(0.009)	(0.058)	(0.039)	(0.039)	(0.095)	(0.127)
N	99082	248260	555772	549865	335779	224124	487632	489475	560721	538205	158142	250850	109635	492068	433851	521488	533472	352353	158523	5918	36954	72212	10892	11401
Clusters	76143	164710	356222	346137	238748	178574	336692	344649	348140	368549	104370	183542	90308	357615	275023	358210	346066	231152	112947	5782	35305	63045	10462	10939

Notes: Standard errors clustered at the choice event level in parentheses.



(a) Fit on Estimation Data **(b)** Fit on Validation Data

Figure A.4: Demand Estimation Fit

Notes: These figures show the fit of the demand model on the data used for estimation and on a random 20% sample of market-line-quarters, which we exclude from estimation. To predict on the training data, we use the estimated patient preference parameters and recover admission rates by one iteration of the admission fixed point. We note that the admission fixed point matches the aggregate length of stay at the hospital-line; hence, we plot the fit at a significantly thinner level: the hospital-line-diagnosis group level.

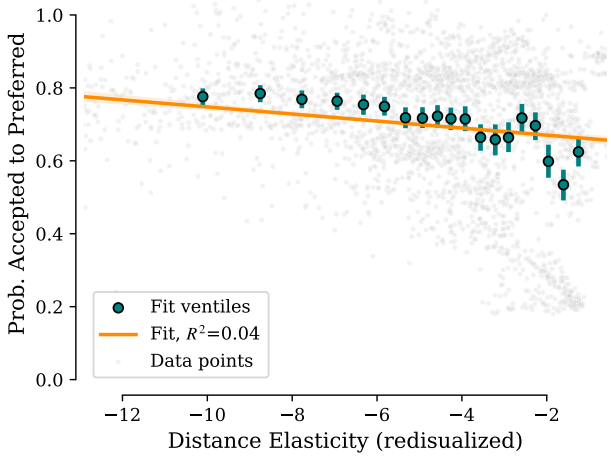
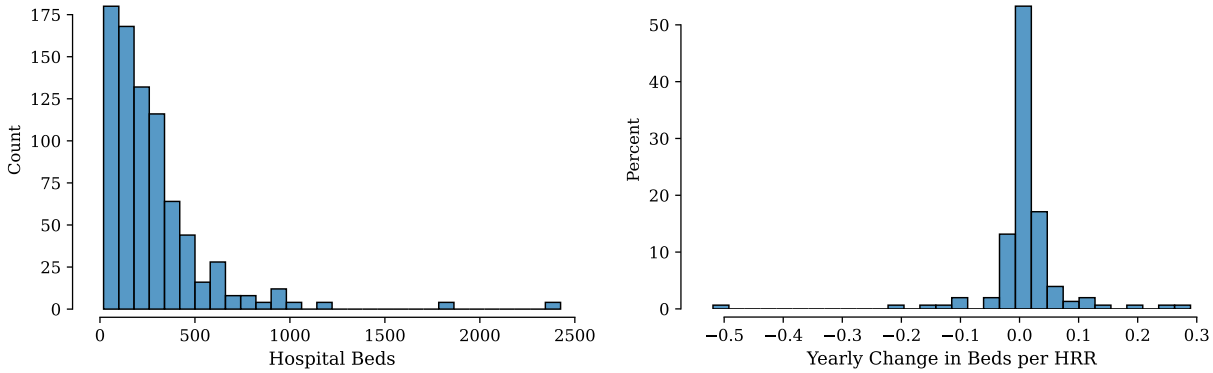


Figure A.5: Correlation Between Estimated Acceptance Rate and Distance Elasticity

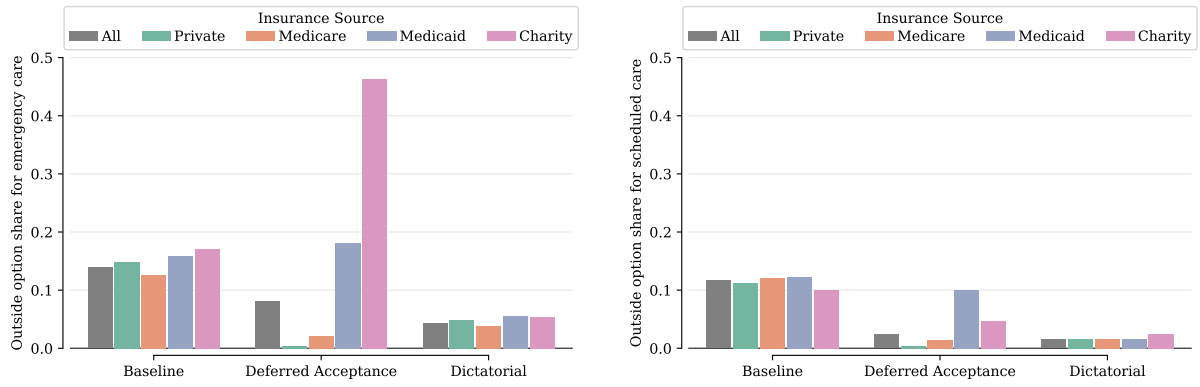
Notes: This figure shows the correlation between the estimated probability that a patient is admitted to her most preferred hospital and her distance elasticity.



(a) Beds per Hospital, 2014 **(b)** Yearly Change in Beds

Figure A.6: Distribution of Beds in Florida – Included Sample Only

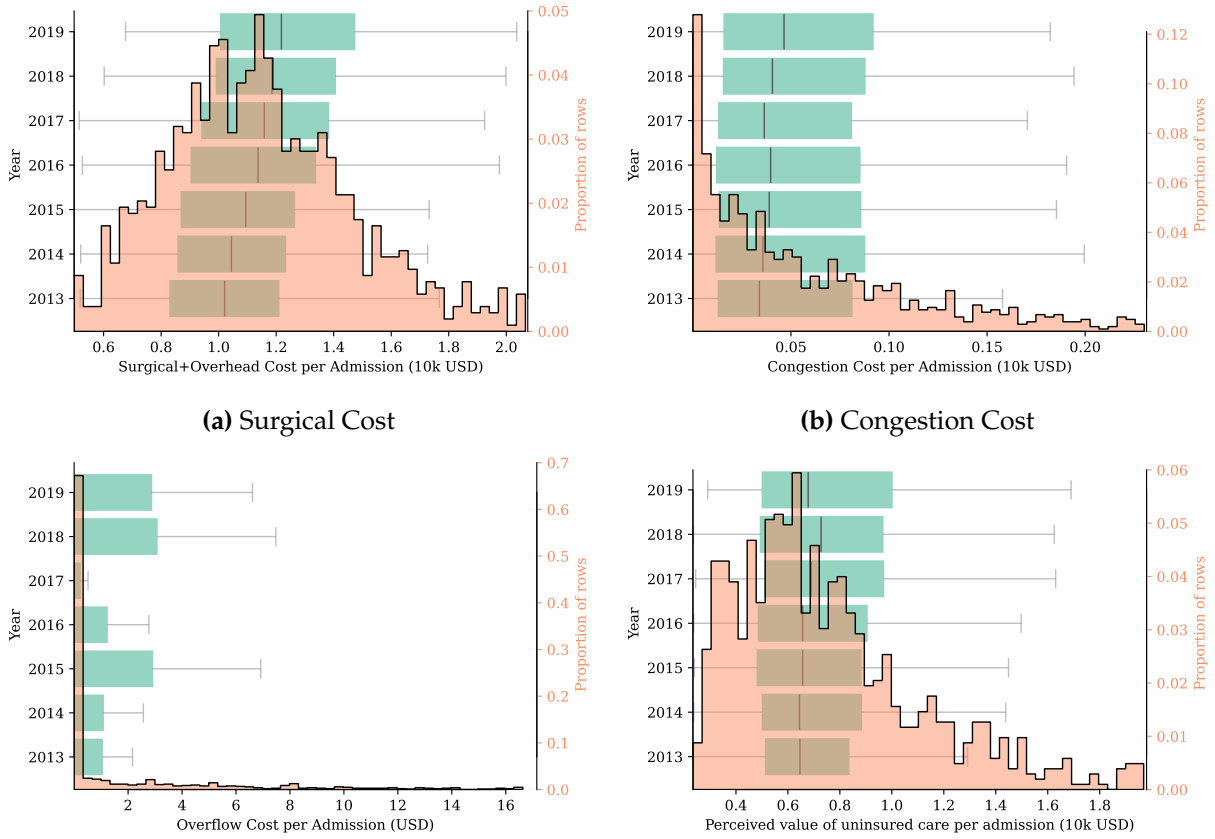
Notes: These figures show the distribution of beds in hospitals included in our analysis. Figure A.6a corresponds to a snapshot in 2014, while A.6b shows the yearly change at the HRR level. The panel of HRR-by-year is balanced—there are no markets that experience a change from zero to positive or vice versa.



(a) Emergency Care **(b)** Scheduled Care

Figure A.7: Outside-of-market Shares Under Alternative Allocative Regimes

Notes: These figures show the share of demand by insurance type and emergency status that is allocated to the out-of-market option under alternative allocative regimes, as discussed in Section 7.1.1. Note that we exclude here patients whose preference is to be assigned to the outside option, as all direct allocation regimes would assign them out of the market.



(a) Surgical Cost

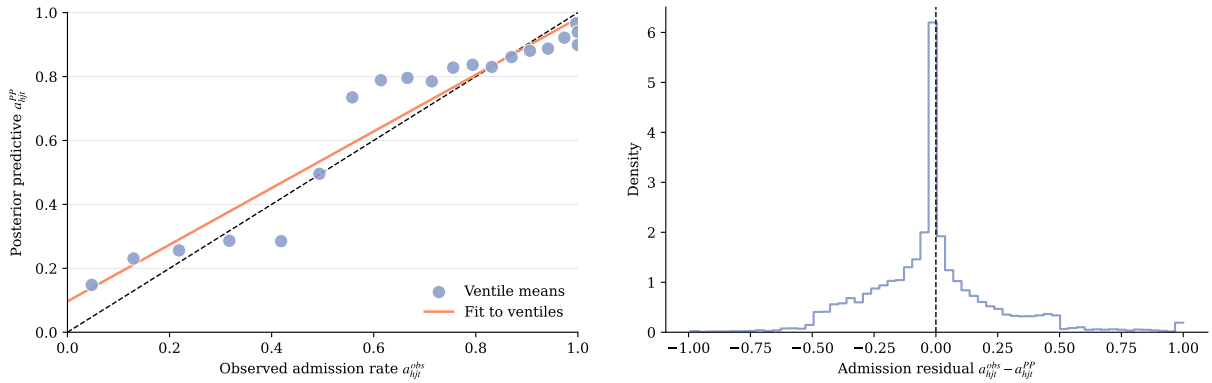
(b) Congestion Cost

(c) Overflow Cost

(d) Perceived Value of Uninsured Care

Figure A.8: Distribution of Estimated Cost Components

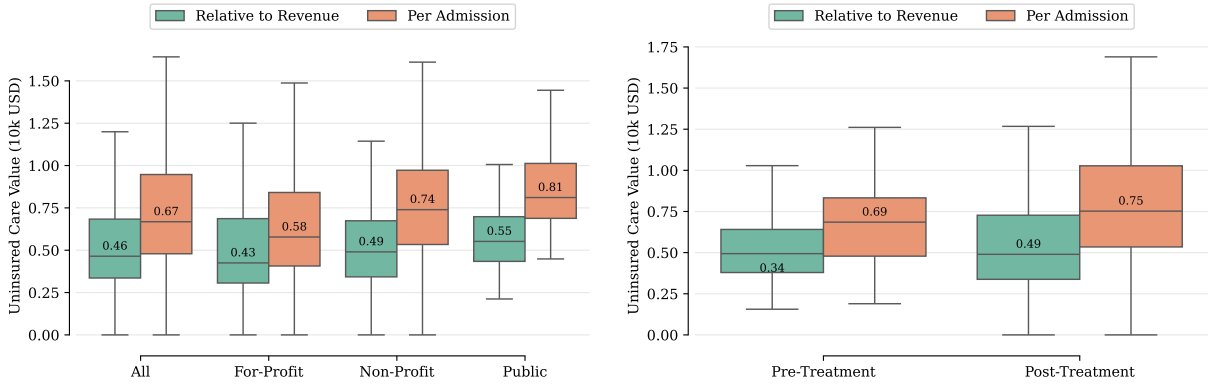
Notes: These figures show the distribution of hospitals' cost estimates from the structural estimates. In each figure, the histogram shows the overall distribution of estimates, while the horizontal boxplots show the yearly distribution. Figure (a) shows the estimates associated with components c^S_{hjt}, c^F_{hjt}. Figure (b) shows the cost associated with c^V_{hjt}. Figure (c) shows the costs associated with component c^o_{ht}. Figure (d) shows the perceived revenue from uninsured care, associated with component τ_{hjt}.



(a) Fit **(b)** Residual

Figure A.9: Cost Estimates Fit — Implied and Observed Admission Rates

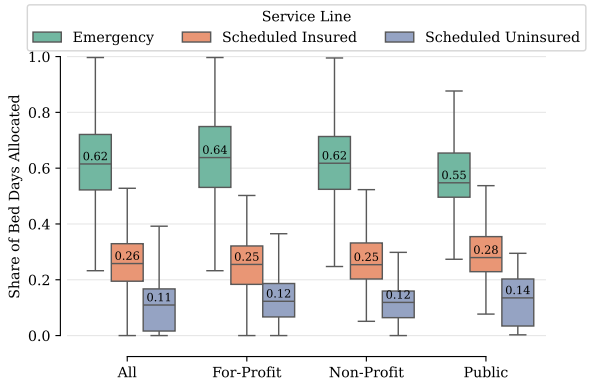
Notes: These figures show the fit of our structural model. Using the estimated cost and τ_{hjt} components, we compute each hospital's optimal admission rate and contrast it with the observed rate.



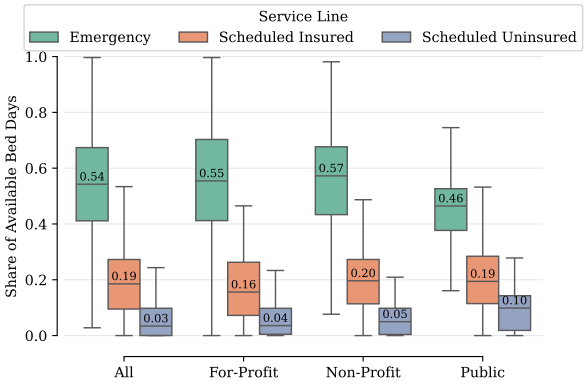
(a) By Hospital Type **(b)** By Spillover Treatment

Figure A.10: Estimate of Hospitals' Preference for Uninsured Care

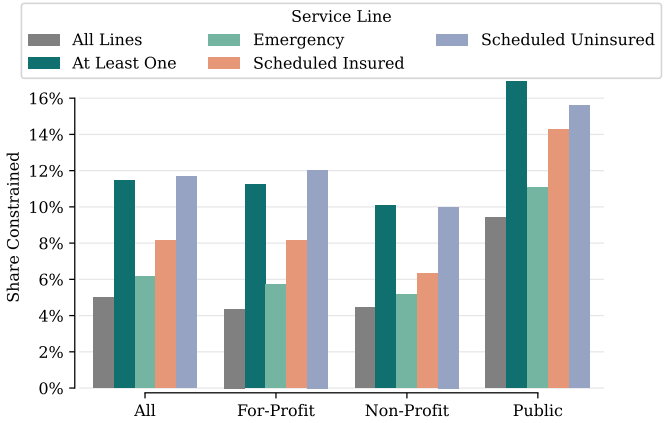
Notes: These figures show estimates associated with τ_{hjt}, the hospital's preference for improving uninsured surplus.



(a) Allocated Bed-Days (b_{htj})



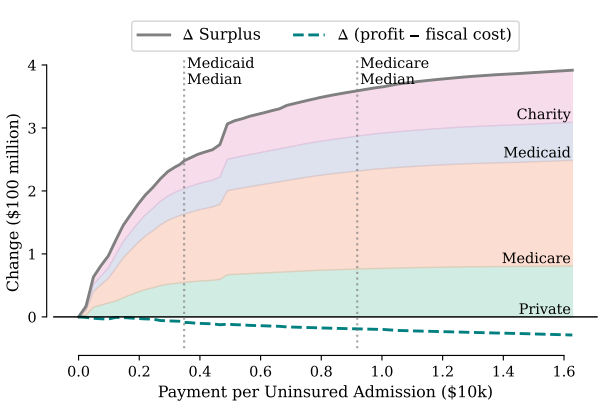
(b) Allocated Available Bed-Day ($a_{hjt}b_{hjt}$)



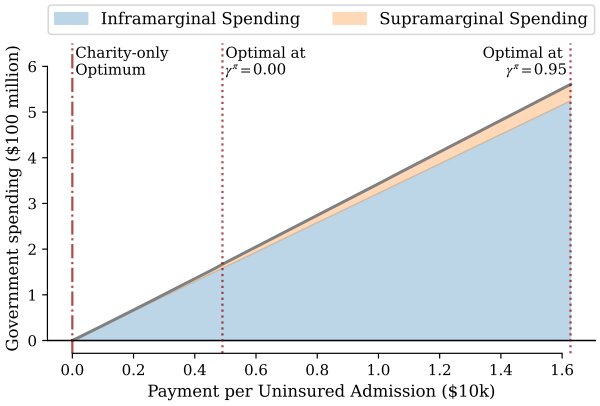
(c) Lines Constrained ($a_{hjt}\Lambda_{hj}\mu_{hj}^{L} = b_{hj}$)

Figure A.11: Estimated Resource Allocation

Notes: These figures show the estimated resource allocation of hospitals among service lines.



