Going for Broker?
Intermediation in Health Insurance Markets*

Anran Li† Tong Liu‡ Anthony LoSasso§ Nicholas Tilipman¶

October 30, 2026

Preliminary – Please Do Not Circulate.
Click here for the latest version.

Abstract

This paper studies how regulating intermediary compensation involves a tradeoff between agency distortions and competitive access. We compile a novel database of contracting relationships among employers, brokers, and insurers in New York State employer-sponsored insurance markets. Exploiting variations in commission schedules, we document two market distortions: First, brokers exhibit agency frictions, steering employers towards more financially lucrative products. Second, lower commission levels shrink broker-insurers' contracting networks and, in turn, soften insurers' competitive pressure for employer demand. We develop and estimate an equilibrium model of employer demand, insurer pricing, and broker network formation. Commission caps reduce steering but can narrow broker networks and weaken upstream competition, making welfare effects non-monotone.

JEL Classification Codes: C78, I11, L13

*We thank Ben Handel, Igal Hendel, Gaston Illanes, Jonanthan Kolstad, Mark Shepard, and seminar participants at various conferences and institutes for their constructive comments. We thank Mark Egan, Atul Gupta, Christoph Herpfer, Adam Jorring, Charles Murry, Cici McNamara, Yanfei Wang, Wesley Yin for their insightful discussions. Liu gratefully acknowledges funding support from MIT Sloan Health Systems Initiative (HSI) and MIT Sloan Junior Faculty Research Assistance Program. Tilipman acknowledges funding from the Johns Hopkins Business of Health Initiative Pilot grant and Carey Business School General Research Support Fund.

†Li: Department of Economics, University of Minnesota Twin Cities and NBER, anranli@umn.edu.

‡Liu: Sloan School of Management, MIT, tongl@mit.edu.

§LoSasso: La Follette School of Public Affairs, University of Wisconsin-Madison, tony.losasso@wisc.edu.

¶Tilipman: Carey School of Business, Johns Hopkins University, tilipman@jhu.edu.

1. Introduction

Intermediaries play a central role in many markets, including finance, housing, cars, and health insurance, serving as skilled experts who assist in consumer decision-making. While intermediaries can reduce consumer search costs, policymakers have increasingly scrutinized their role due to concerns about conflicts of interest and commission-based steering that may distort market outcomes. In several insurance markets where intermediaries interact directly with consumers, state and federal policymakers considered or enacted caps on broker compensation as a means of curbing these agency problems.[1] Yet intermediaries who contract selectively with upstream suppliers do more than advise consumers; they determine which products compete for consumers' business in the first place. When the same commissions that bias advice also sustain these contracting relationships, regulating compensation confronts a tension that curbing steering alone cannot resolve.

This paper examines the welfare effects of health insurance brokers in the employer-sponsored insurance (ESI) market. The ESI market covers over 175 million Americans, nearly half the US population, and accounts for approximately $1.2 trillion in annual health care spending, reflecting roughly 30% of national health expenditures (CMS, 2024). Over the last decade, per-member cost growth for ESI has reached 80%, outpacing that of Medicare and Medicaid (KFF, 2025). A central question is how much of this cost growth reflects frictions in how insurance options are intermediated and chosen by firms, rather than underlying medical costs.

ESI brokers play unique roles. First, they advise employers on which insurer to choose, working with employers' human resources departments rather than employees directly, unlike intermediaries in Medicare Advantage or ACA exchanges. Second, they contract selectively with insurers, carrying only a subset of available carriers in each local market and thereby determining which insurers compete for a given employer's business. Beyond shaping recommendations within a choice set, brokers shape the choice set itself. Market outcomes, therefore, depend not only on consumer preferences but also on the incentives and contracting behavior of employers, brokers, and insurers.

Our contribution is to provide a systematic empirical account of broker intermediation in employer-sponsored insurance and show that broker compensation creates a steering-competition tradeoff. The steering side is the agency friction emphasized in prior work: commission incentives lead brokers to recommend financially lucrative products over higher-surplus ones. We further identify the second, less understood channel: broker compensation shapes the structure of competition itself, a force we term *competitive-access distortion*. Even absent any steering, commission-driven contracting fragments choice sets and softens upstream competition, raising equilibrium premiums. These two distortions pose a fundamental challenge for regulators. Commission caps, a common policy response to conflicted advice, might curb steering distortions, but can also narrow broker-insurer networks and weaken the competition those networks support, a tension we quantify with a structural model.

[1] The Centers for Medicare and Medicaid Services (CMS) recently overhauled agent and broker compensation in Medicare Advantage and Medicare Part D (Sidley Austin LLP, 2024; Commonwealth Funds, 2025), and New York banned contingent commissions in 2005, a policy later reversed (Reuters, 2008). Others instead target conflicts of interest through disclosure or fiduciary duty requirements (New York Times, 2020).

We begin with a stylized model to illustrate how steering and competitive-access distortions arise and interact. The key conceptual insight is that they operate in tension: higher commissions paid by one insurer relative to another exacerbate steering, increasing brokers' incentives to direct employers toward financially lucrative rather than welfare-maximizing products. However, those same commissions encourage brokers to contract with more insurers to capture them, broadening menus, intensifying upstream competition, and lowering prices. Welfare-optimal regulation, therefore, cannot focus solely on eliminating steering, as doing so may inadvertently soften competition; the net effect depends on how brokers' steering and contracting responses interact with insurers' pricing strategies in equilibrium. The framework yields three empirical implications: commissions affect broker recommendations, commissions affect broker-insurer contracting, and contracting breadth affects premiums.

We test these implications using the first comprehensive dataset on the insurer-broker-employer relationship in the New York State ESI market. We focus on small-to-medium-sized fully-insured employers, and document their relations with ten major insurers and over seven hundred brokers. The data combine multiple sources: employer-level Form 5500 filings from the Department of Labor with detailed information on insurers, brokers, premiums, and enrollees; commission schedules obtained through Freedom of Information Act (FOIA) requests, which identify brokers' financial incentives; and records on broker–insurer contracting networks manually scraped from government portals, including which brokers are licensed to sell for which insurers and the timing of these relationships. We supplement these data with annual enrollment records from a large national insurer and claims data from the Clarivate database.

We proceed to test for each empirical marginal in turn. To identify steering, we exploit sharp discontinuities in broker commission schedules using a regression kink design. Critical to this design is the fact that commission schedules vary within-insurer by the number of enrolled employees, mitigating concerns of unobserved insurer quality correlated with commissions; We find no evidence of manipulation of the running variable around the kink cutoff. However, discrete reductions in marginal commission rates lead to systematic shifts in the choice of insurer identity. Differencing brokered against non-brokered employers in a difference-in-kinks design, we find that each additional ten enrollees beyond the kink cutoff reduces the probability that a brokered employer chooses the insurer by about 5.7 percentage points, with no corresponding kink for non-brokered firms. Average premiums fall, suggesting that higher compensation induces brokers to steer employers toward more expensive products..

We next show that commissions shape broker contracting and that contracting, in turn, shapes competition. For the contracting margin, we leverage a quasi-experiment in which one insurer raised its per-member per-month (PMPM) commission from \$5 to \$18 and later reduced it to \$12, while other insurers' commissions remained stable. Using an event study design, we document that a 33% decline in commissions reduces the share of brokers contracting with that insurer by about 5 percentage points (6% relative to its pre-event share), and the increase in commissions has an opposite effect. For the pricing margin, we compare brokers exposed to these commission changes with unexposed brokers: when the insurer's broker network contracts, per-member annual premiums charged by competing insurers rise by about \$912, consistent with narrower broker-mediated choice sets soften insurer competition softening.

These stylized facts establish that all three margins are empirically relevant. Because steering and net-

work breadth respond to commissions in opposite welfare directions, their net effect is ambiguous and requires a model for quantification. We therefore develop and estimate a structural model of demand, pricing, and network formation. The model features four stages. First, given posted commission schedules, brokers select which insurers to include in their product offerings and incur the associated contracting costs. Second, employers quasi-randomly meet with a broker and decide whether to use the brokers' services. This step quantifies how brokers reduce employers' hassle costs of offering insurance plans. Third, insurers simultaneously set product premiums for employers, depending on whether the employer quotes through a broker, and if so, the broker's contracting set. This captures how selective contracting affects insurer competition. Fourth, employers choose insurers, from the broker's product offering if brokered, otherwise from the available set of products on the market. Employers weigh their own premium expenses against their employee's surplus, with the relative weights shifted by commissions. This step measures broker effects on match quality and the degree of steering.