(a) Welfare Change



(b) Spending Change

Figure A.12: Pay-per-Admission Policies

Notes: These figures display the equilibrium welfare effects associated with compensating hospitals per uninsured admission.

Table A.8: Counterfactual per-capita surplus and allocative-efficiency summary.

	Panel A: per-capita surplus change ($10k/patient)				Panel B: share of GAP allocative gains (%)				
	Uninsured	Private	Medicare	Medicaid	Uninsured	Private	Medicare	Medicaid	Overall
Panel A: Pay-per-Complexity									
Medicaid-equivalent	0.1235	0.0486	0.0467	0.0617	10.559	2.305	5.235	3.729	4.284
Medicare-equivalent	0.1789	0.0654	0.0644	0.0820	17.663	3.355	7.743	5.315	6.444
Optimum at $\gamma^{\pi} = 0$	0.1538	0.0583	0.0561	0.0748	14.043	2.846	6.209	4.777	5.301
Optimum at $\gamma^{\pi} = 0.93$	0.2059	0.0710	0.0706	0.0906	22.512	4.153	9.054	6.545	7.823
Panel B: Floor-and-Trade									
Unconstrained	0.1236	0.0355	0.0459	0.0421	15.664	2.485	8.616	3.564	5.993
Homogenous	0.0320	0.0011	0.0050	0.0030	4.219	-0.040	0.990	0.126	0.707

Notes: Panel A reports the change in per-capita consumer surplus in $10,000 per patient, that is $\gamma^{u}(\sum_k S_{rk} - \sum_k S_{bk})/\sum_k \lambda_k$ with γ^{u} already denominated in $10,000 per consumer-mile. Panel A values patients on the modeled logit surplus and Panel B on the deterministic value of the assigned option; both exclude testing-data sentinel choices from their numerator and their denominator. Panel B is the percentage of the 2019Q4 z=0-to-GAP allocation gain captured by each policy, by payor, followed by the pooled share over all payors. A blank Panel B cell marks a payor whose z=0-to-GAP gain is not positive, for which the share of that gain is undefined.

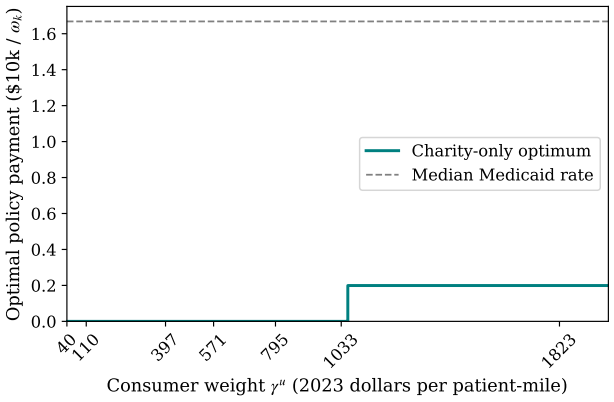
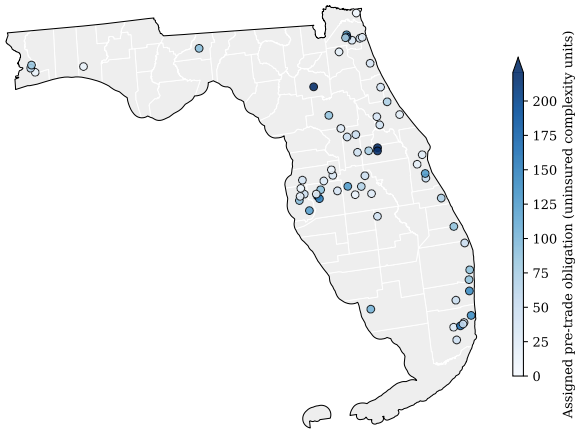
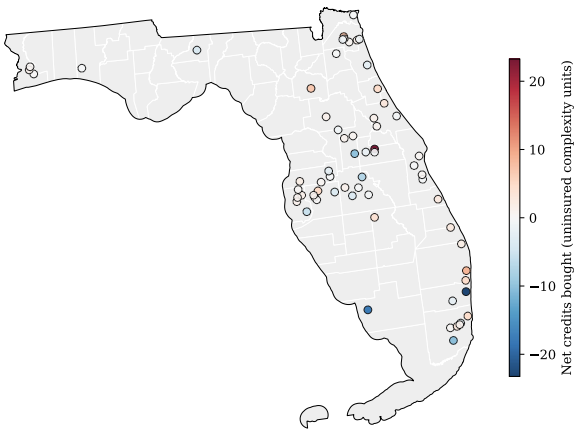


Figure A.13: Optimal Pay-Per-Complexity Charity-Surplus-Only Policy

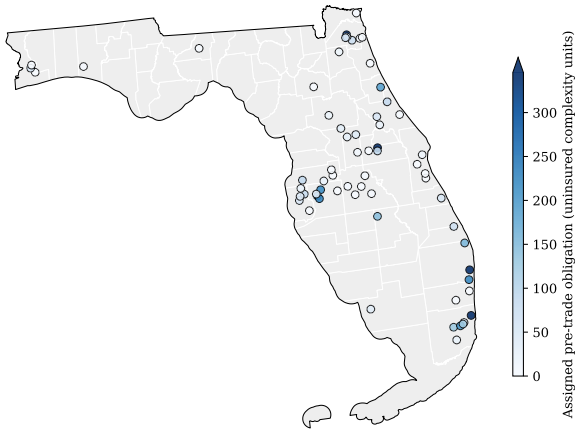
Notes: This figure shows the optimal policy when $\gamma^{\pi} = 0$ and the regulator only values uninsured patient surplus.



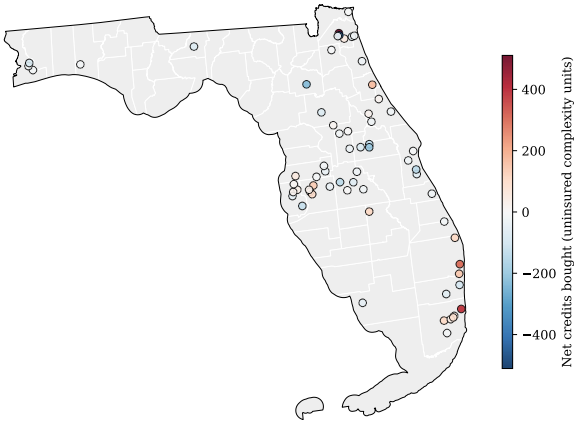
(a) Unconstrained Credit Obligations



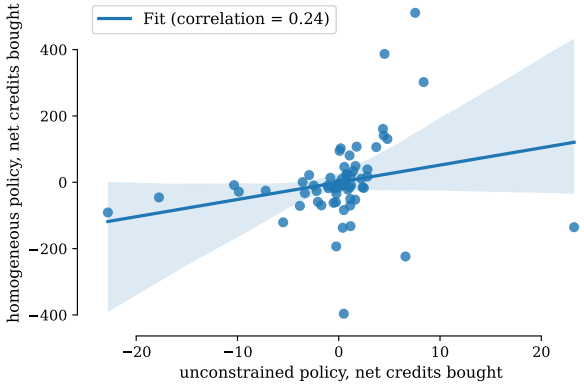
(b) Unconstrained Credit Trade



(c) Homogeneous Credit Obligations



(d) Homogeneous Credit Trade



(e) Homogeneous and Unconstrained Correlation

Figure A.14: Floor-and-Trade Credit Trade

Notes: These figures are for 2019Q4. Figure A.14a shows the assigned obligations under the optimal unconstrained floor-and-trade policy, while Figure A.14b shows the trading position in equilibrium of participating hospitals. Figures A.14c and A.14d show the same for the homogeneous policies. Finally, Figure A.14e shows the correlation between trading positions in the two policies.

Table A.9: Tax-at-Risk Sensitivity: Floor-and-Trade Welfare Effects

Tax at risk	Trading Price	Δ **Welfare** ($100M)				Δ **Available Capacity** (1k bed-days)			
		Patients		Hospitals	Gov.	Emergency	Scheduled		Idle
	$z^\star$	All	Charity				Insured	Uninsured	(p.p.)
Unconstrained									
5%	11.317	2.501	0.543	-0.169	0.000	118.675	-18.918	110.240	-6.167
10%	11.686	2.516	0.545	-0.172	0.000	119.247	-19.013	110.235	-6.181
15%*	11.686	2.516	0.545	-0.172	0.000	119.247	-19.013	110.235	-6.181
20%*	11.686	2.516	0.545	-0.172	0.000	119.247	-19.013	110.235	-6.181
Homogeneous									
5%	0.126	0.203	0.117	-0.002	0.000	6.453	-2.483	22.959	-0.791
10%	0.251	0.310	0.141	-0.002	0.000	13.831	-5.085	25.268	-0.999
15%	0.370	0.432	0.178	-0.003	0.000	21.885	-7.769	29.220	-1.273
20%	0.474	0.613	0.258	-0.018	0.000	30.059	-8.931	76.451	-2.866

Notes: This table shows the effect of increasing the tax-at-risk for non-profit hospitals in the floor-and-trade policies. Welfare and capacity changes are relative to the status quo. All values are quarterly. We note that the last two rows in the unconstrained solve are mechanically identical to the 10% because the slope of the planner's objective in trading-price space is flat to numerical precision at the 10% solution. However, there are minimal gains from increasing the trading price above this level, but the effect is on the fifth decimal of the welfare objectives.

Table A.10: Cumulative True Medicaid Welfare Effects

	Regulatory Step	Δ **Welfare** ($100M)				Δ **Available Capacity** (1k bed-days)			
		Patients		Hospitals	Gov.	Emergency	Scheduled		Idle
		All	Charity				Insured	Uninsured	(p.p.)
Cumulative Medicaid Expansion									
Medicaid Rates	1	2.943	0.591	1.313	1.949	153.521	-5.286	144.756	-8.605
↪ + Medicaid Risk Mix	2	5.428	3.224	1.433	1.974	146.767	-4.642	122.872	-7.783
↪ + Medicaid Networks	3	5.402	3.198	1.429	1.977	148.656	-2.120	118.083	-7.772
↪ + Change of Service Line	4	5.354	3.206	1.227	1.990	313.908	184.760	-211.698	-8.428

Notes: Each row is a cumulative endpoint of the Medicaid transition; all changes are relative to the original status quo baseline, not to the preceding row. Welfare is measured in hundreds of millions. Patient surplus and fiscal cost cover solver choices and fully fixed choices. The Patients columns report the gains for all patients and charitable-care patients, respectively, with "Gov." representing the change in fiscal cost. Available capacity is the sum across hospitals of the product of each line's allocated bed-days and its acceptance rate, measured in thousands. Idle capacity ratio is defined as $(\sum_h B_h - \sum_{h,j} a_{hj} b_{hj}) / \sum_h B_h$ and the change is presented in percentage points. The first column reports the transition step. All values are quarterly.

A.2 Spillover Effects on Quality and Distance Traveled

In this appendix section, we explore how spillovers from Medicaid expansion to non-expansion states affected the quality and distance experienced by patients. Intuitively, if the spillover effects documented in the main text are due to a supply expansion of emergency care at spillover-treated hospitals, then demand should reallocate in favor of its preferences. As patient preferences are not directly identified without further assumptions, a natural proxy is to ask whether their admissions occurred closer to home or at better hospitals.

To measure changes in quality, we examine how the average quality of care changes for patients in treated regions. In particular, we use mortality and readmission metrics estimated and published by CMS and examine how the county-level weighted average quality changes with treatment. We weight each quality metric by the number of admissions of the population of interest, to capture the average quality experienced by patients of different insurance coverage. Formally, we compute $\bar{q}_{ict} = \sum_{h \in \mathcal{H}_c} q_{ht} s_{ihct}$, where q_{ht} is the quality outcome of hospital h in period t, $\mathcal{H}_c$ is the set of hospitals found in county c, and s_{ihct} is the share of admissions of patients with insurance source i from county c that are admitted to hospital h at time t. We measure quality at this aggregate level rather than at the hospital level, because quality metrics vary little over time and our interest is in capturing how the reallocation of patients benefited or disadvantaged certain population groups. To account for local changes in overall quality, we explore the changes in quality perceived by those seeking charitable care and those on government insurance, relative to the privately insured. Thus, we define $\Delta\bar{q}_{ict} = \bar{q}_{ict} - \bar{q}_{private,ct}$ and conduct the regressions

$$\Delta\bar{q}_{ict} = \alpha T_{ct} + \gamma_c + \delta_{s(c)t} + \epsilon_{ict}$$

Where T_{ct} indicates that there is at least one hospital in the county that experiences spillover treatment in a period prior to or equal to t, γ_c is a county fixed effect, and $\delta_{s(c)t}$ is a state-by-year fixed effect. Table A.11 shows the estimates.

The estimates indicate that, relative to privately insured individuals, charitable care patients are admitted to hospitals with lower mortality rates across five of six dimensions. While the magnitude of the effects is modest, the direction supports the idea that charitable care patients move towards better hospitals following Medicaid expansion spillovers. This is consistent with those hospitals obtaining a larger share of the charitable care population by expanding bed-day reservations for patients admitted via the emergency room and with those patients being more likely to seek charity care.

We now turn our attention to the effect of spillover on distance traveled. The first way to assess whether the expansion of charitable care benefited charity patients by offering them care closer to home is to use the average distance among admitted charitable care patients as a dependent variable in the regression of Equation (2). Table A.12 shows that, consistent with the hypothesis

Table A.11: Medicaid Spillover Effect on Admitted Quality by Charitable Care Patients

	Panel A: Mortality Measures					
	AMI	CABG	COPD	Heart Failure	Pneumonia	Stroke
	(1)	(2)	(3)	(4)	(5)	(6)
Spillover	-0.196**	0.017	-0.085	-0.089*	-0.102**	-0.042
Effect (α)	(0.096)	(0.080)	(0.064)	(0.052)	(0.045)	(0.057)
N	853	853	853	853	853	853
R^2	0.641	0.465	0.668	0.269	0.642	0.872
Mean Dep.	11.745	1.499	7.930	11.485	11.762	14.066
N	875	875	875	875	875	875
R^2	0.512	0.486	0.775	0.474	0.500	0.742
Mean Dep.	11.516	1.467	7.853	11.364	11.677	13.952
	Panel B: Readmission Measures					
	AMI	CABG	COPD	Heart Failure	Pneumonia	Stroke
	(1)	(2)	(3)	(4)	(5)	(6)
Spillover	-0.062	0.030	0.163*	-0.053	-0.069	0.111
Effect (α)	(0.159)	(0.279)	(0.096)	(0.201)	(0.063)	(0.114)
N	853	853	853	853	853	853
R^2	0.631	0.494	0.678	0.262	0.462	0.658
Mean Dep.	12.857	6.810	19.566	21.409	16.752	11.792
N	875	875	875	875	875	875
R^2	0.641	0.525	0.809	0.481	0.585	0.487
Mean Dep. Raw	12.613	6.669	19.370	21.189	16.595	11.691

Notes: This table presents the estimated effect of local exposure to spillover effects from Medicaid expansions in non-expansion states on the quality of hospitals accessed by patients seeking charitable care. Quality outcomes are measured by CMS on a rate basis, per 100 patients. The outcome is the county-level average quality for the population relative to the average quality for privately insured patients. All regressions include county fixed effects and state-year fixed effects. Standard errors are clustered at the HRR level. "Mean Dep" shows the average quality outcome experienced by charitable care patients, $\bar{q}_{ict}$. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

above, charitable care patients are the only patient type experiencing a decrease in average distance at spillover hospitals.

Overall, we observe that both distance and quality improved for charitable care patients. This sorting of demand thus appears consistent with elements that might plausibly correlate with their hospital preferences. We note that the interaction between quality and distance might attenuate the effect documented here if higher-quality hospitals are located farther from where the uninsured live. This could help explain the imprecision in the estimated distance effect, as it mixes patients sorting toward quality and away from distance, and vice versa.

Table A.12: Medicaid Spillover Effect on Admitted Distance at Spillover-Treated Hospitals

	Source of Insurance			
	Private	Gov.	Charitable	All
Expansion Effect (α)	0.117 (0.226)	0.332 (0.231)	-0.544* (0.319)	0.116 (0.173)
N	3065	3065	3031	3065
R^2	0.992	0.997	0.971	0.998
Mean Dep.	13.051	11.709	11.659	12.088

Notes: The This table presents the estimated effect of Medicaid expansion in non-expansion states, particularly on spillover-treated hospitals. The estimate shows the effect on the distance patients travel between their ZIP code of residence and the hospital admitting them, as estimated from Equation (2). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

A.3 Model Appendix

This appendix presents developments associated with the model described in section 5.

A.3.1 Admission Rate Approximation. Our main model assumes that hospitals dictate the rate at which patients are admitted rather than control a queuing process. In this appendix, we present two models of hospital behavior in which the hospital line controls a queue rather than the admission rate directly. The first model is simpler and yields exactly the same payoffs as our main model, whereas the second model is more realistic and yields payoffs that can be made arbitrarily close to those of our main model. We present the model below for a single hospital, service line, period, and market, and omit the associated subindices. We begin by restating our main model in a simplified way that captures the core economics of the problem and minimizes the notational burden of our empirically oriented model.

Main Model: In this model, all patients arrive at the hospital at the beginning of the period, with the number of admission requests being distributed according to a Poisson distribution with rate λ. The line has capacity b, and each admitted patient yields a profit of v and occupies a bed for the full period. The hospital decides in advance on a fraction of patients to admit, a, and pays a rental cost r for any additional capacity it requires to serve its admitted population. As in our main model, the implied admission count is Poisson with rate $a\lambda$. Without loss of generality, we can consider the hospital as choosing the admission rate $x = a\lambda$, subject to the constraint $x \in [0, \lambda]$. The expected profit of the hospital is

$$\pi^{main}(x) = vx - r\mathbb{E}[(N(x) - b)_+]$$

where $N(x)$ is the number of admitted patients, and $(y)_+$ denotes the positive part of y.

Posted-Wait Time Model: In this simple queuing model, the hospital manages a queue by posting a wait time $w \geq 0$, indicating how long patients must wait for service. Patients arrive continuously over time according to a Poisson process at rate λ and, upon requesting admission, draw a patience time T_i, which indicates the maximum amount of time they are willing to wait for service. Therefore, a patient waits in the queue only if $T_i \geq w$. The hospital has b beds, and each admitted patient stays for a unit duration. Any patient who must be admitted when no beds are available (due to wait-time commitments) must be served in a rented bed. As before, the hospital earns a profit of v per admission and rents beds at a rate of r. Denote $a(w) = \mathbb{P}(T_i \geq w)$ and assume that T_i is i.i.d. in the population with positive density on the positive reals. Then the number of patients in the committed queue is a Poisson random variable with rate $x(w) \equiv a(w)\lambda$. In stationarity, the number of patients requiring beds at any instant is the number of committed services that start in the previous unit of time, which is a Poisson random variable of rate $x(w)$. Therefore, the hospital experiences the same occupancy demand per period as in our main model, and since the distribution of patients is supported on the positive real line, the hospital can indirectly control the rate $a(w) \in [0, 1]$ by adjusting the posted wait time. Hence, this model yields exactly the same payoff as our main model.

Despite its simplifications, this model is closer to real hospital behavior. In practice, patients arrive continuously over time and are given wait times rather than being admitted at random. While wait times are typically not uniform, the homogeneity of w can be treated as a plug-in approximation of the mean wait time, used by patients to determine whether to commit to the queue and by hospitals to simplify their analysis. The main caveat of this model, however, is that at the beginning of the period, there are beds that sit idle, since every patient is made to wait even if capacity is available. The next model corrects this issue.

FCFS with Wait Time Commitment Model: This model has the same structure as the *Posted-Wait Time Model* except that hospitals commit to having patients wait at most $w \geq 0$ units of time before service, but treat patients as soon as capacity is available on a first-come, first-served (FCFS) basis. As before, patients join the queue only if their patience exceeds the posted wait-time guarantee (i.e., they are pessimistic about wait times) and occupy a bed for a full unit of time. We note that under this model, profits net of rental costs are the same as in the posted-wait time model in stationarity: patients commit to the queue at a rate $x(w)$, and thus, in expectation, the hospital is paid $vx(w)$. The main difference is in the rental cost, since the hospital makes better use of idle capacity with an FCFS queue.

The difference in rental cost between the main model, which serves patients immediately, and the FCFS model, which queues patients, is determined by how likely a patient is to occupy a resource at any given time. To quantify this, fix a time t and consider the indicator that marks whether she is occupying a bed under the FCFS model, versus the indicator that marks her occupancy under the main model. First, if the patient arrived between $t - w$ and t, the main model

might mark her as using a resource, while the FCFS model might have her waiting for service. Alternatively, the patient might have arrived between $t-1-w$ and $t-1$, which implies that the immediate service model already served and discharged the patient, while the FCFS model is still serving her. Each of these two windows has length w, and patients arrive at rate $x(w)$ per period, so the expected number of patients subject to these discrepancies is $2x(w)w$. Now, since for any y, z, b we have that $|(y-b)_+ - (z-b)_+| \leq |y-z|$, the discrepancy between the models in overflow in a period of length one is bounded above by $2x(w)w$. Therefore, denoting $\pi^{FCFS}(w)$ as the expected profit of the firm under the FCFS model, we have that

$$|\pi^{FCFS}(w) - \pi^{main}(w)| \leq 2rx(w)w$$

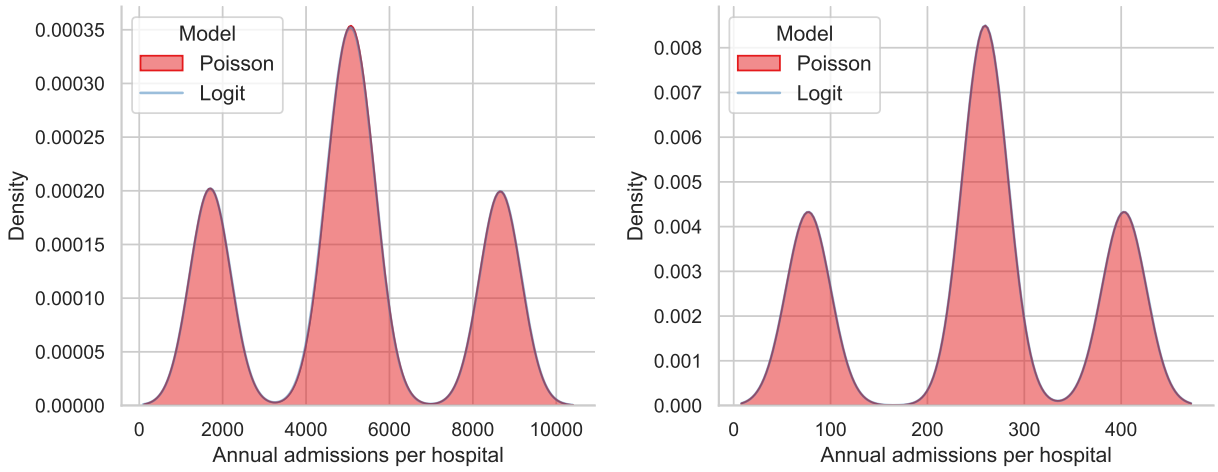
Now suppose that $\mathbb{P}(T_i \geq w) = \frac{1}{1+\eta w}$ (i.e., patience is distributed Pareto type II). Which implies that $x(w) = \frac{\lambda}{1+\eta w}$, which can be reorganized to $wx(w) = \frac{\lambda - x(w)}{\eta}$, which implies that $wx(w) \leq \lambda/\eta$. Then we have:

$$|\pi^{FCFS}(w) - \pi^{main}(w)| \leq \frac{2r\lambda}{\eta}$$

Therefore, the profit discrepancies between the two models decay to zero as consumer patience becomes small ($\eta \to \infty$).

Discussion: These two models show clearly how our admission-rate control model approximates a queue-control model. In particular, admission-rate control yields payoffs similar to those from controlling the wait time patients face by managing queues and rented capacity. However, we note that our admission-rate control model implies a significant simplification relative to a queue management model once competition is incorporated. In particular, when considering estimation, a queue model requires us to distinguish between patients that were admitted directly under idle capacity versus those that are made to wait. Lacking data on wait times, this cannot be resolved without stronger assumptions. Moreover, when considering competition, it becomes natural to consider that patients' patience is not exogenous but rather dictated by the wait times offered at rival hospitals. This links demand across firms in a complex way, making analysis intractable without further simplifications or extensive simulation. Hence, our approach offers a midpoint between realism and tractability. As we show in our estimates, the model yields accurate cost estimates for hospitals, suggesting that our approximations have a modest impact on the characterization of optimal resource allocation.

A.3.2 Poisson Rate Approximation. In this section, we discuss briefly why we consider the Poisson illness process to be mild in our setting. In what follows, we take the independence assumption as given and discuss only the difference between our model and the standard model, which treats each illness draw as a Bernoulli random variable.



(a) No capacity constraints **(b)** Capacity constrained

Figure A.15: Admission rate comparison between multinomial logit model and Poisson

Notes: These figures present a comparison of patient arrival rates at hospitals in an environment calibrated to approximate the city of Madison, WI. The distribution is based on 800 Monte Carlo replicates. The simulated city has a population of 283,000 and four large hospitals, with capacities equal to 672, 440, 426, and 129. We assume a 60% capacity factor and a 9% probability of needing inpatient care. In each simulation, we randomly locate consumers and hospitals across a 9×9 grid, representing square miles in the city. The logit and Poisson are nearly overlapping, making the logit distribution appear as the contour of the Poisson distribution.

First, we note that the population in the empirical application of hospital demand is often very large. For example, in our empirical analysis, we consider the state of Florida, which, in 2019, had a population of over 21.2 million. Second, we note that relative to the population, hospitalizations are rare, with about 2.6 million hospitalization events at general hospitals. By the Poisson limit theorem (Feller, 1968), if n is the size of the population and p_n is the probability of illness, then as $p_n \to 0$ and $np_n \to \lambda$, the distribution of the number of illnesses at any period converges to a Poisson distribution of rate λ.

To illustrate the precision of this assumption, we simulate admission decisions in a market calibrated to the characteristics of Madison, WI. We simulate a large sample of potential patients, each choosing hospitals based on distance and some hospital-specific quality. Following our demand model, we assume that patients are heterogeneous in distaste for distance or out-of-market consumption, with individual characteristics drawn independently from log-normal distributions. Patients choose between the four general inpatient hospitals located in the city or an out-of-market option. Figure A.15 shows the distribution of arrivals implied by the true model with and without a capacity constraint. As can be seen, the distributions are almost exactly identical, as expected by Poisson's law.

A.3.3 Derivations. Here, we present the derivation of Equation (10). Given that the derivation is for a single market period, we omit all t and m indices. We begin by presenting the derivation

of the uninsured surplus, and then show the derivations of the expectations of the other terms.

Denote by $\mathcal{K}^u_{hj}$ the subset of consumer types that are uninsured and require charity care. The expected surplus of the uninsured population is:

$$\begin{aligned}
\mathbb{E}[S^u_{hj}(a_{hj})] &= \sum_{k\in\mathcal{K}^u_{hj}} \lambda_k \left[\mathbb{E}_{i,\boldsymbol{a}}[\ln(1+\sum_{h'\in\mathcal{H}} \tilde{a}_{ih'} e^{v_{ih'}})|(m_i,d_i)=(m_k,d_k)] + \gamma^{euler} \right] \\
&= \sum_{k\in\mathcal{K}^u_{hj}} \lambda_k \Bigg[\mathbb{E}_{i,\boldsymbol{a}_{-h}}[a_{hj}\ln(1+e^{v_{ih}}+\sum_{h'\in\mathcal{H}\setminus\{h\}} \tilde{a}_{ih'} e^{v_{ih'}}) \\
&\qquad +(1-a_{hj})\ln(1+\sum_{h'\in\mathcal{H}\setminus\{h\}} \tilde{a}_{ih'} e^{v_{ih'}})|(m_i,d_i)=(m_k,d_k)] + \gamma^{euler} \Bigg] \\
&= \sum_{k\in\mathcal{K}^u_{hj}} \lambda_k \left[\mathbb{E}_{i,\boldsymbol{a}_{-h}}[a_{hj}\ln(\frac{1+e^{v_{ih}}+\sum_{h'\in\mathcal{H}\setminus\{h\}}\tilde{a}_{h'i}e^{v_{ih'}}}{1+\sum_{h'\in\mathcal{H}\setminus\{h\}}\tilde{a}_{h'i}e^{v_{ih'}}}) + \ln(1+\sum_{h'\in\mathcal{H}\setminus\{h\}}\tilde{a}_{h'i}e^{v_{ih'}})|(m_i,d_i)=(m_k,d_k)] + \gamma^{euler} \right] \\
&=: a_{hj}\Delta_{hj}S^u_{hj} + S^u_{j,-h}
\end{aligned}$$

Since $S^u_{j,-h}$ does not depend on hospital h's actions, it is omitted from the objective in Equation (10).

We now proceed with the remaining terms. To recall, we seek to simplify

$$\mathbb{E}[\sum_{k\in\mathcal{K}_j} n_{hk}(p_{hk}-c_{hk}) - c^V_{hj}N^2_{hj} - \exp(c^o_h(N^L_{hj}(a)-b))]$$

First, by the Poisson distribution of n_{hk}, we have that $\mathbb{E}[n_{hk}] = a_{hj}\lambda_k R_{hk}$. Second, recall that N_{hj} is distributed Poisson with rate $a_{hj}\Lambda_{hj}$. Therefore, by the standard expression for the second uncentered moment of a Poisson random variable, we have that $\mathbb{E}[N^2_{hj}] = a\Lambda_{hj} + a^2\Lambda^2_{hj}$. Finally, the term $\mathbb{E}[\exp(c^o_h(N^L_{hj}(a)-b))]$ can be derived from noting that it is closely related to the moment generating function of the random variable $N^L_{hj}(a)$, which is distributed as a compound Poisson. In particular, we have that $N^L_{hj} = \sum_{i=1}^{N_{hj}} \sum_{k\in\mathcal{K}_j} \frac{\lambda_k R_{hk}}{\Lambda_{hj}} \ell^k_i$ and thus that the moment-generating function of a randomly chosen admitted patient is $M^L_{hj}(t) = \sum_{k\in\mathcal{K}_j} \frac{\lambda_k R_{hk}}{\Lambda_{hj}} \mathbb{E}[e^{t\ell^k_i}]$. Therefore, $\mathbb{E}[e^{tN^L_{hj}}|N_{hj}=n] = (M^L_{hj}(t))^n$, and thus

$$\mathbb{E}[e^{tN^L_{hj}}] = \sum_{n=0}^{\infty} e^{-a_{hj}\Lambda_{hj}} \frac{(a_{hj}\Lambda_{hj})^n}{n!} (M^L_{hj}(t))^n = \exp(a_{hj}\Lambda_{hj}(M^L_{hj}(t)-1))$$

where the second equality follows from the series formula for the exponential. Finally, applying it to the term $\mathbb{E}[\exp(c^o_h(N^L_{hj}(a)-b))]$ yields $\exp(a\Lambda_{hj}(M^L_{hj}(c^o_h)-1) - c^o_h b)$ which is the expression shown in the main text.

A.3.4 Model Properties. Here we present the proof of Proposition 1 from the main text and discuss sufficient conditions for equilibrium uniqueness. For the first three items, the proof is done for a single period, market, and hospital; hence, the t, m, h indices are suppressed. We restore the hospital index only when proving equilibrium existence and discussing uniqueness. Moreover, many of the properties pertain to a single service line; hence, we suppress the j index unless it is relevant for clarification. We note that the objective function of the line can be expressed as

$$\pi(b) = \max_{a \in [0,1]} a\pi^L - a^2\pi^V - \exp(a\chi - c^o b) \quad \text{s.t.} \quad a\Lambda\mu^L \leq b$$

where $\pi^V, c^o > 0$ and $\chi > 0$ by the assumption stated in the main text. We use this simplified notation in the proof below.

Proof of Proposition 1. We start by deriving the closed-form expressions for $a^o(b)$ and $a^*(b)$, which are useful for the empirical analysis but unused in the rest of the proof. Note that the first-order condition of the unconstrained problem above is

$$\pi^L - 2a\pi^V - \chi \exp(a\chi - c^o b) = 0.$$

Define $y = \pi^L - 2\pi^V a$; then the first-order condition simplifies to

$$\begin{aligned}
y &= \chi \exp\left(\chi \frac{\pi^L - y}{2\pi^V} - c^o b\right) \\
y \exp\left(\frac{\chi y}{2\pi^V}\right) &= \chi \exp\left(\frac{\chi \pi^L}{2\pi^V} - c^o b\right) \\
\frac{\chi y}{2\pi^V} \exp\left(\frac{\chi y}{2\pi^V}\right) &= \frac{\chi^2}{2\pi^V} \exp\left(\frac{\chi \pi^L}{2\pi^V} - c^o b\right) \\
\frac{\chi y}{2\pi^V} &= W_0\left(\frac{\chi^2}{2\pi^V} \exp\left(\frac{\chi \pi^L}{2\pi^V} - c^o b\right)\right) \\
\Rightarrow a^o(b) &= \frac{\pi^L}{2\pi^V} - \frac{1}{\chi} W_0\left(\frac{\chi^2}{2\pi^V} \exp\left(\frac{\chi \pi^L}{2\pi^V} - c^o b\right)\right).
\end{aligned}$$

where W_0 is the main branch of the Lambert W function. From this, the expression for $a^*(b)$ is simply

$$a^*(b) = \min\left\{1, \frac{b}{\Lambda\mu^L}, \max\{0, a^o(b)\}\right\}.$$

This concludes the preliminary part of the proposition.

Denote the objective in the maximization by $\tilde{\pi}(a, b) = a\pi^L - a^2\pi^V - \exp(a\chi - c^o b)$. The Hessian

of the objective is

$$\nabla^2\tilde{\pi} = \begin{pmatrix} -2\pi^V - \chi^2\exp(a\chi - c^o b) & \chi c^o \exp(a\chi - c^o b) \\ \chi c^o \exp(a\chi - c^o b) & -(c^o)^2\exp(a\chi - c^o b) \end{pmatrix}.$$

Therefore, for any non-zero vector $\boldsymbol{x} = (x_1, x_2)$, we have that

$$\boldsymbol{x}'(\nabla^2\tilde{\pi})\boldsymbol{x} = -2\pi^V x_1^2 - \exp(a\chi - c^o b)(\chi x_1 - c^o x_2)^2 < 0.$$

Hence, the objective is strictly concave. Now consider the hospital's problem

$$\begin{aligned} \max_{\boldsymbol{a},\boldsymbol{b}} \quad & \sum_{j\in\mathcal{J}_h} \tilde{\pi}_j(a_j, b_j) \\ \text{s.t.} \quad & 0 \le a_j \le 1, \\ & a_j \Lambda_j \mu_j^L \le b_j, \\ & \sum_{j\in\mathcal{J}_h} b_j \le B. \end{aligned}$$

Since the sum of strictly concave functions is strictly concave and the constraint set is convex because it is defined by affine inequalities, the firm's problem is strictly concave. Since the objective is continuous and the constraint set is compact, a solution exists; strict concavity implies that it is unique. This proves the first item of the proposition.