We estimate the model sequentially. On the demand side, we recover employers' preferences using maximum likelihood estimators. By rationalizing differences in employer choice of plan following plausibly exogenous variation in within-insurer commissions, we identify steering, or broker influence on employer product offerings, separately from insurer brand effects. We estimate employers' hassle costs from their decisions to use brokers. We instrument the value of a broker's menu using its value for other clients outside the employer's county, thereby isolating variation in menu value that is less likely to reflect the focal employer's unobserved demand. On the supply side, we recover insurers' marginal costs using first-order conditions consistent with profit maximization. We estimate brokers' contracting costs using moment inequalities that exploit revealed preferences: adding or subtracting an insurer results in profit losses compared to the status quo contracting network.

Model estimates yield several key findings on the role of brokers. First, employers without brokers value \$1 of employee welfare at \$1.33 relative to their own premium expenses. In other words, they pay \$1.33 to generate \$1 in value for an average employee.[2] Second, brokers steer employers towards more generous plans, consistent with our reduced-form results. Brokers increase employers' willingness to pay by \$0.25–\$0.55 for the same \$1 of employee surplus, of which one-fifth is attributable to commission-induced steering. We estimate that a one-percentage-point decrease in commissions reduces the probability of choosing the high-commission insurer by roughly 3 percentage points. On average, brokered employers pay 25% higher premiums than non-broker firms for the same level of employee surplus. Third, employers behave as if brokers provide an additional \$61 per member per month beyond the value of their insurer menus. Under the model, we interpret this residual value as avoided hassle costs. For a typical 200-employee firm in New York City, this amounts to approximately \$171,000 per year, approximately the average salary of hiring two benefits consultant employees. Finally, brokers' selective contracting imposes sizable premium costs, with contracting costs accounting for about 70% of expected commission income.

Using the estimated model, we evaluate proposed regulations, including caps on commissions and prohibitions on exclusive contracts between brokers and insurers, taking into account both steering and competitive-access distortions. We first use partial-equilibrium counterfactuals to isolate the welfare ef-

[2]This is fairly consistent with the magnitude of the tax exclusion on health insurance, which amounts of 30% on average.

fects of broker-induced steering and selective contracting, holding employer-broker matches fixed. Removing commission-driven steering improves plan matching and lowers prices: employer surplus increases by roughly 6%. Changes in employee welfare depend on the alignment btw broker incentives and plan quality: Where high-commission insurers offer high-quality plans, steering directs employers toward them, such that removing steering lowers employee WTP. When insurers are allowed to reoptimize prices in response to the absence of steering, competitive discipline erodes: removing steering reallocates demand away from high-commission insurers, such that share-gaining insurers raise prices. We also examine selective contracting by expanding employers' choice sets to include all insurers active in a market. Broader menus raise employer surplus by more than 10% and modestly increase employee surplus, reflecting increased product variety and stronger price competition.

We then study a policy counterfactual that imposes a one-percentage-point cap on broker commissions, allowing different margins of adjustment to operate sequentially. The cap weakens brokers' incentives to steer employers toward high-commission plans, improving allocative efficiency and initially raising employer surplus by about 3% through better choices and lower insurer markups. However, once brokers are allowed to adjust their menus, the cap reduces the profitability of some insurer-broker relationships, leading brokers to narrow their networks. This contraction amplifies competitive-access forces, raises prices in broker-mediated contracts, and ultimately lowers employer surplus by more than 6%. In equilibrium, the commission cap produces a net decline in employer surplus of roughly 3%, highlighting a trade-off between reducing steering distortions and preserving competitive broker networks.

Our results provide new insight into the mechanisms driving rising health insurance costs in the US and have important policy implications. First, our findings speak directly to the design of intermediary regulation in markets such as Medicare Advantage, where third-party brokers are already subject to extensive federal oversight. Second, they inform antitrust policy by highlighting a vertical friction: selective contracting between insurers and brokers can suppress competition and contribute to higher prices. More broadly, our analysis illustrates how intermediaries, while acting as agents for buyers, can also reshape market structure and competition upstream, amplifying or mitigating welfare losses depending on how their incentives are regulated.

Related Literature. A central theme of the intermediation literature is that commission-based compensation creates agency frictions. This has been documented in car dealers (Li et al., 2025), real estate (Barwick et al., 2017), financial advisory (Bhattacharya et al., 2019), medical devices (Grennan et al., 2024), pension (Boehm, 2023), and mortgage (Robles-Garcia, 2019; Kim, 2024).[3] Notably, despite the policy prominence of broker compensation in health care, there is remarkably little scholarly evidence on how commissions affect broker behavior and market outcomes in health insurance specifically. We contribute on two dimensions. First, we provide a first systematic empirical account of broker intermediation in employer-sponsored insurance. Our novel data collection links upstream commission schedules and broker–insurer contracting relationships to downstream employer choices, premiums, and enrollment, revealing the entire market organization. Second, using distinct quasi-experimental designs, we provide direct evidence on how com-

[3]Most closely related is Robles-Garcia (2019), who studies commission incentives, broker–lender relationships, and upstream competition in mortgage markets. We examine these forces in health insurance using distinct quasi-experimental variation and distinct empirical models that reflect a different institutional environment.

pensation affects both insurer choice and access to competing carriers. [4][5]

We also contribute to the literature on employer-sponsored health insurance. Prior work has documented that consumers face significant frictions when choosing insurance (Handel, 2013; Handel and Kolstad, 2015; Abaluck and Gruber, 2015), and that employer plan design and market structure shape prices and welfare (Bundorf et al., 2012; Shepard, 2022; Tilipman, 2022).[6] Most closely related is Gruber et al. (2021), who study how skilled experts affect consumer choices in markets for complex insurance products; we complement their work by showing that the market-level consequences of intermediation extend beyond choice quality to the structure of upstream competition. More broadly, the ESI literature has largely treated the intermediation layer between insurers and employers as absent. We provide the first evidence on how brokers shape competition and employer welfare in this market.

Outline. The remainder of the paper proceeds as follows. Section 2 describes the institutional background and our data sources. Section 3 presents a conceptual framework that illustrates how brokers influence product assortment and upstream competition. Section 4 presents reduced-form evidence on steering and competitive-access effects. Section 5 develops our structural model. Section 6 discusses estimation and identification, and reports estimation results. Section 7 examines policy counterfactuals, and Section 8 concludes.

2. Empirical Setting

2.1. Institutional Background

The ESI Market. Health insurance brokers serve as intermediaries between employers and insurance carriers in the employer-sponsored insurance (ESI) market. ESI remains the primary source of coverage for working-age adults in the US, with 60% of individuals under age 65, or roughly 165.6 million people, receiving employer-sponsored health insurance in 2026 (KFF, 2026). To provide this coverage, firms either purchase fully insured group policies or self-fund and contract with third-party administrators to manage claims. Because pricing, benefit design, and provider networks are shaped by both local provider competition and federal and state regulations, health insurance brokerage is inherently local.

In New York, the setting of this paper, the ESI market operates under close oversight by the Department of Financial Services, and features a mix of national carriers and regionally dominant insurers operating across downstate and upstate markets. Large employers are more likely to self-insure, while small and mid-sized firms typically purchase fully insured products (see section 2.3). High health care costs and complex regulatory requirements make skilled intermediation particularly valuable, especially for employers without

[4] Karaca-Mandic et al. (2018) find that small firms in more competitive broker markets are more likely to offer insurance and at lower premiums, but do not observe commission schedules or broker-insurer contracting networks. Tong (2022) examines broker advising in health insurance but does not model the competitive effects of broker networks.