The second item of the proposition follows from the cross-partial in the Hessian above, $\chi c^o \exp(a\chi - c^o b) > 0$, which implies that $\tilde{\pi}$ is supermodular in (a, b).[42] Hence, absent constraints, $a^o(b)$ is strictly increasing in b. Introducing the ex-ante constraint and assuming $a^*(b) \in (0, 1)$ as stated in the proposition, the only relevant corner is $a^*(b) = \frac{b}{\Lambda\mu^L}$, which is also strictly increasing in b. This proves the second item of the proposition.

We now prove the third item of the proposition with greater generality. First, we prove that $\pi(b)$ is concave. To see this, take $b_1, b_2 \ge 0$, and denote $a_1 = a^*(b_1)$ and $a_2 = a^*(b_2)$. Let $\lambda \in (0, 1)$, and denote $a_\lambda = \lambda a_1 + (1-\lambda)a_2$ and $b_\lambda = \lambda b_1 + (1-\lambda)b_2$. Since a_1 is feasible under b_1 and a_2 is feasible under b_2, we have that $a_\lambda \Lambda\mu^L \le b_\lambda$. Then

$$\begin{aligned} \pi(b_\lambda) &\ge \tilde{\pi}(a_\lambda, b_\lambda) \\ &= \tilde{\pi}(\lambda a_1 + (1-\lambda)a_2, \lambda b_1 + (1-\lambda)b_2) \\ &\ge \lambda\tilde{\pi}(a_1, b_1) + (1-\lambda)\tilde{\pi}(a_2, b_2), \end{aligned}$$

where the first inequality follows from the optimality of $a^*(b_\lambda)$ and a_λ being feasible, and the second inequality from the concavity of $\tilde{\pi}$.

[42] Alternatively, implicit differentiation of the interior solution for admission rates yields a strictly positive term.

Given that $\pi(b)$ is concave, it is almost everywhere differentiable and has a well-defined superdifferential, which we denote by $\partial\pi(b)$. Denote η as the multiplier on the hospital's total resource constraint. The KKT stationarity condition of the firm's problem is thus

$$\eta \in \partial\pi_j(b_j^*) \quad \forall j \in \mathcal{J}_h \text{ such that } b_j^* > 0.$$

This implies that for any j, j' with $b_{j'}^*, b_{j'}^* > 0$, we have that $\eta \in \partial\pi_j(b_j^*) \cap \partial\pi_{j'}(b_{j'}^*)$. An immediate implication is that, at any point of differentiability,

$$\left.\frac{\partial\pi_j(b)}{\partial b}\right|_{b=b_j^*} = \left.\frac{\partial\pi_{j'}(b)}{\partial b}\right|_{b=b_{j'}^*}.$$

This proves the third item of the proposition.

It remains to prove the fourth item. We now restore the hospital index and make the dependence on rivals' admission rates explicit. All demand and payoff objects in Equations (6) and (10) are evaluated at $\boldsymbol{a}_{-h}$. Denote the expected bed-day load of line j when its admission rate equals one as

$$L_{hj}(\boldsymbol{a}_{-h}) := \Lambda_{hj}(\boldsymbol{a}_{-h})\mu_{hj}^L(\boldsymbol{a}_{-h}).$$

Hospital h's feasible set, conditional on $\boldsymbol{a}_{-h}$, is

$$\begin{aligned}\Gamma_h(\boldsymbol{a}_{-h}) := \Big\{(\boldsymbol{a}_h, \boldsymbol{b}_h) : \boldsymbol{a}_h \in [0,1]^{J_h}, \quad \boldsymbol{b}_h \geq \mathbf{0}, \quad \sum_{j\in\mathcal{J}_h} b_{hj} \leq B_h, \\ a_{hj}L_{hj}(\boldsymbol{a}_{-h}) \leq b_{hj} \quad \text{for every } j \in \mathcal{J}_h\Big\}.\end{aligned} \tag{22}$$

Let

$$\mathrm{BR}_h(\boldsymbol{a}_{-h}) = \left(\mathrm{BR}_h^a(\boldsymbol{a}_{-h}), \mathrm{BR}_h^b(\boldsymbol{a}_{-h})\right)$$

denote the solution to hospital h's joint admission and resource-allocation problem obtained by combining Equations (7) and (10). The first item established above implies that this solution is unique for every $\boldsymbol{a}_{-h}$. An equilibrium is a profile $(\boldsymbol{a}^*, \boldsymbol{b}^*)$ such that $(\boldsymbol{a}_h^*, \boldsymbol{b}_h^*) = \mathrm{BR}_h(\boldsymbol{a}_{-h}^*)$ for every hospital h.[43]

The request probabilities in Equation (6) are continuous in $\boldsymbol{a}_{-h}$. Indeed, after conditioning on a patient's preferences, the expectation over the finitely many Bernoulli acceptance realizations is a multilinear function of the rivals' admission rates. The same argument applies to the demand aggregates and uninsured-surplus term in Equation (10). Hence, the hospital objective and each $L_{hj}(\boldsymbol{a}_{-h})$ are continuous in $\boldsymbol{a}_{-h}$.

[43]Because the expected-load constraints in Equation (22) depend on rivals' admission rates, the hospital game is formally a generalized Nash game. We use "equilibrium" for brevity.

For every $\boldsymbol{a}_{-h}$, the set $\Gamma_h(\boldsymbol{a}_{-h})$ is nonempty because it contains $(\mathbf{0}, \mathbf{0})$, and it is compact and convex. It also varies continuously with $\boldsymbol{a}_{-h}$. Upper hemicontinuity follows from its closed graph. To establish lower hemicontinuity, take $\boldsymbol{a}^n_{-h} \to \boldsymbol{a}_{-h}$ and $(\boldsymbol{a}_h, \boldsymbol{b}_h) \in \Gamma_h(\boldsymbol{a}_{-h})$, and define

$$t_n := \min\left\{1, \min_{\{j:a_{hj}>0\}} \frac{b_{hj}}{a_{hj} L_{hj}(\boldsymbol{a}^n_{-h})}\right\},$$

where the inner minimum over an empty set is interpreted as $+\infty$. Then $(t_n \boldsymbol{a}_h, \boldsymbol{b}_h) \in \Gamma_h(\boldsymbol{a}^n_{-h})$. Continuity of L_{hj} and feasibility at $\boldsymbol{a}_{-h}$ imply $t_n \to 1$, so these feasible points converge to $(\boldsymbol{a}_h, \boldsymbol{b}_h)$.

Berge's maximum theorem and the strict concavity established in the first item therefore imply that BR_h is single-valued and continuous (Aliprantis and Border, 2006, Theorems 17.31 and 17.6). Define

$$\mathcal{A} := \prod_{h \in \mathcal{H}} [0,1]^{J_h}, \qquad \mathcal{T}(\boldsymbol{a}) := \left(\mathrm{BR}^a_h(\boldsymbol{a}_{-h})\right)_{h \in \mathcal{H}}.$$

The set $\mathcal{A}$ is nonempty, compact, and convex, and $\mathcal{T}$ is a continuous self-map of $\mathcal{A}$. By Brouwer's fixed-point theorem, a fixed point $\boldsymbol{a}^*$ exists. Letting $\boldsymbol{b}^*_h := \mathrm{BR}^b_h(\boldsymbol{a}^*_{-h})$ for each h then completes an equilibrium. This proves the fourth item and completes the proof. □

Equilibrium uniqueness. Uniqueness requires imposing conditions under which the rate best-response map is a contraction. These conditions compare the curvature of each hospital's own problem with the effect of rivals' admission rates on its demand and payoff objects; they do not place any restriction directly on the best-response map.

For compactness, denote the coefficient multiplying a_{hj} in Equation (10) by

$$\pi^L_{hj}(\boldsymbol{a}_{-h}) := \sum_{k \in \mathcal{K}_j} \lambda_k R_{hk}(\boldsymbol{a}_{-h})(p_{hk} - c_{hk}) - c^V_{hj} \Lambda_{hj}(\boldsymbol{a}_{-h}) + \tau_{hj} \Delta S^u_{hj}(\boldsymbol{a}_{-h}). \tag{23}$$

For a fixed rival profile, change variables from admission rates to expected bed-day loads,

$$q_{hj} := a_{hj} L_{hj}(\boldsymbol{a}_{-h}).$$

Dropping temporarily the upper bound $a_{hj} \le 1$, the hospital chooses $(\boldsymbol{q}_h, \boldsymbol{b}_h)$ from the fixed set

$$\mathcal{Y}_h := \left\{(\boldsymbol{q}_h, \boldsymbol{b}_h) \ge \mathbf{0} : q_{hj} \le b_{hj} \text{ for every } j, \quad \sum_{j \in \mathcal{J}_h} b_{hj} \le B_h\right\}, \tag{24}$$

and its objective, extended to this set, is

$$\widetilde{U}_h(\boldsymbol{q}_h, \boldsymbol{b}_h; \boldsymbol{a}_{-h}) := \sum_{j \in \mathcal{J}_h} \left[q_{hj} \frac{\pi^L_{hj}(\boldsymbol{a}_{-h})}{L_{hj}(\boldsymbol{a}_{-h})} - q^2_{hj} c^V_{hj} \frac{\Lambda_{hj}(\boldsymbol{a}_{-h})^2}{L_{hj}(\boldsymbol{a}_{-h})^2} - \exp\left(q_{hj} \frac{\chi_{hj}(\boldsymbol{a}_{-h})}{L_{hj}(\boldsymbol{a}_{-h})} - c^o_h b_{hj} \right) \right]. \tag{25}$$

The maintained positive-demand and positive-length-of-stay conditions imply

$$\underline{L}_h := \min_{j \in \mathcal{J}_h} \min_{\boldsymbol{a}_{-h} \in \prod_{g \neq h} [0,1]^{J_g}} L_{hj}(\boldsymbol{a}_{-h}) > 0. \tag{26}$$

Moreover, $\widetilde{U}_h$ is uniformly strongly concave on $\mathcal{Y}_h$. To see this, for each line let

$$\Xi_{hj} := \exp\left(q_{hj} \frac{\chi_{hj}}{L_{hj}} - c^o_h b_{hj} \right),$$

where the dependence of Λ_{hj}, L_{hj}, and χ_{hj} on $\boldsymbol{a}_{-h}$ is suppressed. The negative Hessian block in (q_{hj}, b_{hj}) is

$$\begin{pmatrix} 2c^V_{hj} \dfrac{\Lambda^2_{hj}}{L^2_{hj}} + \dfrac{\chi^2_{hj}}{L^2_{hj}} \Xi_{hj} & -c^o_h \dfrac{\chi_{hj}}{L_{hj}} \Xi_{hj} \\ -c^o_h \dfrac{\chi_{hj}}{L_{hj}} \Xi_{hj} & (c^o_h)^2 \Xi_{hj} \end{pmatrix},$$

whose determinant is $2c^V_{hj}(\Lambda_{hj}/L_{hj})^2 (c^o_h)^2 \Xi_{hj} > 0$. Continuity and compactness therefore imply that

$$\alpha_h := \inf_{\substack{\boldsymbol{a}_{-h} \in \prod_{g \neq h} [0,1]^{J_g} \\ (\boldsymbol{q}_h, \boldsymbol{b}_h) \in \mathcal{Y}_h}} \lambda_{\min}\left(-\nabla^2_{(\boldsymbol{q}_h, \boldsymbol{b}_h)} \widetilde{U}_h(\boldsymbol{q}_h, \boldsymbol{b}_h; \boldsymbol{a}_{-h}) \right) > 0. \tag{27}$$

Thus α_h is a uniform lower bound on the hospital's own curvature.

The first uniqueness condition ensures that the temporarily omitted upper bound on admission rates is never relevant. Write $x^+ := \max\{x, 0\}$ and set

$$\bar{a}_h := \max_{j \in \mathcal{J}_h} \sup_{\boldsymbol{a}_{-h} \in \prod_{g \neq h} [0,1]^{J_g}} \min \left\{ \frac{B_h}{L_{hj}(\boldsymbol{a}_{-h})}, \frac{\left(\pi^L_{hj}(\boldsymbol{a}_{-h}) \right)^+}{2c^V_{hj} \Lambda_{hj}(\boldsymbol{a}_{-h})^2} \right\}. \tag{28}$$

The first term reflects the total capacity constraint; the second reflects the quadratic congestion cost. Either can keep full admission uniformly suboptimal.

We next define the strength of primitive interaction between hospitals. Let $\boldsymbol{y}_h := (\boldsymbol{q}_h, \boldsymbol{b}_h)$ and $\boldsymbol{L}_h(\boldsymbol{a}_{-h}) := (L_{hj}(\boldsymbol{a}_{-h}))_{j \in \mathcal{J}_h}$. For $h' \neq h$, define

$$\phi_{hh'} := \frac{1}{\underline{L}_h \alpha_h} \sup_{\substack{\boldsymbol{a}_{-h} \in \prod_{g \neq h} [0,1]^{J_g} \\ \boldsymbol{y}_h \in \mathcal{Y}_h}} \left\| \nabla^2_{\boldsymbol{y}_h, \boldsymbol{a}_{h'}} \widetilde{U}_h(\boldsymbol{y}_h; \boldsymbol{a}_{-h}) \right\|_2 + \frac{\bar{a}_h}{\underline{L}_h} \sup_{\boldsymbol{a}_{-h} \in \prod_{g \neq h} [0,1]^{J_g}} \left\| \frac{\partial \boldsymbol{L}_h(\boldsymbol{a}_{-h})}{\partial \boldsymbol{a}_{h'}} \right\|_2. \tag{29}$$

This coefficient contains only derivatives of demand and payoff primitives. For example, the line-j entries of the first mixed-derivative block are

$$\frac{\partial^2 \widetilde{U}_h}{\partial q_{hj} \partial a_{h'j'}} = \frac{\partial}{\partial a_{h'j'}} \left(\frac{\pi^L_{hj}}{L_{hj}} \right) - 2 q_{hj} c^V_{hj} \frac{\partial}{\partial a_{h'j'}} \left(\frac{\Lambda^2_{hj}}{L^2_{hj}} \right) - \exp\left(q_{hj} \frac{\chi_{hj}}{L_{hj}} - c^o_h b_{hj} \right) \left(1 + q_{hj} \frac{\chi_{hj}}{L_{hj}} \right) \frac{\partial}{\partial a_{h'j'}} \left(\frac{\chi_{hj}}{L_{hj}} \right),$$

$$\frac{\partial^2 \widetilde{U}_h}{\partial b_{hj} \partial a_{h'j'}} = c^o_h q_{hj} \exp\left(q_{hj} \frac{\chi_{hj}}{L_{hj}} - c^o_h b_{hj} \right) \frac{\partial}{\partial a_{h'j'}} \left(\frac{\chi_{hj}}{L_{hj}} \right).$$

The derivatives of R_{hk} and ΔS^u_{hj} entering these expressions are obtained directly by differentiating the finite expectation over rivals' Bernoulli acceptance draws. Hence $\phi_{hh'}$ can be bounded analytically from utilities, illness rates, lengths of stay, revenues, and costs, or computed numerically over the compact sets in Equation (29); it does not require solving for a best response or an equilibrium.

Assumption 1. *For every hospital h, $\bar{a}_h < 1$, and*

$$\kappa := \max_{h \in \mathcal{H}} \sum_{h' \neq h} \phi_{hh'} < 1. \tag{30}$$

The rate condition is used only to place every hospital problem on the fixed set $\mathcal{Y}_h$. The second condition requires primitive cross-hospital interaction to be dominated by the hospital's own curvature.

Lemma 1. *Under Assumption 1, for any two admission-rate profiles $\boldsymbol{a}$ and $\boldsymbol{a}'$,*

$$\left\| \mathrm{BR}^a_h(\boldsymbol{a}_{-h}) - \mathrm{BR}^a_h(\boldsymbol{a}'_{-h}) \right\|_2 \leq \sum_{h' \neq h} \phi_{hh'} \left\| \boldsymbol{a}_{h'} - \boldsymbol{a}'_{h'} \right\|_2. \tag{31}$$

Proof. Fix $\boldsymbol{a}_{-h}$. Strong concavity gives a unique maximizer $(\boldsymbol{q}^*_h, \boldsymbol{b}^*_h)$ of $\widetilde{U}_h$ over $\mathcal{Y}_h$. At this maximizer, $q^*_{hj} \leq b^*_{hj} \leq B_h$, so $q^*_{hj}/L_{hj} \leq B_h/L_{hj}$. If $q^*_{hj} > 0$, lowering q_{hj} is feasible, and first-order optimality implies

$$0 \leq \frac{\pi^L_{hj}}{L_{hj}} - 2 q^*_{hj} c^V_{hj} \frac{\Lambda^2_{hj}}{L^2_{hj}} - \frac{\chi_{hj}}{L_{hj}} \exp\left(q^*_{hj} \frac{\chi_{hj}}{L_{hj}} - c^o_h b^*_{hj} \right).$$

Therefore, $q^*_{hj}/L_{hj} \leq (\pi^L_{hj})^+ / (2 c^V_{hj} \Lambda^2_{hj})$; the same conclusion is immediate if $q^*_{hj} = 0$. By Equation (28) and the first part of Assumption 1, $q^*_{hj}/L_{hj} \leq \bar{a}_h < 1$. Thus the omitted constraint $q_{hj} \leq L_{hj}$ is slack, the fixed-set maximizer solves the original hospital problem, and

$$\mathrm{BR}^a_{hj}(\boldsymbol{a}_{-h}) = \frac{q^*_{hj}(\boldsymbol{a}_{-h})}{L_{hj}(\boldsymbol{a}_{-h})}.$$

Let $\boldsymbol{y}^*_h(\boldsymbol{a}_{-h})$ denote the fixed-set maximizer. The variational inequalities for the maximizers at

$\boldsymbol{a}_{-h}$ and $\boldsymbol{a}'_{-h'}$, together with α_h-strong concavity, imply

$$\begin{aligned}\left\|\boldsymbol{y}_h^*(\boldsymbol{a}_{-h}) - \boldsymbol{y}_h^*(\boldsymbol{a}'_{-h})\right\|_2 &\le \frac{1}{\alpha_h}\left\|\nabla_{\boldsymbol{y}_h}\widetilde{U}_h(\boldsymbol{y}_h^*(\boldsymbol{a}'_{-h}); \boldsymbol{a}_{-h}) - \nabla_{\boldsymbol{y}_h}\widetilde{U}_h(\boldsymbol{y}_h^*(\boldsymbol{a}'_{-h}); \boldsymbol{a}'_{-h})\right\|_2 \\ &\le \frac{1}{\alpha_h}\sum_{h'\neq h}\left(\sup_{\substack{\widetilde{\boldsymbol{a}}_{-h}\in\prod_{g\neq h}[0,1]^{J_g}\\ \boldsymbol{y}_h\in\mathcal{Y}_h}}\left\|\nabla^2_{\boldsymbol{y}_h,\boldsymbol{a}_{h'}}\widetilde{U}_h(\boldsymbol{y}_h;\widetilde{\boldsymbol{a}}_{-h})\right\|_2\right)\left\|\boldsymbol{a}_{h'} - \boldsymbol{a}'_{h'}\right\|_2,\end{aligned}$$

where the second inequality follows by integrating the mixed derivative along the line segment joining the two rival profiles. The same argument applied to $\boldsymbol{L}_h$ gives

$$\left\|\boldsymbol{L}_h(\boldsymbol{a}_{-h}) - \boldsymbol{L}_h(\boldsymbol{a}'_{-h})\right\|_2 \le \sum_{h'\neq h}\left(\sup_{\widetilde{\boldsymbol{a}}_{-h}\in\prod_{g\neq h}[0,1]^{J_g}}\left\|\frac{\partial \boldsymbol{L}_h(\widetilde{\boldsymbol{a}}_{-h})}{\partial \boldsymbol{a}_{h'}}\right\|_2\right)\left\|\boldsymbol{a}_{h'} - \boldsymbol{a}'_{h'}\right\|_2.$$

Finally, applying the triangle inequality componentwise to $\mathrm{BR}^a_{hj} = q^*_{hj}/L_{hj}$ yields

$$\begin{aligned}&\left\|\mathrm{BR}^a_h(\boldsymbol{a}_{-h}) - \mathrm{BR}^a_h(\boldsymbol{a}'_{-h})\right\|_2 \\ &\quad\le \frac{1}{\underline{L}_h}\left\|\boldsymbol{q}^*_h(\boldsymbol{a}_{-h}) - \boldsymbol{q}^*_h(\boldsymbol{a}'_{-h})\right\|_2 + \frac{\overline{a}_h}{\underline{L}_h}\left\|\boldsymbol{L}_h(\boldsymbol{a}_{-h}) - \boldsymbol{L}_h(\boldsymbol{a}'_{-h})\right\|_2.\end{aligned}$$

Substituting the preceding two bounds and using Equation (29) proves Equation (31). □

Proposition 3 (Uniqueness)**.** *Under the maintained conditions of Proposition 1 and Assumption 1, the equilibrium is unique.*

Proof. Existence follows from the fourth item of Proposition 1. By Lemma 1, for any two admission-rate profiles,

$$\|\mathcal{T}(\boldsymbol{a}) - \mathcal{T}(\boldsymbol{a}')\|_{\infty,2} \le \kappa\|\boldsymbol{a} - \boldsymbol{a}'\|_{\infty,2}, \qquad \|\boldsymbol{a}\|_{\infty,2} := \max_{h\in\mathcal{H}}\|\boldsymbol{a}_h\|_2.$$

Assumption 1 gives $\kappa < 1$, so $\mathcal{T}$ is a contraction and has a unique fixed point. The corresponding bed allocations are also unique because each hospital's joint best response is unique by the first item of Proposition 1. □

Assumption 1 is sufficient, not necessary. Its interaction coefficients are constructed from the primitive payoff rather than from the best-response Jacobian. The rate condition removes a moving constraint, while the row-sum condition is a primitive diagonal-dominance requirement: own curvature must dominate the demand and payoff changes induced by rivals. Failure of the condition therefore does not imply multiplicity; it means only that this global contraction argument does not establish uniqueness.

A.4 Identification and Estimation Appendix

A.4.1 Derivations. We begin by showing the derivation of the likelihood in Equation (11). As in the main text, define $\bar{\sigma}_{idht}$ as the standard logit choice probability without the admissions constraint. The proof is for a single patient, diagnosis, and time, so we omit those subindices for clarity.

$$\begin{aligned}
\sigma_h &= \mathbb{E}[\frac{\tilde{a}_h e^{v_h}}{1+\sum_{h'\in\mathcal{H}}\tilde{a}_{h'}e^{v_{h'}}}] \\
&= \mathbb{E}[\frac{\tilde{a}_h \bar{\sigma}_h}{\bar{\sigma}_0+\sum_{h'\in\mathcal{H}}\tilde{a}_{h'}\bar{\sigma}_{h'}}] \\
&= a_h\bar{\sigma}_h\mathbb{E}[(1+(\bar{\sigma}_0+\bar{\sigma}_h-1)+\sum_{h'\in\mathcal{H}\setminus\{h\}}\tilde{a}_{h'}\bar{\sigma}_{h'})^{-1}] \\
&= a_h\bar{\sigma}_h\mathbb{E}[(1-\sum_{h'\in\mathcal{H}\setminus\{h\}}(1-\tilde{a}_{h'})\bar{\sigma}_{h'})^{-1}] \\
&= a_h\bar{\sigma}_h\mathbb{E}[\int_0^\infty \exp(-u(1-\sum_{h'\in\mathcal{H}\setminus\{h\}}(1-\tilde{a}_{h'})\bar{\sigma}_{h'}))du] \\
&= a_h\bar{\sigma}_h\int_0^\infty e^{-u}\mathbb{E}[\prod_{h'\in\mathcal{H}\setminus\{h\}}\exp(u(1-\tilde{a}_{h'})\bar{\sigma}_{h'})du] \\
&= a_h\bar{\sigma}_h\int_0^\infty e^{-u}\prod_{h'\in\mathcal{H}\setminus\{h\}}\left(a_{h'}+(1-a_{h'})e^{u\bar{\sigma}_{h'}}\right)du
\end{aligned}$$

Where the second equality follows from dividing the numerator and denominator by $(1+\sum_{h\in\mathcal{H}}e^{v_h})$, the third follows from expanding the two values of $\tilde{a}_h$, the fourth from re-organizing shares, the fifth from the identity $y^{-1}=\int_0^\infty e^{-ty}dt$ for $y>0$, the sixth from the linearity of expectations, and the last one from the independence of admission events.

A.4.2 Admission Fixed Point. Here we prove that the log-transformed Equation (12) is a contraction mapping, under some regularity conditions. We begin by establishing these conditions, and then prove the result.

Assumption 2 (Acceptance Fixed Point Assumptions)**.** *The following hold for every period t,*

1. ***Regularity****: every hospital $h\in\mathcal{H}$ and line $j\in\mathcal{J}_h$, has $N^{L\star}_{hj}>0$ and $\mathcal{I}^*_{hj}\neq\emptyset$. For every patient $i\in\mathcal{I}_{jt}$, $\ell^*_i>0$ and $|v_{ih}|<\infty$ for every $h\in\mathcal{H}_i$.*
2. ***Interior Admission Rate****: $N^{L\star}_{hj}<\sum_{i\in\mathcal{I}_{hj}}\bar{\sigma}_{ih}\ell^*_i$ for every h,h*
3. ***Bounded Diversion****: $\sum_{i\in\mathcal{I}^*_{jt}}\frac{e^{v_{ih}}}{1+e^{v_{ih}}}\ell^*_i<2\sum_{i\in\mathcal{I}^*_{jt}}\bar{\sigma}_{ih}\ell^*_i$*

The first assumption is minimal and simply states that every hospital we observe in the data has positive total length of stay, every patient admitted has positive length of stay, and finite preferences. The second item in the assumption states that the total admitted length of

stay is strictly smaller than would be predicted under simple logit demand, which implies that $a_{hj} \in (0,1)$. Finally, the last assumption states that a monopolist hospital obtains less than twice its present demand. Essentially, this implies that the market is spread out enough such that if all other hospitals were at capacity, more consumers would prefer the out-of-market option than the monopolist hospital.

Before proving the main result, we present an intermediary lemma.

Lemma 2 (Elasticity Bound)**.** *Under Assumption 2.1, it holds that*

$$0 \leq \sum_{h' \in \mathcal{H}\setminus\{h\}} a_{h'j(i,h')}\left(-\frac{\partial R_{ih}}{\partial a_{h'j(i,h')}}\right) \leq \frac{e^{v_{ih}}}{1+e^{v_{ih}}} - R_{ih}$$

Proof. In what follows, we omit the line index for notational simplicity, leaving it implicit that all the arguments are for the line in which a fixed patient i seeks care.

Note that

$$R_{ih} = \mathbb{E}\left[\frac{e^{v_{ih}}}{1+e^{v_{ih}}+\sum_{h'\neq h} e^{v_{ih'}}\tilde{a}_{h'}}\right] = a_{h'}\mathbb{E}\left[\frac{e^{v_{ih}}}{1+e^{v_{ih}}+e^{v_{ih'}}+\sum_{h''\notin\{h,h'\}} e^{v_{ih'}}\tilde{a}_{h''}}\right] + (1-a_{h'})\mathbb{E}\left[\frac{e^{v_{ih}}}{1+e^{v_{ih}}+\sum_{h''\notin\{h,h'\}} e^{v_{ih'}}\tilde{a}_{h''}}\right]$$

Denote temporarily $D_{ih'} = 1+e^{v_{ih}}+\sum_{h''\notin\{h,h'\}} e^{v_{ih'}}\tilde{a}_{h''}$. We have that $\frac{\partial R_{ih}}{\partial a_{h'}} = -\mathbb{E}\left[\frac{e^{v_{ih}+v_{ih'}}}{D_{ih'}(D_{ih'}+e^{v_{ih'}})}\right] < 0$. Now note that $D_{ih'} \geq 1+e^{v_{ih}}$ therefore

$$\begin{aligned}
\sum_{h'\in\mathcal{H}\setminus\{h\}} a_{h'}\left(-\frac{\partial R_{ih}}{\partial a_{h'}}\right) &= \sum_{h'\in\mathcal{H}\setminus\{h\}} a_{h'}\mathbb{E}\left[\frac{e^{v_{ih}+v_{ih'}}}{D_i(D_i+e^{v_{ih'}})}\right] = \sum_{h'\in\mathcal{H}\setminus\{h\}} \mathbb{E}\left[\frac{\tilde{a}_{h'}e^{v_{ih}+v_{ih'}}}{D_i(D_i+\tilde{a}_{h'}e^{v_{ih'}})}\right] \\
&= \mathbb{E}\left[\frac{e^{v_{ih}}}{\bar{D}_i}\sum_{h'\in\mathcal{H}\setminus\{h\}}\frac{\tilde{a}_{h'}e^{v_{ih'}}}{D_{ih'}}\right] \leq \mathbb{E}\left[\frac{e^{v_{ih}}}{\bar{D}_i}\frac{\sum_{h'\in\mathcal{H}\setminus\{h\}}\tilde{a}_{h'}e^{v_{ih'}}}{1+e^{v_{ih}}}\right] \\
&= \mathbb{E}\left[\frac{e^{v_{ih}}}{\bar{D}_i}\frac{\bar{D}_i-(1+e^{v_{ih}})}{1+e^{v_{ih}}}\right] = \frac{e^{v_{ih}}}{1+e^{v_{ih}}} - R_{ih}
\end{aligned}$$

where $\bar{D}_i = 1+e^{v_{ih}}+\sum_{h'\neq h} e^{v_{ih'}}\tilde{a}_{h'}$. □

Denote $T_{hj}(\boldsymbol{a}) = \frac{N_{hj}^{L\star}}{\sum_{i\in\mathcal{I}_{jt}^*} R_{ih}(\boldsymbol{a})\ell_i^*}$ the base fixed-point. Note that $R_{ih}(\boldsymbol{a}) \geq \bar{o}_{ih} > 0$ and $R_{ih}(\boldsymbol{a}) < R_{ih}(\mathbf{0}) = \frac{e^{v_{ih}}}{1+e^{v_{ih}}}$. Therefore, $\frac{N_{hj}^{L\star}}{\sum_i \bar{o}_{ij}\ell_i^*} \geq T_{hj}(\boldsymbol{a}) \geq \frac{N_{hj}^{L\star}}{\sum_i R_{ih}(\mathbf{0})\ell_i^*}$. We denote the bounds on the map as $\underline{a}_{hj}$ and $\bar{a}_{hj}$. We are now ready to prove the main result

Proposition 4 (Admission Rate Contraction Mapping)**.** *Suppose Assumption 2 holds.Let* $\mathcal{A} = \prod_{(h,j)}[\ln(\underline{a}_{hj}), \ln(\bar{a}_{hj})]$ *and define the log-transformation of* T*,* $\phi : \mathcal{A} \to \mathcal{A}$ *as*

$$\phi_{hj}(\alpha) = \ln(N_{hj}^{L\star}) - \ln\left(\sum_{i\in\mathcal{I}_{hj}^*} R_{ih}(e^{\alpha})\ell_i^*\right)$$

Then, ϕ *is a contraction on* $(\mathcal{A}, \|\cdot\|_\infty)$*.*

Proof. Note that $\mathcal{A}$ is a non-empty, convex, and compact subset of $\mathbb{R}^{HJ}$; endowed with the sup-metric, it is a complete metric space. It is direct to see that ϕ is C^∞ on the open set $(-\infty, 0)^{HJ}$. Since $T([0,1]^{hj}) \subseteq \prod[\underline{a}_{hj}, \bar{a}_{hj}]$, $\phi(\mathcal{A}) \subseteq \mathcal{A}$.

Next, we note that

$$\frac{\partial \phi_{hj}}{\partial \alpha_{h'j'}} = \frac{\sum_{i \in \mathcal{I}_j^*} a_{h'j'} \ell_i^* (-\frac{\partial R_{ih}}{\partial a_{h'j'}})}{\sum_{i \in \mathcal{I}_j^*} \ell_i^* R_{ih}} \geq 0$$

We note that the diagonal of the implied Jacobian is zero, since R_{ih} is not a function of a_h, and the off-diagonal items are positive since a higher acceptance at rivals decreases admission at h. Now, summing the off-diagonal terms, we get

$$\|\nabla \phi_{hj}(\boldsymbol{\alpha})\|_1 = \frac{\sum_{i \in \mathcal{I}_j^*} \sum_{h' \neq h} a_{h'j'} \ell_i^* (-\frac{\partial R_{ih}}{\partial a_{h'j'}})}{\sum_{i \in \mathcal{I}_j^*} \ell_i^* R_{ih}} \leq \frac{\sum_i \ell_i^* (R_{ih}(\mathbf{0}) - R_{ih})}{\sum_{i \in \mathcal{I}_j^*} \ell_i^* R_{ih}} = \frac{\sum_i \ell_i^* R_{ih}(\mathbf{0})}{\sum_{i \in \mathcal{I}_j^*} \ell_i^* R_{ih}} - 1 \leq \frac{\sum_i \ell_i^* R_{ih}(\mathbf{0})}{\sum_{i \in \mathcal{I}_j^*} \ell_i^* \bar{\sigma}_{ih}} - 1$$

Where the first inequality follows from the lemma above and the second from $R_{ih} \geq \bar{\sigma}_{ih}$. By assumption 2 the last term $\beta \equiv \frac{\sum_i \ell_i^* R_{ih}(\mathbf{0})}{\sum_{i \in \mathcal{I}_j^*} \ell_i^* \bar{\sigma}_{ih}} - 1$ is positive and less than 1. Therefore, by the mean value theorem, we have that for every $\boldsymbol{\alpha}, \boldsymbol{\alpha}' \in \mathcal{A}$ there exists $\tilde{\boldsymbol{\alpha}}$ on a segment between $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}'$ such that

$$|\phi_{hj}(\boldsymbol{\alpha}) - \phi_{hj}(\boldsymbol{\alpha}')| = |\nabla \phi_{hj}(\tilde{\boldsymbol{\alpha}})(\boldsymbol{\alpha} - \boldsymbol{\alpha}')| \leq \|\nabla \phi_{hj}(\tilde{\boldsymbol{\alpha}})\|_1 \|\boldsymbol{\alpha} - \boldsymbol{\alpha}'\|_\infty \leq \beta \|\boldsymbol{\alpha} - \boldsymbol{\alpha}'\|_\infty$$

where the first inequality is Hölder's inequality and the last one follows from the derivation above. Since β is positive and strictly smaller than unity, ϕ is a contraction map. □

A.4.3 Demand Estimation Details.