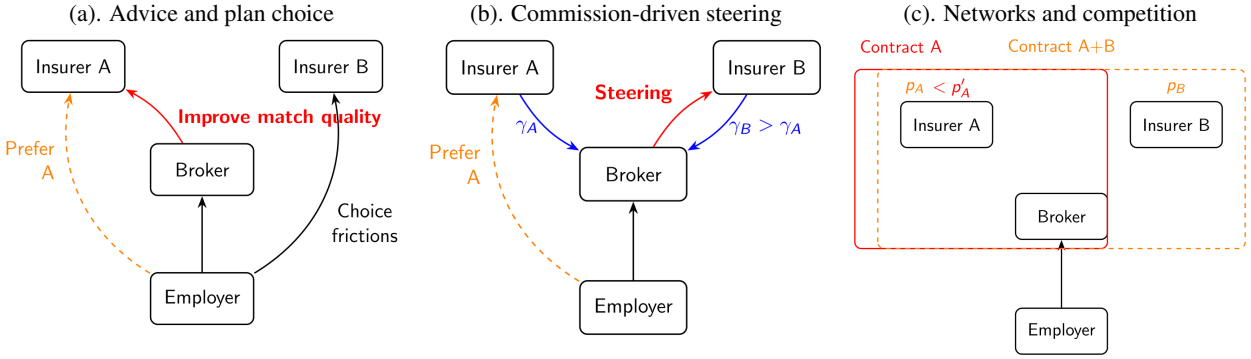
[5] We connect to the literature on vertical relationships and endogenous product assortment (Conlon and Mortimer, 2021; Ho and Lee, 2017; Hristakeva, 2022), but the competitive-access channel arises from contracting incentives rather than physical capacity constraints. Ho and Lee (2017) study endogenous insurer-provider network formation; we study a parallel problem where brokers' contracting decisions with insurers shape the competitive environment facing employers.

[6] See also Craig (2022); Fleitas et al. (2025); Marone and Sabety (2022); Dafny et al. (2013) on insurer competition; Li et al. (2022); Myerson and Li (2022) on information frictions and decision support.

specialized benefits staff[7]. As a result, brokers are often hired to help employers navigate plan selection, pricing, and administration.

Figure 1 summarizes brokers' roles in the ESI market. Brokers reduce employers' administrative burden and help them evaluate insurance options. Their recommendations may also respond to insurer-paid commissions. Their contracting relationships also determine which insurers are available through the broker, shaping competition among insurers for employers' business. We describe the institutional relationships and compensation arrangements underlying these roles below.

Figure 1. The Role of Brokers in the Employer-Sponsored Insurance Market



Notes: This figure illustrates the relationships among an employer, a broker, and two insurers. Panel (a) shows how broker advice can help an employer identify suitable coverage when choice frictions are present. Panel (b) illustrates the incentive to recommend insurer B when it pays a higher commission than insurer A, even if the employer would otherwise prefer A. Panel (c) compares contracting with both insurers with contracting only with A. The premiums p_A and p'_A correspond to these broader and narrower menus, respectively. The price comparison is illustrative.

Broker-Employer Relationships. Although some brokerage chains operate nationally, employer–broker relationships are typically local: employers generally work with brokers in close geographic proximity (see Appendix D3). Employers that retain broker services overwhelmingly use conventional brokers rather than benefit consulting companies or internet-based brokerages (Gates et al., 2010). These relationships are often formed through referrals from current clients or professional partners.[8] These local matching patterns motivate the broker-meeting process later in the model.

Broker-of-record agreements give the broker legal authority to represent the employer with carriers. Once formed, broker–employer relationships tend to persist (see Section 2.3). Changes in ownership or broker retirement can prompt employers to switch[9]. However, broker-of-record agreements are typically terminable on short notice, often on a month-to-month basis, even when compensation is negotiated annually. Thus, persistence reflects relationship value and switching or search frictions rather than binding long-term contracts.

Brokers provide administrative services that reduce the hassle costs employers would otherwise incur to

[7] Liu and Sydnor (2022) documents firms offering dominated options in ESI, consistent with substantial complexity in plan design and evaluation.

[8] Industry trades and marketing materials often coach employers to search for brokers "in their area" through recommendations or trusted advisors, such as accountants, attorneys, or other businesses (HealthPartners, 2021; PeopleKeep, 2013) and emphasize the advantages of local brokers as offering a 'more personalized service' (HealthPartners, 2021)

[9] Industry guidance describes employers' reluctance to replace brokers because of perceived hassle, while explaining that changing brokers need not require changing the underlying employee-benefit program (Marsh McLennan Agency, 2026).

offer coverage. They handle enrollment, regulatory compliance, renewal paperwork, and ongoing benefits management.[10] The administrative role gives brokers value independent of plan choice.

Brokers can also improve match quality through advisory services. They assess firm risk and workforce preferences, collect quotes, present plan options, and recommend products to human resources departments or chief financial officers. Figure 1a illustrates how this advice can help employers identify coverage that better matches their needs. At the same time, the advisory role creates scope for commission-based steering.

Broker-Insurer Relationships. Broker-insurer contracting further shapes the set of products available to employers. To complete a sale for a specific insurer, a broker must be licensed in the relevant state and enter a carrier-specific agreement to sell that insurer's plans.[11] These agreements are generally nonexclusive, and brokers need not contract with every insurer. Brokers may collect quotes without such an agreement, but they cannot complete a sale without a formal contract.

Maintaining these relationships requires recurring time and administrative costs. For example, some insurers require brokers to pass periodic product tests to keep their contracts active. Brokers often choose not to contract with certain carriers when expected client volume is low or when administrative costs are high. This introduces variation in the choice sets available to employers depending on which broker they work with. In practice, brokers tend to favor a limited set of carriers with whom they have long-standing relationships. As one broker put it, "There is an optimal number of insurers to contract with; it's not one or two, but it's also not fifteen."

These contracting relationships can also affect employer premiums. Figure 1c illustrates this channel: a broker contracting with both A and B can offer competing insurance options, whereas a broker contracting only with A offers a narrower menu. Brokers can use competing quotes to negotiate lower premiums. As one broker explained in our interviews: "The majority of what I do is to use competition to create a marketplace. The initial quote serves as a starting point for negotiations, and I use quotes from other insurers to negotiate reductions. The more you work with, the better your prices, but only up to a point." Broader broker networks can therefore strengthen competitive pressure, while the costs of maintaining insurer relationships limit network breadth.

Commissions and Regulations. Brokers are typically compensated by insurers rather than by employers, either through percentage commissions on premiums or through flat fees per enrollee. Percentage-based commissions are most common for serving small to mid-sized fully-insured firms, while fee-based contracts are more prevalent among large self-insured groups, reflecting the greater complexity and scale of services provided.[12] Carriers generally pass these payments through to employers in premiums and have limited

[10] Brokers may also assist with claims resolution, coordinate with carriers on behalf of employers, and, for larger employers, coordinate with pharmacy benefit managers or other specialized vendors. Industry guidance describes dispute resolution and employee claims support as continuing responsibilities of employee-benefits brokers (Marsh McLennan Agency, 2026)

[11] Often, this requires submitting documentation to the carrier, agreeing to their terms and conditions, completing their training, and meeting certain requirements. The process, requirements, paperwork, and terms vary carrier to carrier. Many carriers require brokers to be affiliated with a local Field Marketing Organization to facilitate this (Marketing, 2025).

[12] Two other compensation types are feasible but not applicable in our context: Brokers can receive customized commissions for serving large employers, which are negotiable and often depend on group size or premium volume. In addition, brokerage firms can receive contingent commissions tied to their total volume of new business with a carrier. These payments are largely opaque to employers, episodic and contingent in nature rather than stable components of broker compensation, and usually apply to large employers outside the scope of our sample. Moreover, industry sources suggest that these bonuses are typically small relative to

incentive to reduce commissions, as doing so may discourage broker engagement. For small and mid-sized employers, these commissions are the primary implicit price of broker services.

Standard commission schedules are rigid, observable to brokers, and rarely vary over time (see section 2.2). In New York, insurers file these schedules with the Department of Financial Services, and schedules apply uniformly across brokers for a given insurer. Brokers can therefore compare commission terms before deciding which carriers to contract with or recommend.

Commission differences can influence brokers' insurer recommendations. Figure 1b illustrates this incentive. Holding the broker's menu and insurer quotes fixed, a higher commission from insurer B gives the broker a financial incentive to recommend B even when the employer would otherwise prefer A. Commissions also affect the expected return to maintaining insurer relationships and therefore brokers' incentives to incur the associated contracting costs.

Regulators have long used compensation schemes to address broker agency concerns. Prior to 2010, New York imposed a cap of 4 percent on commissions for health maintenance organizations (HMOs) and the in-network portion of point-of-service (POS) plans (New York State Insurance Department, 2010). The Centers for Medicare and Medicaid Services (CMS) sets explicit maximum payments to agents and brokers for Medicare Advantage (MA) and Part D enrollment under 42 C.F.R. §422.2274 (Ritter Insurance Marketing, 2024).[13] These policies illustrate that broker compensation is a natural regulatory lever. In Section 3, we formalize how commissions shape both brokers' product recommendations and insurer-broker contracting decisions.

2.2. Data

We construct a new employer–broker–insurer panel for New York over 2011–2020. Existing research on ESI has generally lacked linked data on contracting relationships and compensation arrangements underlying employers' insurance purchases. Our data make the intermediary layer observable by jointly recording employer–broker matches, broker–insurer contracts, employers' insurer choices and premiums, and complete insurer commission schedules. To our knowledge, this is the first comprehensive dataset linking these relationships, transactions, and compensation arrangements in ESI.

We combine five primary data sources. The first is the Department of Labor's Form 5500 filings in 2011-2020. Fully insured employers that offer health insurance and have at least 100 enrolled employees must submit these forms, while self-insured plans may choose to submit. These filings report employer identifiers, the brokers with whom they contract, insurance carriers and plan types they select, the number of enrollees and total premiums paid for each carrier-year, and employers' characteristics such as firm size and industry. We focus on fully insured employers in New York, most of which have between 100 and 500 employees, with a small number in the 500–1000 range. These data form the backbone of our analysis as they allow us to link employers, realized brokers, and realized insurers over time. Appendix B1 describes

regular commissions, which remain the primary form of broker compensation.

[13] A CMS final rule in 2024 explicitly noted that "excessive compensation, and other bonus arrangements, offered by plans to agents and brokers can result in individuals being steered tos ome Medicare advantage and Part D plans over others based on the agent or brokers's financial interststs, rather than the prospective enrolee's health care needs." See https://www.cms.gov/newsroom/fact-sheets/contract-year-2025-medicare-advantage-part-d-final-rule-cms-4205-f

our data construction process in detail.

The second source is insurer commission schedules, obtained through a Freedom of Information Act (FOIA) request to New York State (NYS). Under New York Insurance Law Section 4235, insurers must file their commission schedules and rate manuals with the NYS Department of Financial Services to ensure compliance with compensation limits. We collected 76 documents covering all major health insurers in New York over the past two decades and hand-coded their compensation terms. The schedules vary across insurers, employer size, and premium volume. They also differ in form, with some carriers paying percentage commissions and others paying flat per-enrollee fees. These heterogeneities provide the main source of variations in broker financial incentives for our analysis.

The third source is broker-insurer contracting networks. We begin with Discovery Data, a proprietary database that records the set of carriers with which each broker has contracts and is licensed to sell. We supplement these records with hand-collected data from the NYS Department of Financial Services portal, which reports broker identifiers, business types (individual or business entity), addresses, and lines of business (e.g., life, accident, health, or property and casualty), insurer names and identifiers that the broker contracts with, and most importantly, the start and end dates of broker-insurer contracts. Observing contracted menus separately from realized purchases helps distinguish competitive access from insurer choice within those menus.

Finally, we use two datasets to measure employee demand and insurance plan characteristics. One is from a large insurer operating in New York. We observe all employers that choose this insurer, with employer-year level plan offerings, eligible and enrolled employees, plan attributes such as plan type (EPO, PPO, POS), coinsurance, deductible, out-of-pocket maximum, and total premium. We also observe enrollee-year level demographics, plan choice from the employer's offerings, and linked enrollee claims data. Although these data records come from a single large insurer, they complement the Form 5500 data, which contain no within-insurer enrollment records across plans. This lets us recover employee preferences over plan attributes. Appendix B2 reports further details on the dataset and summary statistics.

We also draw on the proprietary Clarivate DRG Real World Data (Liu, 2022; Arnold et al., 2025). The data are sourced from a claims-processing intermediary and are broadly representative across a wide range of commercial insurers. These granular claims records report providers, patients, payers, diagnoses, procedures, and payer and patient payments for deductibles, coinsurance, and copayments. Using these records, we construct plan characteristics, such as deductibles and coinsurance (see Appendix B3 for details). The Clarivate data complements the Form 5500 data, which reports realized insurer choices and premiums but not detailed product attributes. This allows us to accurately characterize preferences for distinct insurers.

2.3. Descriptive Patterns

Employers. Table 1 reports sample statistics for NY employers by broker status. Three patterns emerge. First, fully insured employers, the focus of this paper, account for 74% of employer-year pairs and 81% of total premiums, but only 34% of covered lives. This pattern reflects the fact that small employers are more likely to offer fully insured plans: among such firms, 84% have fewer than 500 enrollees and only 5% exceed 1,000. Second, brokers are widespread in ESI. Among fully insured employers, 79% use a broker. Brokered

employers are smaller on average, with 378 enrolled employees and a 78% take-up rate, compared to 605 employees and a 75% take-up rate for non-brokered firms. Most of these employers offer plans from only one insurer. Third, premiums differ by broker status. Average annual premiums per enrollee are $7,495 for brokered firms versus $6,563 for non-brokered firms. This $932 gap is descriptive, not causal: brokered and non-brokered employers differ in size and other characteristics, and Sections 4-5 isolate broker effects from selection. The gap is nonetheless large relative to average broker payments of $201 per enrollee, suggesting that brokered and non-brokered firms purchase different insurers or products, rather than simply paying an added broker fee.

Table 1. Employer Sample Statistics

	(1)	(2)	(3)	(4)
	All employers	Fully-insured employers		
	All	All	With broker	No broker
Sample share, by empl-yr pairs	100%	77%	60%	16%
Sample share, by enrollee	100%	34%	24%	10%
Sample share, by premium	100%	81%	58%	22%
Employer-year pairs	22,150	16,989	13,368	3,621
Average number of enrollees	963	442	378	605
Average takeup rate	77%	77%	78%	75%
Average number of carriers offered	1.2	1.2	1.2	1.4
Share offering only one carrier	86%	87%	86%	73%
Average premium per enrollee	3,043	7,213	7,495	6,563
Average broker payments per enrollee	57	140	201	-

Notes: This table reports employer-year level summary statistics for New York State employer-sponsored insurance from Form 5500 filings (group health filings and Schedule A) in 2011-2020. Column (1) covers all employers, i.e., both fully and self-insured; columns (2)-(4) restrict to fully insured employers, the focus of our analysis, split by whether the employer uses a broker. Sample shares are computed relative to all employer-years. Take-up is the ratio of enrolled to eligible employees, and carriers offered counts the distinct insurers in an employer's health insurance benefits offering. Premiums and broker payments are annual dollars per enrollee. Premiums are lower for self-insured plans, lowering the column (1) averages.

Insurers. Table 2 reports descriptive statistics for the major insurers in our sample and highlights four patterns. First, the insurer market is relatively concentrated: the top four carriers account for 73.3% market share. As a result, our structural analysis focuses on the ten largest insurers. By contrast, the broker market is far more decentralized, with more than 700 brokers serving in the NY ESI markets. Second, insurer choices differ systematically between employers that work with brokers and those that do not. While the overall distribution across insurers is broadly similar, the national insurer captures a larger market share among brokered firms, whereas lower-priced regional insurers account for a greater share among non-brokered firms. This may reflect brokers' ability to match employers with higher-quality insurers, but it may also indicate that brokers steer employers to higher-priced insurers. Third, insurers differ markedly in both pricing and contracting strategies. Average annual per-member premiums range from $5,268 to $9,412, and broker networks differ widely: some insurers contract with more than 500 brokers, while others engage fewer than a hundred. Notably, the largest insurers, such as Oxford, tend to maintain broader broker networks and charge higher premiums. Finally, contracted networks are broader than realized placements. For example, of Oxford's 551 contracted brokers, only 319 ultimately placed clients into an Oxford plan in our sample.

Table 2. Insurer Sample Statistics

	(1) Market Share	(2) Market Share	(3) Premium	(4) Number of brokers	(5) Number of brokers
	No broker	With broker	With broker	Contracted	Recommending
Oxford/United	23.1%	25.3%	9,341	551	319
Aetna	13.3%	13.1%	7,242	474	184
Excellus	11.0%	19.7%	6,562	234	129
Connecticut/Cigna	13.4%	15.1%	6,464	401	177
Empire/Anthem	12.9%	11.5%	9,412	417	177
Healthnow	8.5%	5.9%	6,255	229	76
HIP of NY/Emblem	7.7%	5.9%	7,232	341	127
MVP Health	7.4%	6.1%	5,937	251	82
Independent Health	5.8%	2.3%	5,268	71	36
CDPHP	7.6%	3.0%	6,528	132	51
Other	3.8%	4.9%	5,734	-	205

Notes: This table reports insurer-year level statistics for fully insured New York State ESI employers in 2011-2020. Columns (1) and (2) report market shares separately within the non-brokered and brokered employer-year units. Column (3) reports the average annual premium per enrollee among brokered employers, in dollars. Column (4) counts the number of brokers who have ever had an active contract to sell the insurer during the sample period. Column (5) counts the number of brokers who have ever placed at least one client with the insurer during the sample period. Carrier labels group affiliated entities (e.g., Oxford/United), and other aggregates all remaining insurers.