A.4.4 Supply Identification. In this appendix section, we prove the global identification of the hospital's cost parameters. The proof is hospital-by-hospital and omits a time index. In what follows, we fix a hospital h and denote by s the state of the hospital, which indexes all the demand and revenue elements that the hospital conditions on when solving its optimization problem. We note that we are abusing notation here, since s in the main text is briefly used for a population segment. We denote the distribution of s by $\mathcal{D}_h$, which is a marginal of the data-generating process underlying the data. The argument exploits the structure of the first-order condition in Equation (13): the net marginal revenue of an admission splits into a term that is polynomial in the admission rates and carries the linear parameters $\boldsymbol{\theta}_{hj} = (c_{hj}^S, c_{hj}^F, c_{hj}^V, \tau_{hj})$, and the exponential marginal capacity cost term which carries the overflow rate c_h^o. Because polynomials and exponentials with distinct frequencies are linearly independent on open sets, sufficient variation in reimbursements—holding demand composition and capacity fixed—pins

down the overflow rate; variation in case mix and in charitable-care demand then separates the remaining parameters.

We begin by establishing notation. The first-order condition for the admission rate problem of Equation (10) states that for any line with $a_{hj} \in (0,1)$ it holds that:

$$\bar{p}_{hj} = \boldsymbol{x}'_{hj}\boldsymbol{\theta}_{hj} + k_{hj}(c^o_h) \tag{32}$$

Where $\bar{p}_{hj} \equiv \sum_{k\in\mathcal{K}_j} \lambda_k R_{hk} p_{hk}$ is the expected revenue of the line, $\boldsymbol{x}'_{hj} \equiv (\sum_{k\in\mathcal{K}_j} \lambda_k R_{hk}\omega_k, \Lambda_{hj}, \Lambda_{hj}(1 + 2\Lambda_{hj}a_{hj}), -\Delta S^u_{hj})$ collects all the linear terms, and $k_{hj}(c^o) \equiv \chi_{hj}(c^o_h)\exp(\chi_{hj}(c^o)a_{hj} - c^o_h b_{hj}) + \nu_{hj}\Lambda_{hj}\mu^L_{hj}$, where ν_{hj} is the Lagrange multiplier on the lines ex-ante feasibility constraint. We note that, unlike in the main text, here we work with the first-order conditions not expressed in terms of per-expected admission level. In addition, we establish three shorthands. First, the average resource intensity of the line's potential admissions,

$$\bar{\omega}_{hj} = \Lambda^{-1}_{hj} \sum_{k\in\mathcal{K}_j} \omega_k \lambda_k R_{hk},$$

so that the first entry of $\boldsymbol{x}_{hj}$ in Equation (32) is $\Lambda_{hj}\bar{\omega}_{hj}$. Second, we make the dependence of the main text's χ_{hj} on the overflow rate explicit, writing, for a rate t,

$$\chi_{hj}(t) = \Lambda_{hj}\Big(M^L_{hj}(t) - 1\Big), \qquad \text{so that } \chi_{hj} = \chi_{hj}(c^o_h).$$

Third, we define the slack margin $\delta_{hj}(t) = \chi_{hj}(t) - t\,\Lambda_{hj}\mu^L_{hj}$, which will separate slack from binding lines in the bed program below.

Combining the allocation stage (7) with the deterministic-equivalent line problem (10) gives the hospital's joint problem

$$\max_{\substack{\boldsymbol{a}\in[0,1]^{J_h}\\ \boldsymbol{b}\geq\boldsymbol{0}}} \sum_{j\in\mathcal{J}_h} \Big[\Big(\bar{p}_{hj} - \Lambda_{hj}(c^S_{hj}\bar{\omega}_{hj} + c^F_{hj} + c^V_{hj}) + \tau_{hj}\Delta S^u_{hj}\Big)a_{hj} - c^V_{hj}\Lambda^2_{hj}a^2_{hj} - e^{\chi_{hj}(c^o_h)a_{hj} - c^o_h b_{hj}}\Big] \tag{33}$$

subject to $\Lambda_{hj}\mu^L_{hj}a_{hj} \leq b_{hj}$ for every j and $\sum_j b_{hj} \leq B_h$. In what follows, we omit the h index from c^o_h unless needed for clarity. When $c^V_{hj} > 0$ for all j the objective is strictly jointly concave in $(\boldsymbol{a}, \boldsymbol{b})$ and the optimizer $\Big(\boldsymbol{a}_h(s;\boldsymbol{\theta}, c^o), \boldsymbol{b}_h(s;\boldsymbol{\theta}, c^o)\Big)$ is unique by main text Proposition 1. We denote $\boldsymbol{a}^*_h(s) := \boldsymbol{a}_h(s;\boldsymbol{\theta}^*, c^{o*})$ the optimal admission-rate at the true parameter.

Let $\nu_{hj} \geq 0$ denote the multiplier on line j's ex-ante capacity constraint and $\eta_h \geq 0$ the multiplier on the hospital's total resource constraint. Because the overflow term is strictly decreasing in b_{hj}, the first-order condition in b_{hj},

$$c^o\, e^{\chi_{hj}(c^o)a_{hj} - c^o b_{hj}} = \eta_h - \nu_{hj}, \tag{34}$$

forces $\eta_h > 0$, meaning that capacity always binds. Intuitively, capacity is always better allocated

than not allocated, given each line's control over admission rates.

At any interior admission rate $a_{hj} \in (0,1)$, stationarity in a_{hj} yields Equation (32), restated here for clarity

$$\bar{p}_{hj} = \boldsymbol{x}_{hj}(a_{hj})'\boldsymbol{\theta}_{hj} + k_{hj}, \qquad \boldsymbol{x}_{hj}(a) := \begin{pmatrix} \Lambda_{hj}\bar{\omega}_{hj} \\ \Lambda_{hj} \\ \Lambda_{hj}(1 + 2\Lambda_{hj}a) \\ -\Delta S^{u}_{hj} \end{pmatrix}, \tag{35}$$

where

$$k_{hj} := \chi_{hj}(c^o)\, e^{\chi_{hj}(c^o)a_{hj} - c^o b_{hj}} + \nu_{hj}\Lambda_{hj}\mu^{L}_{hj} \tag{36}$$

is the marginal capacity cost: the marginal expected overflow cost of an admission plus the shadow value of the bed-days it consumes. Although k_{hj} is defined through the unobserved allocation (b_{hj}, ν_{hj}), Lemma 4 below shows that it is a known closed-form functional of the admission rates, the total capacity, and the rate c^o alone; the internal allocation never needs to be observed.

Before stating our results, we define what a regular state is, what we mean by global identification, and our key set of assumptions.

Definition 1 (Regular states; global identification)**.** *The set of* regular states *is* $\mathcal{R}_h := \left\{ s \in \mathrm{supp}(\mathcal{D}_h) : \boldsymbol{a}^*_h(s) \in (0,1)^{J_h} \text{ and } \underline{B}\left(\boldsymbol{a}^*_h(s)\right) < B_h(s) \right\}$*, where* $\underline{B}(\boldsymbol{a}) := \sum_j \Lambda_{hj}\mu^L_{hj}a_{hj}$ *is the expected admitted load. The parameter* $(\boldsymbol{\theta}^*, c^{o*})$ *is globally identified on a parameter space* Θ_h *if for any* $(\boldsymbol{\theta}', c^{o\prime}) \in \Theta_h$ *such that* $\boldsymbol{a}_h(s; \boldsymbol{\theta}', c^{o\prime}) = \boldsymbol{a}_h(s; \boldsymbol{\theta}^*, c^{o*})$ *for all* $s \in \mathrm{supp}(\mathcal{D}_h)$ *implies that* $(\boldsymbol{\theta}', c^{o\prime}) = (\boldsymbol{\theta}^*, c^{o*})$.

Assumption 3 (Identification conditions)**.**

1. (Affine surgical costs.) $c_{hk} = c^S_{hj(k)}\,\omega_k + c^F_{hj(k)}$ *for every* k*, where* $j(k)$ *is the line serving type* k.
2. (Common overflow rate.) c^o_h *applies to every* $j \in \mathcal{J}_h$.
3. (Nondegenerate primitives.) *For every* $s \in \mathrm{supp}(\mathcal{D}_h)$ *and* $j \in \mathcal{J}_h$*,* $\Lambda_{hj}(s) > 0$*; and each type's length of stay* ℓ^k *is strictly positive with a finite exponential moment,* $\bar{c}_k := \sup\{t \geq 0 : \mathbb{E}[e^{t\ell^k}] < \infty\} > 0$*, so that* $\bar{c}^o_h := \min_k \bar{c}_k > 0$.
4. (Parameter space.) $(\boldsymbol{\theta}, c^o) \in \Theta_h := \left\{ (\boldsymbol{\theta}, c^o) :\ c^V_{hj} > 0,\ \tau_{hj} \geq 0\ \forall j \in \mathcal{J}_h,\ c^o \in (0, \bar{c}^o_h) \right\}$.
5. (Case-mix and charitable-demand relevance.) *For each line* j*, the realized set*

$$\left\{ \left(\bar{\omega}_{hj}(s),\ \Lambda_{hj}(s)\, a^*_{hj}(s),\ \Delta S^u_{hj}(s)/\Lambda_{hj}(s) \right) : s \in \mathcal{R}_h \right\} \subseteq \mathbb{R}^3$$

   is not contained in a single affine plane (equivalently, it contains four affinely independent points).
6. (Conditional price variation in a full-slack region.) *There exist values* $\left(\Lambda^0_{hj}, \bar{\omega}^0_{hj}, M^{L,0}_{hj}(\cdot), \Delta S^{u,0}_{hj} \right)_{j \in \mathcal{J}_h}$ *and* B^0_h *of the line-level composites, and a nonempty open set* $U \subseteq \mathbb{R}^{J_h}$*, such that the conditioned state*

set

$$O := \Big\{ s \in \text{supp}(\mathcal{D}_h) : \Big(\Lambda_{hj}, \bar{\omega}_{hj}, M^L_{hj}, \Delta S^u_{hj}, B_h\Big)(s) = \Big(\Lambda^0_{hj}, \bar{\omega}^0_{hj}, M^{L,0}_{hj}, \Delta S^{u,0}_{hj}, B^0_h\Big)\ \forall j,\ \ \bar{\boldsymbol{p}}_h(s) \in U \Big\}$$

satisfies: (a) price richness, $\bar{\boldsymbol{p}}_h(O) = U$*; and (b)* interior, full-slack truth, *for every* $s \in O$, $\boldsymbol{a}^*(s) \in (0,1)^{J_h}$ *and*

$$E_{\text{fs}}\Big(\boldsymbol{a}^*(s), B^0_h; c^{o*}\Big) < \delta_{hj}(c^{o*})\, a^*_{hj}(s) \quad \textit{for every } j, \qquad E_{\text{fs}}(\boldsymbol{a}, B; c^o) := \frac{1}{J_h}\Big[\sum_j \chi_{hj}(c^o) a_{hj} - c^o B\Big].$$

Parts (i)–(iv) are maintained model restrictions: (i) restates the cost parametrization adopted in Section 6; (ii) restates the overflow-cost specification of Section 5, which makes bed-days a fungible resource priced by a single internal shadow price, so that the clearing analysis below aggregates the hospital's capacity market into a one-dimensional index; (iii) restates the primitives of the model, matching the hypotheses of Proposition 1; and (iv) defines the space on which identification is claimed. Parts (v)–(vi) are the substantive variation requirements. Part (vi) is an "identification at a configuration" condition: reimbursement variation (e.g., from fee-schedule and payer-mix changes) that is rich in all J_h directions while market size, case mix, length-of-stay composition, charitable demand, and capacity are held fixed, on a region where the hospital operates with slack in every line; part (vi)(b) holds when capacity is large relative to expected load at the truth, for instance when arrival intensities are small relative to B^0_h. The conditions in this assumption imply that lines are slack at states in O, which implies that they are regular as we defined above. Note also that part (v) cannot be satisfied by states in O alone: on O the triples of part (v) all lie on the vertical line through $(\bar{\omega}^0_{hj}, \cdot, \Delta S^{u,0}_{hj}/\Lambda^0_{hj})$, so the case-mix and charitable-demand variation it requires must come from states outside O.

Having established our assumption, we move to prove two preliminary lemmas that will be used to establish identification.

Lemma 3 (Curvature of the length-of-stay transform)**.** *Under Assumption 3.3, for every line* j*, every state, and every* $t \in (0, \bar{c}^o_h)$*:* $\delta_{hj}(t) > 0$*, and both* $\chi_{hj}(t)$ *and* $\delta_{hj}(t)$ *are strictly increasing on* $(0, \bar{c}^o_h)$*.*

Proof. The moment generating function for length of stay, $M^L_{hj}(t) = \sum_{k \in \mathcal{K}_j} \frac{\lambda_k R_{hk}}{\Lambda_{hj}} \mathbb{E}[e^{t\ell^k}]$ is finite on $[0, \bar{c}^o_h)$ with $M^L_{hj}(0) = 1$, $(M^L_{hj})'(0) = \mu^L_{hj}$, and $(M^L_{hj})''(t) = \mathbb{E}_{hj}[\ell^2 e^{t\ell}] > 0$ because lengths of stay are strictly positive. Strict convexity places M^L_{hj} strictly above its tangent at zero, $M^L_{hj}(t) > 1 + t\mu^L_{hj}$ for $t > 0$; multiplying by $\Lambda_{hj} > 0$ gives $\delta_{hj}(t) > 0$. Moreover $\chi'_{hj}(t) = \Lambda_{hj}(M^L_{hj})'(t) > 0$ and $\delta'_{hj}(t) = \Lambda_{hj}\Big((M^L_{hj})'(t) - \mu^L_{hj}\Big) > 0$, the latter because $(M^L_{hj})'$ is strictly increasing with $(M^L_{hj})'(0) = \mu^L_{hj}$. □

Lemma 4 (Bed allocation and capacity clearing)**.** *Suppose Assumptions 3.2 and 3.3 hold. Fix a state, a rate* $c^o \in (0, \bar{c}^o_h)$*, and* $\boldsymbol{a} \in (0,1)^{J_h}$ *with slack* $\underline{B}(\boldsymbol{a}) < B_h$*. Given* $\boldsymbol{a}$*, the bed allocation solving* (33) *is the unique*

solution of the strictly concave program $\max_{\boldsymbol{b}}\{-\sum_j e^{\chi_{hj}(c^o)a_{hj}-c^o b_{hj}}\}$ *subject to the two constraint families, and:*

1. $\sum_j b_{hj} = B_h$ *and* $b_{hj} = \max\Big\{(\chi_{hj}(c^o)a_{hj} - E)/c^o,\ \Lambda_{hj}\mu_{hj}^L a_{hj}\Big\}$, *where* $E = E(\boldsymbol{a}, B_h; c^o)$ *is the unique solution of the capacity-clearing equation*

$$\sum_{j\in\mathcal{J}_h} \max\Big\{\frac{\chi_{hj}(c^o)a_{hj} - E}{c^o},\ \Lambda_{hj}\mu_{hj}^L a_{hj}\Big\} = B_h. \tag{37}$$

   The multipliers are unique and given by $\eta_h = c^o e^E$ *and* $\nu_{hj} = c^o \max\Big\{e^E - e^{\delta_{hj}(c^o)a_{hj}},\ 0\Big\}$; *consequently*

$$k_{hj}(\boldsymbol{a}, B_h; c^o) = \begin{cases} \chi_{hj}(c^o)\, e^E, & E \le \delta_{hj}(c^o)a_{hj} \ \ (\textit{slack line}), \\ c^o\Lambda_{hj}\mu_{hj}^L\, e^E + \delta_{hj}(c^o)\, e^{\delta_{hj}(c^o)a_{hj}}, & E \ge \delta_{hj}(c^o)a_{hj} \ \ (\textit{binding line}), \end{cases} \tag{38}$$

   the two branches agreeing when $E = \delta_{hj}(c^o)a_{hj}$.

2. *At the clearing price* E *at least one line is strictly slack:* $E < \delta_{hj}(c^o)a_{hj}$ *for some* j.

3. $(\boldsymbol{a}, B_h) \mapsto E(\boldsymbol{a}, B_h; c^o)$ *is continuous on* $\{\underline{B}(\boldsymbol{a}) < B_h\}$.

4. *If* $\tilde{\mathcal{J}}_h \subseteq \mathcal{J}_h$ *is nonempty with* $E \le \delta_{hj}(c^o)a_{hj}$ *for all* $j \in \tilde{\mathcal{J}}_h$ *and* $E \ge \delta_{hj}(c^o)a_{hj}$ *for all* $j \notin \tilde{\mathcal{J}}_h$, *then*

$$E = E_{\tilde{\mathcal{J}}_h}(\boldsymbol{a}, B_h; c^o) := \frac{1}{\tilde{J}_h}\Big[\sum_{j\in\tilde{\mathcal{J}}_h} \chi_{hj}(c^o)a_{hj} + c^o\Big(\sum_{j\notin\tilde{\mathcal{J}}_h} \Lambda_{hj}\mu_{hj}^L a_{hj} - B_h\Big)\Big]. \tag{39}$$

   In particular, under full slack *(*$\tilde{\mathcal{J}}_h = \mathcal{J}_h$*),* $E = E_{\mathrm{fs}}(\boldsymbol{a}, B_h; c^o)$ *and* $k_{hj} = \chi_{hj}(c^o)e^E$ *for every* j.