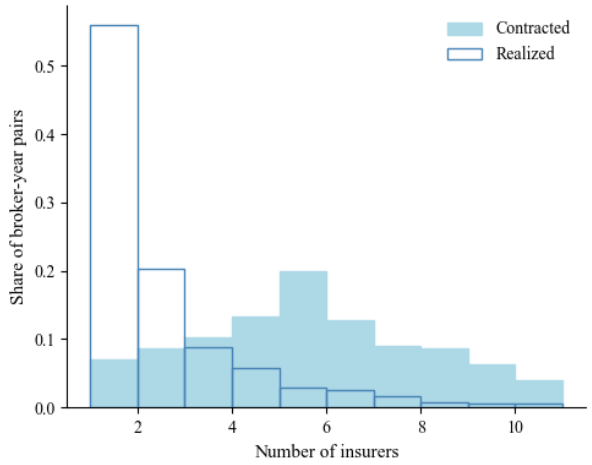
Broker-Insurer Contracts. Figure 2 illustrates the distribution of broker-insurer contracts and highlights three key patterns. First, selective contracting is widespread. Only 3% of brokers contract with all ten major insurers, and the median broker contracts with five insurers. Second, the set of insurers that brokers ultimately recommend (white bars) is only a subset of those with whom they contract (blue bars). Brokers place clients with fewer insurers than they contract with, underscoring the value of our data with precise contract initiation and termination dates: treating either all realized or all potential insurers as the relevant choice set would generate substantial measurement error. Finally, the broker-insurer networks evolve considerably over time, as contracts are formed and dissolved. These dynamics provide useful variation for identifying brokers' primitives in our network formation model.

Commission Schedules. Brokers face heterogeneous commission schedules across insurers. Of the ten major carriers in our sample, nine pay percentage-based commissions, with rates varying discontinuously with employer size or total premiums; only one pays a flat per-enrollee fee. Figure 3 depicts some examples: one insurer pays a 4% commission below \$0.5 million in premiums and 3% for additional premiums sold above that threshold; another pays 4% on policies below \$1 million and 3.5% above. Meanwhile, there are meaningful variations in the average commission percentages at the insurer-broker-market level. These discontinuities generate variation in broker incentives within an insurer, which we exploit in Section 4.

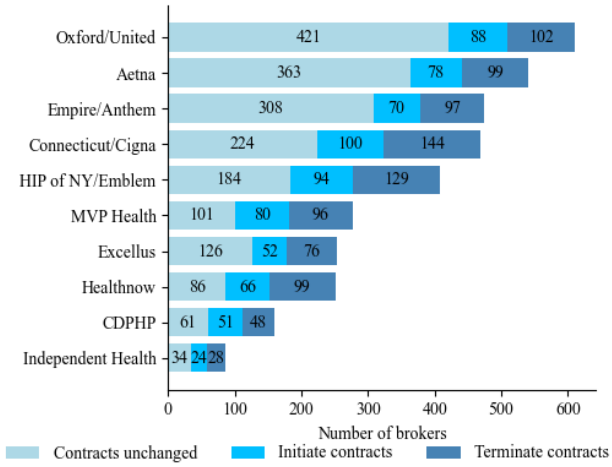
Limited Employer-Broker Switching. Table 3 reports statistics on employer-broker matching dynamics. Two patterns emerge. First, broker switching is infrequent. Employers work with an average of 1.68 brokers over the sample period, and only 13% of employer-year observations involve a switch from one broker to another. Transitions into and out of broker use are symmetric: 6.4% of employers move from no broker to broker, and the same share move from broker to no broker. Notably, these rates are nearly identical between the NYC region and non-NYC regions, despite the fact that the NYC market is substantially thicker. If broker competition disciplines matches, we would expect thicker markets to exhibit more frequent realloca-

Figure 2. Broker-Insurer Contracts Among Fully-Insured Employers

(a). Insurers contracted versus recommended



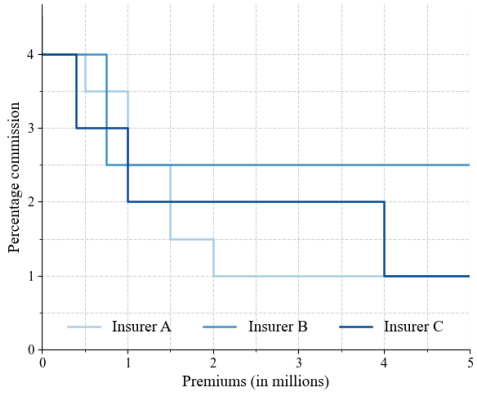
(b). Changes in contracting networks over time



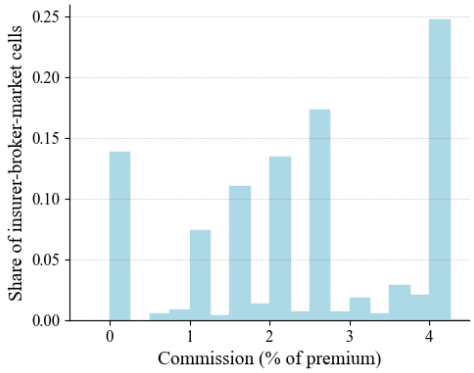
Notes: The figure depicts broker-insurer relationships among brokers that serve fully insured New York State ESI employers in 2011-2020. The sample is restricted to the ten major insurers reported in Table 2 in the empirical analysis. Panel (a) plots the broker-year-level distribution of the number of insurers each broker contracts with (blue) versus the number it places with at least one of its clients (white). Contracted pairs are measured with active broker-insurer contracts. Realized pairs are observed in Form 5500 filings. Panel (b) reports the number of broker relationships by insurer and decomposes these links into continuing, newly formed, and terminated contracts over the sample period.

Figure 3. Examples of Commission Schedules and Commission Dispersion

(a). Example marginal commission schedules



(b). Dispersion in average commission rates



Notes: Panel (a) plots example commission schedules for selected insurers in ESI. The x-axis shows the total annual premium per employer; the y-axis shows the marginal percentage commission specified in the schedule. Marginal commission rates step down at insurer-specific premium thresholds. Panel (b) plots the distribution of effective (average) percentage commissions across insurer-broker-market-year cells. Flat per-enrollee commissions are converted into premium-equivalent percentage commissions where applicable. Commission schedules were obtained via a FOIA request to the NYS Department of Financial Services, and we report the 2011-2020 period.

tion. Using market thickness as a proxy for competitive pressure, the absence of a meaningful urban-rural gradient suggests that broker competition, even where it should be most intense, does not generate frequent reallocation of employer-broker relationships.

Second, switching does not appear to substantially change the insurer menu facing the employer. Among employers that switch brokers, the average change in network breadth following a broker switch is 0.36 insurers, indicating that new brokers offer roughly the same number of carriers as their predecessors. More-

over, 77% remain with the same insurer after switching brokers, and only 6.7% move to an insurer that was not previously in the old broker's network. These patterns suggest limited employer shopping across brokers on network breadth. This is consistent with the hassle-cost interpretation underlying our model: if brokers differentiated themselves by offering access to novel insurers, they would effectively reimpose on employers the very search frictions that motivate employers to hire a broker in the first place. This interpretation is reinforced by the fact that switches are largely shock-driven, prompted by governance changes or broker retirement (see Section 2.1), rather than by employers shopping across brokers on network breadth.

Table 3. Broker switching and network changes

	(1) Full Market	(2) NYC Region	(3) Non-NYC Region
(a). Broker Switching			
Average number of brokers per employer	1.68	1.67	1.68
Share switching brokers (A to B), unforced	10%	10.4%	9.7%
Share switching brokers (A to B), forced	2.5%	2.7%	2.3%
Share switching (broker to no broker), unforced	1.6%	2.0%	1.3%
Share switching (broker to no broker), forced	0.2%	0.3%	0.1%
Share switching (no broker to broker)	2.7%	2.8%	2.6%
(b). Among Switchers (A to B)			
Average network pre-switch minus post-switch	−0.36	−0.39	−0.33
Share staying with same insurer	77%	74%	80%
Share changing to new insurer	23%	26%	20%
of which, new insurer not previously offered	6.7%	0%	10%

Notes: This table reports broker switching at the employer-year level for fully insured New York State ESI employers in 2011-2020. The denominator in panel (a) is all single-carrier fully-insured employer-year pairs that exist in two consecutive years; the denominator in panel (b) further filters to those who switch from A to B. The broker switch occurs when an employer is observed with broker A in the previous year and with broker B in the current year. Transitions into and out of broker use are defined analogously. Forced switches are those whose prior broker exited the market. The NYC region includes employers located in the New York City business region; the non-NYC region includes all other New York business regions. Among switchers, network change is the difference between the number of major insurers in the new broker's contracted network and the number in the previous broker's contracted network. New insurer not previously offered means that the employer's post-switch insurer was not in the previous broker's contracted network.

3. Conceptual Framework

In this section, we illustrate why commission regulation has ambiguous welfare effects. We use a stylized model to isolate three forces: how commissions shape broker menu formation, how commissions distort broker steering, and how these forces combine to impact welfare. We state the results in the main text and relegate the derivations to Appendix C.

Consider a market with two insurers, A and B, each offering a single product, and a broker who receives flat commissions, γ_A and γ_B.[14] The game proceeds in three stages. First, the broker observes commissions[15] and chooses one of three possible menus: contracting with both insurers A+B, contracting only with A, or contracting only with B. Second, employers meet with the broker and receive price quotes from insurers. If

[14] The intuition and qualitative predictions still hold if the broker instead receives commissions as a percentage of premiums.

[15] We treat commissions as exogenous because our focus is how commission caps affect menus and welfare, rather than how insurers set commissions. In the data, we barely observe changes in commission schedules throughout our sample period. Appendix C discusses the strategic-commission case and the associated equilibrium-existence concerns.

both insurers are in the menu,prices are p_A^c, p_B^c; if only one insurer is in the menu, prices are p_A^m, p_B^m, with $p_j^m > p_j^c,\ j \in \{A, B\}$. Third, employers choose among insurers A, B, or remain uninsured (U).

The flow utility of employers from option $j \in \{A, B, U\}$ under contracting arrangement $l \in \{c, m\}$ is:

$$V_A^l = \delta_A - \alpha p_A^l - \rho\gamma_B + \epsilon_A, \quad V_B^l = \delta_B - \alpha p_B^l + \epsilon_B, \quad V_U^l = \delta_U - \rho\gamma_B + \epsilon_U.$$

where δ_j is the employer's brand preferences for product j, p_j^l is the price under contracting arrangement l, and ϵ_j is an independent and identically distributed draw from a type-I extreme-value distribution.

Steering enters employer utility through the term $\rho\gamma_{-j}$, which reduces the perceived utility of the non-steered option when the rival insurer offers a higher commission γ_{-j}. The parameter ρ captures brokers' steering ability: a larger ρ implies that commissions exert greater sway over employers' choices through broker recommendations. We include the commission level γ_{-j} in employer utility as a reduced-form way of capturing broker influence within a single employer choice equation, consistent with the full model in Section 5.1. Appendix C6 shows that this specification can be microfounded by a model in which employers maximize their own objective while brokers separately maximize a payoff that combines commission income and employer surplus.

The broker's profit equals the expected commission income minus contracting costs. Depending on the contracting strategy, the broker earns:

$$\Pi_{AB} = s_A^c\gamma_A + s_B^c\gamma_B - 2c, \quad \Pi_A = s_A^m\gamma_A - c, \quad \Pi_B = s_B^m\gamma_B - c,$$

where c is the fixed cost of contracting with a single insurer, and s_j^l is the market share of insurer j under arrangement l.

We impose three conditions:

$$\delta_A > \delta_B, \quad s_B^m > s_B^c, \quad s_A^c\gamma_A - c > 0. \tag{1}$$

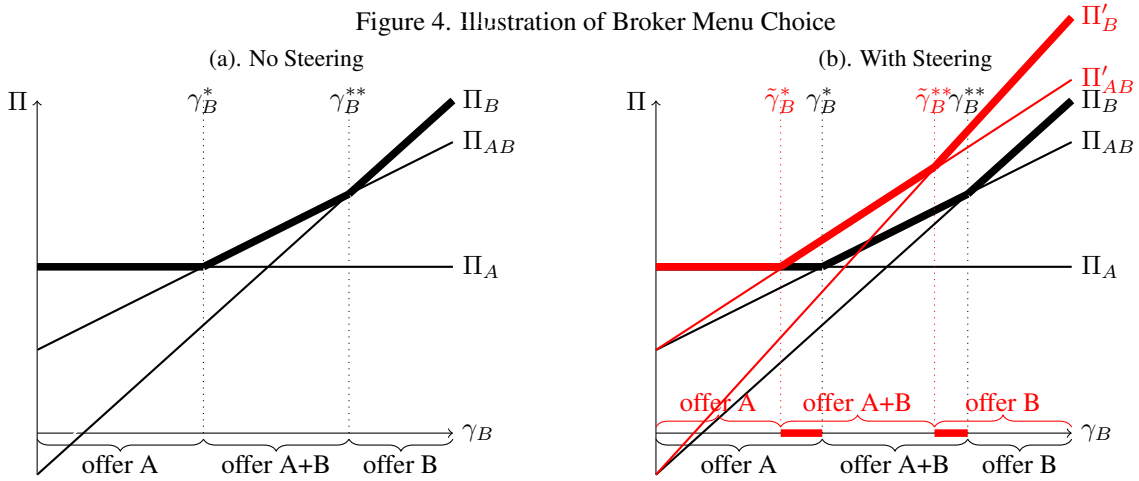
The first condition states that employers inherently prefer A to B. The second condition states that B is chosen less often than it would be if offered exclusively against the outside option. The third condition states that A is sufficiently attractive that under combined contracting, it remains profitable for the broker even if B offers no commission and incurs no contracting costs.

3.1. Broker Incentives and Menu Formation

We start by considering how commissions shape the broker's menu, abstracting from steering ($\rho = 0$). Offering an insurer generates commission revenue but requires paying the fixed contracting cost c. The broker chooses the menu that maximizes profit, corresponding to the upper envelope of the three profit lines Π_A, Π_{AB}, and Π_B in Figure 4a. Holding γ_A fixed, raising insurer B's commission γ_B raises the return to including B. As a result, the broker's optimal menu shifts in stages: when γ_B is low, the broker offers only A, reflecting employers' inherent preference for A; as γ_B rises past a threshold γ_B^*, the broker expands the menu to A+B; and at a higher threshold γ_B^{**}, expected revenue from B dominates and the broker contracts only with B, forgoing the contracting cost of A. Appendix C reports detailed derivations.

This analysis highlights the ambiguous effect of commission capping on brokers' menu choices. Small

Figure 4. Illustration of Broker Menu Choice



Notes: This figure plots how brokers' profits from alternative menus as insurer B's commission γ_B varies, holding other primitives fixed. The broker chooses the upper envelope. Panel (a) abstract from steering, $\rho = 0$ (black lines); Panel (b) allows steering toward B, $\rho > 0$ (red lines). Tilded thresholds denote menu-switching cutoffs. Curves are illustrative; Appendix C gives derivations and the local ordering conditions.

changes in γ_B may leave contracting decisions unchanged unless they cross thresholds γ_B^* or γ_B^{**}. Starting from a high baseline commission, an exogenous reduction from above to below γ_B^{**} would expand the broker's menu from B alone to A+B, as catering to employer preferences increases expected profits. Conversely, when the baseline commission is low, an exogenous reduction from above to below γ_B^* would shrink the menu from A+B to only A, as the expected commission income from B is no longer sufficient to justify contracting costs.

3.2. Steering

We next allow brokers to influence employer choices, $\rho > 0$. Steering enters employer utility through the term $\rho\gamma_{-j}$ in equation (1): when insurer B pays a higher commission, the broker shifts employer choice toward B, away from the option the employer would otherwise pick. Without loss of generality, we assume steering occurs toward B.

With steering, commissions affect broker profits in two ways: by raising per-employer payments and raising the likelihood that employers follow the broker's recommendation toward B. As a result, broker profits respond more strongly to γ_B than in the no-steering case: steering steepens both profit slopes, which widens the range of commissions over which the broker offers B, either exclusively or alongside A, as illustrated in Figure 4b. A higher steering ability ρ enlarges this range. Appendix C reports detailed derivations.