Proof. The bed program is strictly concave with linear constraints and a Slater point ($b_{hj} = \Lambda_{hj}\mu_{hj}^L a_{hj} + (B_h - \underline{B})/J_h$), so the Karush–Kuhn–Tucker conditions are necessary and sufficient, and the solution is unique. If $\nu_{hj} = 0$, (34) gives $b_{hj} = (\chi_{hj}(c^o)a_{hj} - E)/c^o$ with $E := \log(\eta_h/c^o)$, and feasibility $b_{hj} \ge \Lambda_{hj}\mu_{hj}^L a_{hj}$ reads $E \le \delta_{hj}(c^o)a_{hj}$. If $\nu_{hj} > 0$, complementary slackness gives $b_{hj} = \Lambda_{hj}\mu_{hj}^L a_{hj}$ and (34) gives $\nu_{hj} = c^o(e^E - e^{\delta_{hj}(c^o)a_{hj}}) > 0$, i.e., $E > \delta_{hj}(c^o)a_{hj}$. Both cases are summarized by the stated max formula, whose sum over lines must exhaust the capacity that (34) forces to bind, giving (37); substituting the multiplier formulas into (36) yields (38) after collecting $\chi_{hj} - c^o\Lambda_{hj}\mu_{hj}^L = \delta_{hj}$.

Denote $D(E)$ the left of (37), which is continuous and nonincreasing in E, diverges as $E \to -\infty$, equals $\underline{B} < B_h$ for $E \ge \max_j \delta_{hj}(c^o)a_{hj}$, and is strictly decreasing below that threshold (the line attaining the maximum is interior there); hence the clearing equation $D(E) = B_h$ has a unique solution, and since $D(E) = B_h > \underline{B}$ some summand exceeds its floor, proving item 2 of the lemma.

For item 3, fix $(\boldsymbol{a}_0, B_0)$ with $\underline{B}(\boldsymbol{a}_0) < B_0$, write $E_0 := E(\boldsymbol{a}_0, B_0)$, and let $\varepsilon > 0$. By item 2, E_0 lies strictly below the threshold $\max_j \delta_{hj}(c^o)a_{0,hj}$, below which $D(\cdot\,; \boldsymbol{a}_0)$ is strictly decreasing; hence

$D(E_0-\varepsilon;\boldsymbol{a}_0) > B_0$, and $D(E_0+\varepsilon;\boldsymbol{a}_0) < B_0$. Since $D(E;\boldsymbol{a})$ is jointly continuous, both strict inequalities are preserved on a neighborhood N of $(\boldsymbol{a}_0, B_0)$ within $\{\underline{B}(\boldsymbol{a}) < B\}$: $D(E_0-\varepsilon;\boldsymbol{a}) > B > D(E_0+\varepsilon;\boldsymbol{a})$ on N. For any $(\boldsymbol{a}, B) \in N$, monotonicity of $D(\cdot\,;\boldsymbol{a})$ then excludes roots outside $(E_0-\varepsilon, E_0+\varepsilon)$, so $|E(\boldsymbol{a}, B) - E_0| < \varepsilon$, proving continuity.

For item 4, under the stated weak inequalities, each summand of D equals the indicated branch (the branches coincide where an inequality binds with equality), so the clearing equation is linear in E and solves to (39); the expressions for k_{hj} follow from (38). □

Lemma 4 has a consequence worth emphasizing before proceeding: given the admission rates and the total capacity, the entire congestion side of the problem—the bed allocation, the clearing price E, the multipliers, and hence every k_{hj}—is a known functional of observed objects and of the overflow rate c^o alone. No component of $\boldsymbol{\theta}$ enters it, and no bed-level data are required to evaluate it. Now consider two parameter vectors $(\boldsymbol{\theta}^*, c^{o*}), (\boldsymbol{\theta}', c^{o\prime}) \in \Theta_h$ that generate the same admission-rate response at a state where $\boldsymbol{a}^*(s)$ is interior and $\underline{B}(\boldsymbol{a}^*(s)) < B_h(s)$. Each parameter's bed allocation must solve the bed program of Lemma 4 at the common $\boldsymbol{a}^*(s)$ under its own rate, so subtracting the two versions of (35) at the common observed $(\bar{p}_{hj}, \boldsymbol{a}^*)$ eliminates the price:

$$\boldsymbol{x}_{hj}\left(a^*_{hj}\right)'\Delta_j = k_{hj}\left(\boldsymbol{a}^*, B_h; c^{o*}\right) - k_{hj}\left(\boldsymbol{a}^*, B_h; c^{o\prime}\right), \qquad \Delta_j := \boldsymbol{\theta}'_{hj} - \boldsymbol{\theta}^*_{hj}, \tag{R}$$

where each side is evaluated at the same observed data.[44]

Finally, we require one additional technical result to establish that, given the assumptions we have established so far, the observed set of acceptance rates would form an open set.

Proposition 5 (Open acceptance set)**.** *Suppose Assumptions 3.1 to 3.4 and 3.6 hold. Then the response map $\bar{\boldsymbol{p}} \mapsto \boldsymbol{a}^*(\bar{\boldsymbol{p}})$ is C^1 on U with everywhere nonsingular derivative*

$$\partial \boldsymbol{a}^* / \partial \bar{\boldsymbol{p}} = \left(\operatorname{diag}\left(2c^{V*}_{hj}(\Lambda^0_{hj})^2\right) + \tfrac{1}{J_h}\,\boldsymbol{\chi} e^E \boldsymbol{\chi}'\right)^{-1}$$

, where $\boldsymbol{\chi} = \left(\chi_{hi}(c^{o*})\right)_i$. Consequently $\mathcal{W} := \{\boldsymbol{a}^*(s) : s \in O\}$ is open in $\mathbb{R}^{J_h}$.

Proof. The feasible set of (33) is compact and does not depend on $\bar{\boldsymbol{p}}$, and the objective is jointly continuous, so by strict concavity and Berge's maximum theorem the optimizer is single-valued and upper hemicontinuous, hence a continuous function of $\bar{\boldsymbol{p}}$ on U (Aliprantis and Border, 2006, Theorems 17.31 and 17.6). By Assumption 3.6(b), for every $\bar{\boldsymbol{p}} \in U$, the optimum is interior and full

[44] Stationarity at an interior optimum requires only differentiability, so (R) does not presume anything about the rival beyond membership in Θ_h. The requirement $\underline{B}(\boldsymbol{a}^*) < B_h$ ensures, via Lemma 4(i)–(ii), that the multipliers—and hence k_{hj}—are uniquely determined by the data and the rate; if every line constraint held with equality, the capacity multiplier would not be point-identified from the bed program and (R) could fail.

slack, so it solves the smooth system $F(\boldsymbol{a};\bar{\boldsymbol{p}}) = \mathbf{0}$ with

$$F_j(\boldsymbol{a};\bar{\boldsymbol{p}}) = \bar{p}_j - \boldsymbol{x}_{hj}(a_j)'\boldsymbol{\theta}^*_{hj} - \chi_{hj}(c^{o*})\,\exp\Big(E_{\mathrm{fs}}(\boldsymbol{a}, B^0_h; c^{o*})\Big).$$

Its Jacobian in $\boldsymbol{a}$ is $-\Big(\mathrm{diag}\Big(2c^{V*}_{hj}(\Lambda^0_{hj})^2\Big) + \frac{1}{J_h}\boldsymbol{\chi}e^E\boldsymbol{\chi}'\Big)$, a negative of a positive diagonal plus a nonnegative rank-one matrix. Denote $\rho_j = 2c^{V*}_{hj}(\Lambda^0_{hj})^2$; then the absolute value of the determinant of the Jacobian is $\Big(\prod_j \rho_j\Big)\Big(1 + \frac{e^E}{J_h}\sum_i \chi_{hi}(c^{o*})^2/\rho_i\Big) > 0$ by the matrix determinant lemma (Horn and Johnson, 2013, §0.8.5). Fix $\bar{\boldsymbol{p}}_0 \in U$. The implicit function theorem yields neighborhoods $W_0 \ni \boldsymbol{a}^*(\bar{\boldsymbol{p}}_0)$ and $U_0 \ni \bar{\boldsymbol{p}}_0$ and a C^1 map $\phi : U_0 \to W_0$ that is the unique root of $F(\cdot\,;\bar{\boldsymbol{p}})$ in W_0; since $\boldsymbol{a}^*$ is continuous and solves $F = \mathbf{0}$ for every $\bar{\boldsymbol{p}} \in U$, it coincides with ϕ near $\bar{\boldsymbol{p}}_0$ and inherits the stated derivative. A C^1 map with nonsingular derivative is open by the inverse function theorem, so $\mathcal{W} = \boldsymbol{a}^*(U)$ is open. □

We are now in a position to prove identification. The identification of the overflow rate rests on the following classical independence property, which we state in the form we need. Its content is standard; we include a short proof for completeness. It invokes no model assumptions.

Lemma 5 (Exponential–polynomial independence)**.** *Let $r_1, \dots, r_m \in \mathbb{R}^d$ be pairwise distinct and $q_1, \dots, q_m$ real polynomials. If $\sum_{k=1}^m q_k(w)\, e^{\langle r_k, w\rangle} = 0$ for all w in a nonempty open set $V \subseteq \mathbb{R}^d$, then $q_k \equiv 0$ for every k.*

Proof. Step 1: reduction to a line segment with distinct projected frequencies. The hypothesis is an identity on an open subset of $\mathbb{R}^d$; to bring one-dimensional methods to bear, we restrict it to a segment, choosing the direction so that the frequencies remain distinct after projection. For $k \neq l$ the vector $r_k - r_l$ is nonzero, so $\{v \in \mathbb{R}^d : \langle r_k - r_l, v\rangle = 0\}$ is a proper hyperplane, and a finite union of proper hyperplanes cannot cover $\mathbb{R}^d$. Any v outside their union therefore satisfies $\langle r_k, v\rangle \neq \langle r_l, v\rangle$ for all $k \neq l$; write $\mu_k := \langle r_k, v\rangle$, pairwise distinct. Fix $w \in V$ and $\varepsilon > 0$ with $w + tv \in V$ for $|t| < \varepsilon$. Evaluating the hypothesis at $w + tv$ and using $e^{\langle r_k, w+tv\rangle} = e^{\langle r_k, w\rangle}e^{\mu_k t}$, we obtain, with $\tilde{q}_k(t) := q_k(w + tv)\, e^{\langle r_k, w\rangle}$ a univariate polynomial of degree $d_k := \deg \tilde{q}_k \leq \deg q_k$,

$$y(t) := \sum_{k=1}^m \tilde{q}_k(t)\, e^{\mu_k t} = 0 \qquad \text{for all } |t| < \varepsilon. \tag{40}$$

Step 2: y solves a linear constant-coefficient equation on all of $\mathbb{R}$. Consider the order-N operator $L := \prod_{k=1}^m (d/dt - \mu_k)^{1+d_k}$, $N := \sum_k (1 + d_k)$. For any smooth f and any μ, the product rule gives $(d/dt - \mu)\Big(f(t)e^{\mu t}\Big) = f'(t)e^{\mu t}$: the factor $(d/dt - \mu)$ passes through $e^{\mu t}$ and differentiates the prefactor. Iterating, $(d/dt - \mu_k)^{1+d_k}\Big(t^j e^{\mu_k t}\Big) = \Big(t^j\Big)^{(1+d_k)} e^{\mu_k t} = 0$ for $0 \leq j \leq d_k$, since the $(1 + d_k)$-th derivative of a polynomial of degree at most d_k vanishes. Because the factors of L are polynomials in the single operator d/dt, they commute, so applying the k-th factor first shows $L\Big(t^j e^{\mu_k t}\Big) = 0$ for every k and

$0 \le j \le d_k$. By linearity, $Ly = 0$—and this holds on all of $\mathbb{R}$, since y is an entire function of t, even though (40) asserts $y = 0$ only on $(-\varepsilon, \varepsilon)$.

Step 3: vanishing on an interval upgrades to vanishing on $\mathbb{R}$. Because $Ly = 0$ is a linear homogeneous equation of order N with constant coefficients, its initial value problem admits a unique global solution (Coddington and Levinson, 1955, Ch. 3). By (40), y vanishes identically near $t = 0$, so $y(0) = y'(0) = \cdots = y^{(N-1)}(0) = 0$; the zero function solves the same initial value problem, and uniqueness forces $y \equiv 0$ on $\mathbb{R}$.

Step 4: linear independence forces every coefficient to vanish. Since the μ_k are pairwise distinct, the characteristic polynomial of L is $\prod_k (z - \mu_k)^{1+d_k}$ with roots μ_k of multiplicity $1 + d_k$, and the N functions $\{t^j e^{\mu_k t} : 1 \le k \le m,\ 0 \le j \le d_k\}$ form a fundamental system for $Ly = 0$: a basis of its N-dimensional solution space, in particular linearly independent on $\mathbb{R}$ (Coddington and Levinson, 1955, Ch. 3). Expanding each polynomial as $\tilde{q}_k(t) = \sum_{j=0}^{d_k} c_{kj} t^j$, Step 3 states that $\sum_{k,j} c_{kj}\, t^j e^{\mu_k t}$ is the zero element of that solution space; expanding zero in a basis has only the trivial representation, so $c_{kj} = 0$ for all k, j, i.e., $\tilde{q}_k \equiv 0$ for every k.

Step 5: the base point sweeps out V. Evaluating $\tilde{q}_k \equiv 0$ at $t = 0$ gives $q_k(w)\, e^{\langle r_k, w\rangle} = 0$, hence $q_k(w) = 0$. Since $w \in V$ was arbitrary, each polynomial q_k vanishes on the nonempty open set V, and a polynomial vanishing on a nonempty open set is identically zero. □

Proposition 6 (The overflow rate is identified)**.** *Suppose $J_h \ge 2$ and Assumptions 3.1 to 3.4 and 3.6 hold. If $(\boldsymbol{\theta}', c^{o\prime}) \in \Theta_h$ is observationally equivalent to $(\boldsymbol{\theta}^*, c^{o*})$, then $c^{o\prime} = c^{o*}$.*

Proof. Work on the open set $\mathcal{W}$ of Proposition 5 and write $w = \boldsymbol{a}^* \in \mathcal{W}$. On O, the composites are fixed, so the left side of (R) is a degree-one polynomial in w with constant coefficients that depends only on its own coordinate,

$$q_j(w) := \boldsymbol{x}_{hj}(w_j)'\Delta_j = \gamma_{0j} + \gamma_{1j} w_j, \qquad \gamma_{0j} = \Lambda^0_{hj}\Big(\Delta^S_j \bar{\omega}^0_{hj} + \Delta^F_j + \Delta^V_j\Big) - \Delta S^{u,0}_{hj}\, \Delta^\tau_j, \quad \gamma_{1j} = 2\Delta^V_j (\Lambda^0_{hj})^2,$$

while, by Assumption 3.6(b) and Lemma 4.4, the truth's side of (R) is

$$k_{hj}(w; c^{o*}) = \chi_{hj}(c^{o*})\, e^{-c^{o*} B^0_h / J_h}\, e^{\langle r^*, w\rangle}, \qquad r^* := \Big(\chi_{hj'}(c^{o*})/J_h\Big)_{j' \in \mathcal{J}_h} \in \mathbb{R}^{J_h}_{++}.$$

Since $\underline{B}(w) < B^0_h$ on $\mathcal{W}$, the rival's clearing price $E'(w) := E(w, B^0_h; c^{o\prime})$ is well defined and continuous there by Lemma 4.1 and 4.3.

Step 1: the rival's regime is constant on a nonempty open subset. For each nonempty $\tilde{\mathcal{J}}_h \subseteq \mathcal{J}_h$ define

$$G_{\tilde{\mathcal{J}}_h} := \Big\{ w \in \mathcal{W} :\ E'(w) \le \delta_{hj}(c^{o\prime}) w_j \ \forall j \in \tilde{\mathcal{J}}_h,\ \ E'(w) \ge \delta_{hj}(c^{o\prime}) w_j \ \forall j \notin \tilde{\mathcal{J}}_h \Big\},$$

a relatively closed subset of $\mathcal{W}$. Every $w \in \mathcal{W}$ lies in $G_{\tilde{\mathcal{J}}_h(w)}$, where $\tilde{\mathcal{J}}_h(w) := \{j : E'(w) < \delta_{hj}(c^{o\prime}) w_j\}$ is nonempty by Lemma 4.2. An open subset of $\mathbb{R}^{J_h}$ is a Baire space, so the finite closed cover

$\{G_{\tilde{\mathcal{J}}_h}\}_{\tilde{\mathcal{J}}_h}$ contains a set $G_{\tilde{\mathcal{J}}_h^0}$ with a nonempty open subset V (Munkres, 2000, §48). On V, Lemma 4.4 applies to the rival with slack set $\tilde{\mathcal{J}}_h^0$, of cardinality $\tilde{J}_h^0$.