This analysis reveals that the effect of steering on brokers' menu choices is ambiguous. If the baseline commission level lies between γ_B^* and $\tilde{\gamma}_B^*$, the additional revenue from shifting employers to B outweighs contracting costs, expanding the menu from A to A+B. If the baseline commission level lies between γ_B^{**} and $\tilde{\gamma}_B^{**}$, revenue from B is high enough that the broker forgoes contracting costs with A, shrinking the menu from A+B to B. A higher steering ability ρ increases the range of commissions where menu changes may occur, amplifying the sensitivity of broker contracts to commissions.

3.3. The Welfare Tradeoff

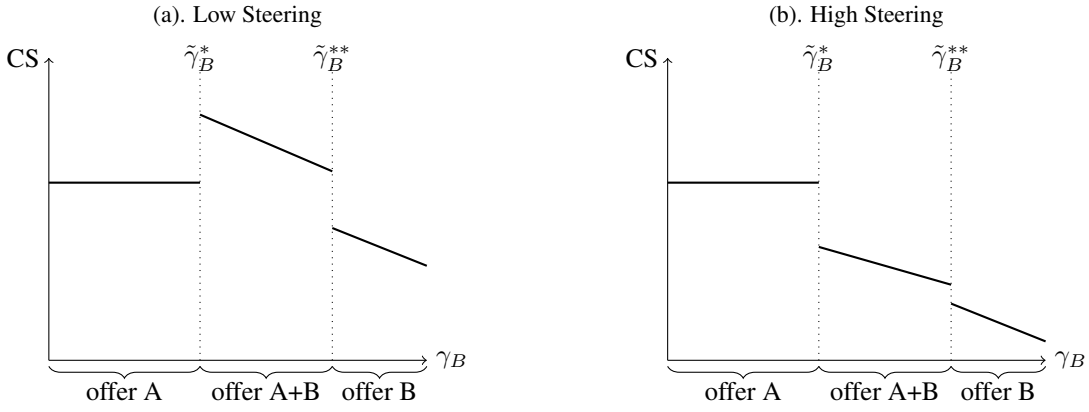
We now consider employer surplus, taking into account the broker's menu choice. We measure surplus in utility units and suppress the common $1/\alpha$ scaling. Expanding the menu from A to A+B creates two benefits: more product variety and stronger price competition, while steering toward the high-commission plan creates one cost: employers may choose a plan that is not best for them. Evaluating employer surplus at the commission $\tilde{\gamma}_B^*$, where the broker is indifferent between the A and A+B menus, we find the welfare change decomposes into exactly these components:

$$\begin{aligned} \mathrm{CS}_{A+B} - \mathrm{CS}_A = & \underbrace{\ln\left[\frac{\exp(\delta_A - \alpha p_A^c) + \exp(\delta_B - \alpha p_B^c) + \exp(\delta_0)}{\exp(\delta_A - \alpha p_A^m) + \exp(\delta_0)}\right]}_{\text{Competitive Effect } (+)} \\ & + \underbrace{\ln\left[\frac{\exp(\delta_A - \alpha p_A^c - \rho\tilde{\gamma}_B^*) + \exp(\delta_B - \alpha p_B^c) + \exp(\delta_0 - \rho\tilde{\gamma}_B^*)}{\exp(\delta_A - \alpha p_A^c) + \exp(\delta_B - \alpha p_B^c) + \exp(\delta_0)}\right]}_{\text{Steering Effect } (-)}. \end{aligned} \tag{2}$$

The competitive effect is always positive, and the steering effect is always negative; Appendix C unpacks the competitive effect into its price-competition and product-variety components and signs each term.

Whether the steering or the competitive effect dominates is an empirical question. Figure 5a and 5b illustrate two contrasting cases. When steering ability is low, menu expansion improves welfare: surplus jumps up at $\tilde{\gamma}_B^*$ as B enters and competition lowers prices, so the welfare-maximizing policy permits the A+B menu. By contrast, when steering ability is high, the steering distortion outweighs the competitive benefit: surplus under A+B falls below the A-only level, and the welfare-maximizing policy is a low commission at which the broker offers only the employer's preferred insurer. A commission cap thus faces a genuine tension: it may reduce steering, but it can also weaken brokers' incentives to maintain broad networks and the competition those networks bring.

Figure 5. Employer Surplus Conditional on Broker Menu



Notes: This figure is illustrative. It plots employer surplus under the broker's endogenous menu as insurer B's commission γ_B varies. The vertical lines denote the with-steering menu thresholds $\tilde{\gamma}_B^*$ and $\tilde{\gamma}_B^{**}$. Panel (a) shows a low-steering case in which the competitive-access gain from adding B dominates; Panel (b) shows a high-steering case in which the steering loss dominates. Surplus is in log-sum units, and curves are not drawn to scale.

3.4. Empirical Implications

The framework yields three implications that organize our empirical analysis. First, when a broker is involved, employer choices should respond to commission differences, i.e., the steering margin. Second, broker-insurer contracting should respond to commission levels, i.e., the network-formation margin. Third, premiums should respond to broker network breadth because the menu determines how intensely insurers compete for a given employer, i.e., the insurer-pricing margin.

These three margins, i.e., steering, broker-insurer network formation, and insurer pricing, are precisely the forces in tension in the welfare decomposition above. In what follows, we first present reduced-form evidence on each force in Section 4, and then quantify their joint welfare consequences by structurally estimating a model in Section 5.

4. Evidence of Steering and Competitive-Access Distortions

As Section 3 analyzes, commissions can affect market outcomes through three margins: brokers' recommendations, broker-insurer contracting, and insurer pricing. Below, we examine each margin in turn. First, we use a regression kink design to test whether brokered employers are more likely to choose insurers that pay brokers higher commissions. Second, we use time-series changes in commission schedules to test whether brokers' contracting networks respond to compensation. Third, we ask whether broker cohorts exposed to commission-driven network changes experience changes in rival insurers' premiums.

4.1. Effects of Commissions on Broker Steering

Empirical Strategy. To examine whether brokers steer employers toward insurers that offer higher commissions, we exploit within-insurer variation in commission rates across enrollment size. As depicted in Figure 6, Insurer 1 reduces its marginal broker commission from 4% to 3% once enrollment exceeds 500 employees. This discrete drop in the marginal rate produces a kink in the average commission a broker earns on Insurer 1: average rate is flat at 4% below 500, then declines toward 3% above. Because no other insurer changes commissions at the cutoff, Insurer 1's attractiveness to brokers declines in slope at the threshold, relative to employers just below it. This kink in relative attractiveness allows us to implement a regression kink design (RKD; Card et al., 2015).

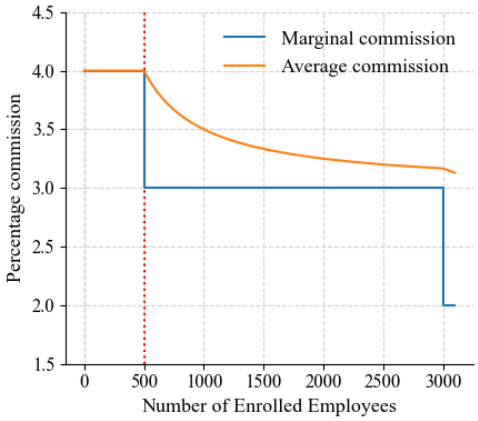
We exploit this kink to test whether employers on either side of the threshold differ in their propensity to choose Insurer 1. We focus on employers whose broker has Insurer 1 in their network and estimate:

$$Y_f = \alpha + \tau \cdot \mathbb{1}[X_f \geq K] \cdot (X_f - K) + g(X_f - K) + \varepsilon_f, \tag{3}$$

where Y_f is an indicator for whether employer f chooses Insurer 1 and $K = 500$ is the commission kink cutoff[16]; $g()$ is a flexible polynomial in the distance to the cutoff. The running variable X_f is the number of enrolled employees, the count of covered employee lives reported on Form 5500 filing. This is the precise margin on which Insurer 1's schedule is written, and it is neither eligible employees (i.e., the firm's total

[16]Figure A1 reports the universe of commission discontinuities in our data; the Insurer 1 threshold is the only kink that combines a clean non-price running variable, sufficient support both around the cutoff and for a matched non-brokered control group.