Step 2: $\tilde{\mathcal{J}}_h^0 = \mathcal{J}_h$. Suppose some line $j \notin \tilde{\mathcal{J}}_h^0$. By Lemma 4.4, on V the line-j restriction (R) reads

$$\chi_{hj}(c^{o*})e^{-c^{o*}B_h^0/J_h}\, e^{\langle r^*, w\rangle} - c^{o\prime}\Lambda_{hj}\mu_{hj}^L e^{-c^{o\prime}B_h^0/\tilde{J}_h^0}\, e^{\langle r^0, w\rangle} - \delta_{hj}(c^{o\prime})\, e^{\langle d_j, w\rangle} - q_j(w)\, e^{\langle 0, w\rangle} = 0,$$

where $r^0 \in \mathbb{R}_{++}^{J_h}$ has entries $\chi_{hj'}(c^{o\prime})/\tilde{J}_h^0$ for $j' \in \tilde{\mathcal{J}}_h^0$ and $c^{o\prime}\Lambda_{hj'}\mu_{hj'}^L/\tilde{J}_h^0$ for $j' \notin \tilde{\mathcal{J}}_h^0$, and $d_j := \delta_{hj}(c^{o\prime})\boldsymbol{e}_j$, with $\boldsymbol{e}_j$ the j-th standard basis vector. Since $J_h \geq 2$, there is a line $j' \neq j$ at whose coordinate d_j vanishes while r^* and r^0 are strictly positive, and $d_j \neq 0$ by Lemma 3; hence $d_j \notin \{r^*, r^0, 0\}$. Grouping any coincident frequencies among $\{r^*, r^0, 0\}$ and applying Lemma 5 on V, every grouped coefficient must vanish; in particular the coefficient of $e^{\langle d_j, w\rangle}$, the nonzero constant $\delta_{hj}(c^{o\prime})$, must vanish—a contradiction. Hence $\tilde{\mathcal{J}}_h^0 = \mathcal{J}_h$: the rival is full slack on V, and for every line j and $w \in V$, $k_{hj}(w; c^{o\prime}) = \chi_{hj}(c^{o\prime})e^{-c^{o\prime}B_h^0/J_h}e^{\langle r', w\rangle}$, with $r' := \left(\chi_{hj'}(c^{o\prime})/J_h\right)_{j' \in \mathcal{J}_h}$.

Step 3: $c^{o\prime} = c^{o*}$. Suppose not. Each $\chi_{hj}(\cdot)$ is strictly increasing (Lemma 3), so $r' \neq r^*$, and both differ from 0. Fixing any line j, restriction (R) on V is then a vanishing combination of exponentials with three pairwise-distinct frequencies and polynomial coefficients; by Lemma 5 the coefficient of $e^{\langle r^*, w\rangle}$, namely $\chi_{hj}(c^{o*})e^{-c^{o*}B_h^0/J_h} > 0$, must vanish—a contradiction. Hence $c^{o\prime} = c^{o*}$. □

Proposition 7 (The linear parameters are identified)**.** *Suppose Assumptions 3.1 to 3.5 hold and* $(\boldsymbol{\theta}', c^{o\prime}) \in \Theta_h$ *is observationally equivalent to* $(\boldsymbol{\theta}^*, c^{o*})$ *with* $c^{o\prime} = c^{o*}$. *Then* $\Delta_j = 0$ *for every* $j \in \mathcal{J}_h$, *and hence* $\boldsymbol{\theta}' = \boldsymbol{\theta}^*$.

Proof. At any regular state $s \in \mathcal{R}_h$, Lemma 4 determines the bed allocation, the multipliers, and hence every k_{hj} as functionals of the observed data and of the rate alone. With $c^{o\prime} = c^{o*}$ the two parameter vectors therefore share k_{hj} at every regular state, and (R) reduces to $\boldsymbol{x}_{hj}(a_{hj}^*(s))'\Delta_j = 0$. Dividing by $\Lambda_{hj}(s) > 0$,

$$\Delta_j^S\, \bar{\omega}_{hj}(s) + \left(\Delta_j^F + \Delta_j^V\right) + 2\Delta_j^V\, \Lambda_{hj}(s)\, a_{hj}^*(s) - \Delta_j^\tau\, \frac{\Delta S_{hj}^u(s)}{\Lambda_{hj}(s)} = 0 \qquad \text{for all } s \in \mathcal{R}_h,$$

an affine function of $(x, y, z) = \left(\bar{\omega}_{hj},\ \Lambda_{hj}a_{hj}^*,\ \Delta S_{hj}^u/\Lambda_{hj}\right) \in \mathbb{R}^3$ vanishing on the realized set of triples. By Assumption 3.5, that set contains four affinely independent points, so the affine function is identically zero: $\Delta_j^S = 0$, $\Delta_j^V = 0$, $\Delta_j^\tau = 0$, and then $\Delta_j^F = 0$. □

Theorem 1 (Global identification)**.** *Suppose* $J_h \geq 2$ *and Assumption 3 holds. Then* $(\boldsymbol{\theta}^*, c^{o*})$ *is globally identified on* Θ_h *in the sense of Definition 1.*

Proof. Immediate from Propositions 6 and 7. □

We conclude with a few remarks regarding the proof and its structure. First, we note that

the identifying restrictions do not rely on the line-level allocations b_{hj} and the multipliers (η_h, ν_{hj}), which appear only as internal optimality objects: at any regular state, Lemma 4.1 shows they are the unique solution of the bed program given $(\boldsymbol{a}^*, B_h)$ and a candidate rate, so the marginal capacity cost $k_{hj}(\boldsymbol{a}^*, B_h; c^o)$ is computable without observing $\boldsymbol{b}_h$. However, we note two caveats to this approach. First, it relies on the regularity condition $\underline{B}(\boldsymbol{a}^*) < B_h$: at the knife edge $\underline{B}(\boldsymbol{a}^*) = B_h$, the multipliers are not point-identified from the bed program, which is why Definition 1 and Assumption 3.5 restrict attention to regular states. Second, it relies on the common overflow rate of Assumption 3.2, which collapses the internal capacity market into the single index $E(\boldsymbol{a}, B_h; c^o)$; with line-specific rates, recovering the implied allocation from $(\boldsymbol{a}, B_h)$ alone would be substantially more challenging.

The second thing to note about the result is that the full-slack requirement in Assumption 3.6(b) is expositionally convenient but not essential. What matters is that the truth's regime be constant on some open subset of $\mathcal{W}$. If the truth's constant regime contains a binding line j, identification of c^o is immediate: the truth's own-coordinate exponential $\delta_{hj}(c^{o*})e^{\delta_{hj}(c^{o*})w_j}$ can only be matched, by Lemma 5, by a rival term with the same frequency, which requires the rival to bind at j with $\delta_{hj}(c^{o\prime}) = \delta_{hj}(c^{o*})$; strict monotonicity of $\delta_{hj}(\cdot)$ (Lemma 3) then gives $c^{o\prime} = c^{o*}$. If it contains none, the regime is full slack and Steps 2–3 of Proposition 6 apply.

Finally, we note that conditional reimbursement variation on the full-slack region does not, by itself, identify all supply-side primitives. Once $c^{o\prime} = c^{o*}$, restriction (R) on the open set V of Proposition 6 reduces to $q_j \equiv 0$, i.e., $\Delta_j^V = 0$ and $\Lambda_{hj}^0(\Delta_j^S \bar{\omega}_{hj}^0 + \Delta_j^F) - \Delta S_{hj}^{u,0} \Delta_j^\tau = 0$. Therefore, price variation identifies the overflow rate c^o, the congestion cost c_{hj}^V, and the composite net per-admission cost $c_{hj}^S \bar{\omega}_{hj}^0 + c_{hj}^F - \tau_{hj} \Delta S_{hj}^{u,0} / \Lambda_{hj}^0$, but cannot separate the case-mix–scaled cost, the fixed cost, and the charitable preference. Variation in case mix and in charitable-care demand across regular states supplies the missing directions. This formalizes the observation in Section 6 that, absent variation in demand among those seeking charitable care, neither the cost of charitable care nor the hospital's preference for uninsured surplus is identified.

A.4.5 Supply Estimation.

A.5 Counterfactual Appendix

This appendix contains the additional information and evidence on our counterfactual analyses.

A.5.1 Dollarizing Patient Surplus. To present dollar-value patient surplus estimates and contrast the gains of charitable care regulation with government spending, we require a conversion from mile-equivalent patient surplus to dollars. Our demand estimates allow us to compute the compensating variation in miles that would leave patients indifferent between status quo and being allocated to different hospitals in counterfactual exercises. These estimates capture patient

heterogeneity in patient travel and — by using travel distance as the numeraire — allow comparison across patients. However, dollarizing these values requires an estimate of how patients trade off expenditure and travel.

Several articles estimate hospital-choice or referral models that allow estimating patients' dollar-valued disutility from travel. These estimates, however, share common empirical challenges relevant to the hospital setting. In particular, hospital patients rarely know which prices they will face when seeking care because hospital prices are opaque, mediated by coinsurance schedules, and are associated with an uncertain bundle of services related to each medical condition. For example, a patient choosing where to have a delivery needs to not only form expectations about the degree of complications she or the baby might have, but also about the price she will pay for each of these services, which requires knowledge of hospital prices and her position in her coinsurance schedule. In addition, patient choices are often mediated by referring physicians or insurers, mixing their preferences with those of the intermediaries.

Given these challenges, Prager (2020) stands out as the most appropriate for our setting. This article studies inpatient non-emergency admissions in Massachusetts during 2009 and 2012 in a setting where patients have fixed, hospital-tier-specific out-of-pocket copays. The copays are clearly stated, predictable, and observable before admission, allowing patients to forecast their out-of-pocket payment accurately. Using Prager's preferred specification and adjusting the estimates to our 2023-based dollars using CPI measures, we obtain a range of $1,033 to $1,823 per mile. While the article's estimates include substantial heterogeneity across patient types, we rely on the lower of these estimates as a conservative measure of patients' disutility from travel. We caveat, however, that these estimates might contain the effects of physician referrals.

We benchmark our estimates to a range of alternative conversions obtained from Ho and Pakes (2014) and Gowrisankaran et al. (2015). The former studies privately insured maternity admissions from California during 2003. They use a revealed-preference moment-inequality estimator to recover a measure of travel disutility in dollars. Their estimator is designed to accommodate measurement error in expected hospital prices and flexible hospital-by-severity quality differences, but they rely on the insurer price rather than the consumer price in their estimation. Converting from their 2003 dollars to 2023 dollars, the estimates imply an insurer-payment equivalent of approximately $110 to $571 per additional mile. However, the authors explicitly characterize the estimates as summarizing the preferences generated by this combined decision-making process rather than separately identifying patient preferences.

Gowrisankaran et al. (2015) use hospital-discharge and claims data from Northern Virginia covering 2003–2006. They construct patient-specific expected coinsurance rates from observed claims information and estimate a model in which utility depends on travel time and each patient's expected out-of-pocket payments. The paper reports that increasing travel time to all hospitals by one minute reduces estimated consumer surplus by approximately $167.28 in the nominal dollars

of the pooled sample period. Using our average-sample-period CPI adjustment, this corresponds to approximately $265 per additional travel minute in 2023 dollars. Converting this estimate to miles requires an external assumption about average driving speed: at 30 miles per hour, it corresponds to approximately $530 per mile, while assumed speeds of 20–40 miles per hour imply approximately $397–$795 per mile.

Given the clear observability of prices, the general inpatient setting, and the recency of the data, we consider the estimates Prager (2020) as the most appropriate for our setting. Nevertheless, we present the robustness of our findings to the full range of estimates in the main text.

A.5.2 Pay-per-Admit Policy.

A.5.3 Floor-and-Trade Proposition. Here we present a short proof of Proposition 2. First, define

$$Q_h^u(\boldsymbol{a}) := \sum_{j\in\mathcal{J}} \sum_{k\in\mathcal{K}_j^c} a_{hj}\lambda_k R_{hk}(\boldsymbol{a})\omega_k, \qquad Q^u(\boldsymbol{a}) := \sum_{h\in\mathcal{H}^{np}} Q_h^u(\boldsymbol{a}).$$

Market clearing imposes $\sum_{h\in\mathcal{H}^{np}} u_h = Q^u(\boldsymbol{a})$. Denote the left-hand side of the inequality (21) for hospital h by $\Delta\Pi_h(z^p)$. Note that given any positive market clearing price z^p, we have $\Delta\Pi_h(z^p) > 0$ since it is the difference between the hospital's optimal profit when per-unit uninsured revenue is higher but without any obligations, relative to the optimal revenue under non-profit status after (positive) taxes and no revenue increase on the uninsured.

With the above notations, the inequality (21) implies that $\sum_{h\in\mathcal{H}^{np}} \Delta\Pi_h(z^p) \geq z^p Q^u(\boldsymbol{a})$. If letting $u_h = \frac{\Delta\Pi_h}{\sum_{h'}\Delta\Pi_{h'}} Q^u(\boldsymbol{a})$, then we have trivially that market clearing holds and that

$$\Delta\Pi_h \geq z^p u_h \iff \Delta\Pi_h \geq z^p \frac{\Delta\Pi_h}{\sum_{h'}\Delta\Pi_{h'}} Q^u(\boldsymbol{a}) \iff \sum_{h'} \Delta\Pi_{h'} \geq z^p Q^u(\boldsymbol{a})$$

Hence there exists a market-clearing price and an obligations vector that solves the floor-and-trade problem if and only if the aggregate condition holds.

A.5.4 Policies with Public Responses. This section briefly discusses the key results when allowing public hospitals to respond to the counterfactual policies. The main results are shown in Table A.13 which can be compared to main text estimates from Table 4. The broad patterns of the main text evidence are replicated in this sample. First, providing payment to hospitals that care for the uninsured benefits patients on average, with only a fraction of the value falling on patients receiving charitable care. Second, as in the main text, a policymaker who focuses solely on the surplus value of the pseudo-insurance policy on the uninsured would not implement any of the listed policies because the surplus to the uninsured is an order of magnitude lower than the public cost. Third, paying for otherwise charitable care increases available capacity in

Table A.13: Counterfactual Policy Welfare Effects With Public Hospital Participation

	Regulatory	Δ **Welfare** ($100M)				Δ **Available Capacity** (1k bed-days)			
	Price	Patients		Hospitals	Gov.	Emergency	Scheduled		Idle
	z	All	Charity				Insured	Uninsured	(p.p.)
Panel A: Pay-per-Complexity									
Medicaid-equivalent	1.676	2.302	0.495	1.392	1.543	99.059	-9.260	116.264	-5.070
Medicare-equivalent	3.759	3.071	0.745	3.205	3.475	157.479	-18.545	188.198	-8.048
Optimum at $\gamma^{\pi} = 0$	1.104	1.882	0.357	0.914	1.014	78.725	-5.792	60.048	-3.272
Optimum at $\gamma^{\pi} = 1.00$	4.165	3.300	0.782	3.543	3.851	164.859	-11.967	193.840	-8.530
Panel B: Floor-and-Trade									
Unconstrained	12.108	2.563	0.407	-0.199	0.000	82.915	-5.661	117.434	-4.790
Homogeneous	0.241	0.359	0.110	-0.014	0.000	13.158	-1.916	24.947	-0.890

Notes: Welfare is measured as the change relative to the baseline, in hundreds of millions. The Patients columns report the gains for all patients and charitable-care patients, respectively, with "Gov." representing the change in fiscal cost. Available capacity is the sum across hospitals of the product of each line's allocated bed-days and its acceptance rate, measured in thousands. Idle capacity ratio is defined as $(\sum_h B_h - \sum_{h,j} a_{hj} b_{hj}) / \sum_h B_h$ and the change is presented in percentage points. The first column reports the regulatory payment or credit price in $10,000 units. All values are quarterly.

the emergency room, decreases capacity for scheduled care for the uninsured, and increases it for scheduled care for the uninsured, overall decreasing idle capacity. Fourth, an unconstrained floor-and-trade policy reaches consumer surplus near Medicaid-equivalent rates at zero public cost, while a homogeneous policy achieves only a small fraction of those values.

The results also highlight some differences associated with public hospital participation. First, because public hospitals are more responsive to uninsured demand, patient-optimal pseudo-insurance pays slightly less than Medicaid rates. Second, because more hospitals participate in public payment programs, running Medicaid- or Medicare-equivalent pseudo-insurance systems is slightly more expensive in this scenario. Third, because of how public participation affects private hospitals' incentives to provide care to the uninsured, patient surplus gains from pseudo-insurance policies are somewhat lower than in the main text result. Finally, because public hospitals in this scenario respond but do not participate in the floor-and-trade scheme, the patient surplus gain is slightly larger, with more of the gains falling on the insured.

A.5.5 Implementation.

References

Charalambos D. Aliprantis and Kim C. Border. *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer, Berlin, third edition, 2006. doi: 10.1007/3-540-29587-9.

Earl A. Coddington and Norman Levinson. *Theory of Ordinary Differential Equations*. McGraw–Hill,

New York, 1955.

William Feller. *An Introduction to Probability Theory and Its Applications*, volume 1. John Wiley & Sons, New York, third edition, 1968.

Gautam Gowrisankaran, Aviv Nevo, and Robert Town. Mergers when prices are negotiated: Evidence from the hospital industry. *The American Economic Review*, 105(1):172–203, 2015. ISSN 0002-8282.

Kate Ho and Ariel Pakes. Hospital choices, hospital prices, and financial incentives to physicians. *American Economic Review*, 104(12):3841–3884, 2014.

Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, second edition, 2013.

James R. Munkres. *Topology*. Prentice Hall, Upper Saddle River, NJ, second edition, 2000.

Elena Prager. Healthcare demand under simple prices: Evidence from tiered hospital networks. *American Economic Journal: Applied Economics*, 12(4):196–223, October 2020. doi: 10.1257/app.20180422. URL https://www.aeaweb.org/articles?id=10.1257/app.20180422.

Coady Wing, Seth M Freedman, and Alex Hollingsworth. Stacked difference-in-differences. Working Paper 32054, National Bureau of Economic Research, January 2024. URL http://www.nber.org/papers/w32054.