Figure 6. Commission Schedules, Insurer 1



Notes: The figure plots Insurer 1's broker commission as a function of enrolled employees. The marginal commission rate drops from 4% to 3% at 500, generating a kink in the average commission rate. The vertical line marks the 500-enrollee cutoff used in the regression kink design.

headcount, which enters only as a predetermined covariate) nor a constructed size index. The parameter of interest, τ, captures the change in the slope of the choice probability at the threshold. Identifying τ requires that every other determinant of insurer choice other than commissions evolve smoothly through the cutoff K, such that any slope break is attributable to the change in commission rates. This smoothness requirement is demanding in the simple RKD. We therefore difference against a natural control group.

The control group consists of employers of similar size who do not use a broker: they face the same enrollment margin but bear no broker's financial incentive. Pooling the brokered and non-brokered groups yields a "difference-in-kinks" (DiK) design (Böckerman et al., 2025), the regression kink analog of the difference-in-discontinuities (Grembi et al., 2016):

$$\begin{aligned} Y_f =& \beta_0 + \beta_1 \mathbb{1}_{fb} + g_0(X_f - K) + \mathbb{1}_{fb} \cdot g_1(X_f - K) + \tau_0 \cdot \mathbb{1}[X_f \geq K] \cdot (X_f - K) \\ &+ \tau^{DiK} \cdot \mathbb{1}_{fb} \cdot \mathbb{1}[X_f \geq K] \cdot (X_f - K) + \varepsilon_f, \end{aligned} \tag{4}$$

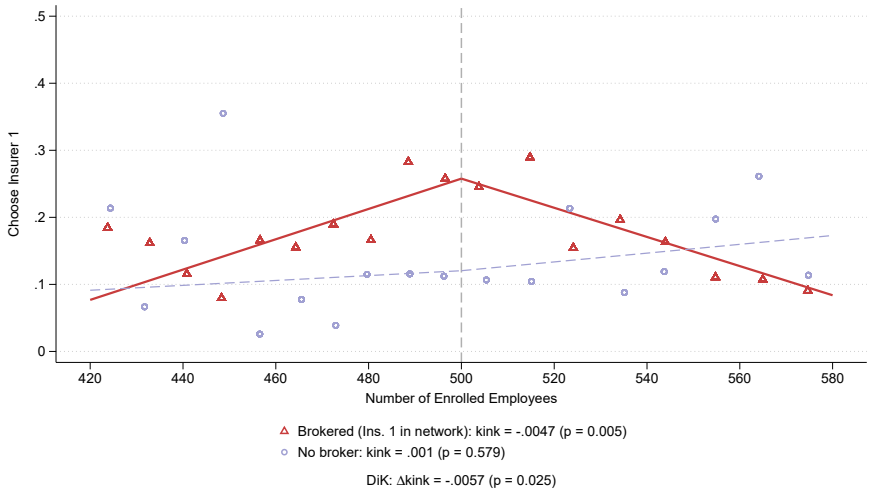
where $\mathbb{1}_{fb}$ indicates whether employer f uses a broker. The parameter of interest, τ^{DiK}, measures the excess slope change at the commission kink for brokered relative to non-brokered employers. Commission-driven steering implies $\tau^{DiK} \neq 0$: only brokered employers, whose insurer choice is intermediated by a commission-earning agent, should exhibit a differential kink at the commission threshold.

Identification of τ^{DiK} relies on a parallel-kinks assumption: absent commission-driven steering, the slope of insurer choice with respect to enrollment would change at the cutoff by the same amount for both groups. This does not require the two groups to share the level or slope of insurer choice, only that neither would exhibit a differential kink at the cutoff for reasons unrelated to the commission schedule.

Our preferred specification uses a local linear polynomial with a bandwidth of 80 enrolled employees on each side of the cutoff, slightly below the CER-optimal bandwidth of 97.[17] We restrict to fully insured, single-carrier employers, who are best represented in Form 5500, and include covariate adjustments for the

[17] In an RKD, the object of interest is a derivative; it is more sensitive than a level to bias from observations far from the cutoff (Card et al., 2015). We therefore undersmooth relative to the CER-optimal bandwidth, trading precision for bias reduction (Calonico et al., 2014).

Figure 7. Difference-in-Kinks: Brokered vs. Non-Brokered Employers, Probability of Choosing Insurer 1



Notes: This figure plots binned means and fitted values from the difference-in-kinks specification in equation (4). The outcome is an indicator for choosing Insurer 1. The treatment group (red dots, solid lines) consists of brokered employers with Insurer 1 in the broker's network. The control group (grey dots, dashed lines) consists of employers who do not use a broker. The running variable is the number of enrolled employees, centered at Insurer 1's 500-enrollee commission cutoff (Figure 6). Bins are local polynomial means. Fitted lines use a local linear specification with triangular kernel weights and an 80-enrollee bandwidth on each side of the cutoff. Reported kinks are slope changes at the cutoff; the DiK estimate is the brokered minus non-brokered slope change.

NYC and Mid-Hudson regions to absorb compositional differences between downstate and upstate NY.

Results. Figure 7 reports the main estimates. The differential slope change is sharp and statistically significant: beyond the threshold, the probability that a brokered employer chooses Insurer 1 bends downward relative to its pre-kink trend. The components are consistent with the design: brokered employers kink downward at the cutoff, whereas the non-brokered control exhibits no discernible kink. These slope changes are estimated locally at the cutoff, not as a level shift, and are identified from the bend in the choice probability rather than their extrapolation across the window. These estimates indicate that brokers respond to financial incentives by steering employers toward the insurer that compensates them more generously.

Figure A2 traces the consequences of this steering. Panel (a) examines premiums per enrollee. Below the cutoff, Insurer 1 offers the highest commission and is also a relatively expensive carrier, so steering places more brokered employers in a costly plan and raises their premiums; above the cutoff, the lower commission weakens this pull and shifts placements toward cheaper carriers. Premiums should therefore kink downward at the threshold, and they do: for brokered employers, the slope of log premiums per enrollee decrease beyond the cutoff, while the non-brokered control group shows no comparable kink. Panel (b) examines the probability of choosing Insurer 2, the carrier offering the second-highest commission in this firm size range, and the same carrier whose flat-fee schedule we exploit in Section 4.2. Its choice probability kinks upward at the same cutoff, consistent with brokers substituting toward the next most lucrative carrier once Insurer 1's advantage weakens. The brokered-group kinks in both panels are individually significant; DiK estimates carry the predicted signs but are less precise, reflecting the thinner non-brokered control sample for these outcomes. Taken together, the choice, premium, and substitution patterns indicate that the commission kink changes brokers' incentives, and employer choices move accordingly.

As an additional test, we examine brokered employers whose brokers do not contract with Insurer 1. For

these employers, the probability of choosing Insurer 1 is mechanically zero, so they cannot serve as a control group for the choice of Insurer 1. We therefore reframe the test by detecting kinks in *conditional* densities. In Appendix Figure A6, we plot in panel (a) the density of firm size conditional on firms with Insurer 1 in their choice set (via their contracted broker) but who do not choose Insurer 1. Here, there is a significant upward kink the density at the commission cutoff at 500, as this is the mirror image of the RKD in Figure 7. Before the cutoff, steering towards Insurer 1 suppresses the mass of firms choosing rival carriers, but once Insurer 1's commissions fall at 500 employees, this suppression flattens. In Panel (b) we plot a placebo test: the density conditional on firms not having Insurer 1 in their choice set, and therefore facing no change in Insurer 1's commissions. Here, we find that there is no analog kink in the conditional density.

Validity and Robustness Checks. A potential threat to our design is a determinant of choice that bends specifically at 500, specifically for brokered employers, for reasons other than commissions. Three features make such a confound unlikely. First, the threshold lies well inside New York's large-group market, whose small-group segment ends at 100 employees, so no rating, medical-loss-ratio, or guaranteed-issue rule changes nearby. We are not aware of any regulatory or institutional threshold of 500 employees in this market, apart from Insurer 1's commission schedule. Second, any change that does occur at 500 for all employers, including any insurer pricing, underwriting, or plan-design rule, is removed by the difference-in-kinks comparison, since it would show up in the control as well. Third, as shown above, the non-brokered control group exhibits no discernible kink; because these employers also purchase Insurer 1, the absence of a kink argues against a supply-side change by Insurer 1 at the cutoff.

We next examine two additional concerns: manipulation of enrollment around the cutoff and differences in predetermined characteristics across the threshold. First, enrollment must not be manipulated around the cutoff. Because the running variable is enrolled lives, the natural concern is intensive-margin manipulation (Lee and Lemieux, 2008): a broker could steer an employer just above 500 toward less generous plans within Insurer 1, depress take-up, and thereby return the firm's enrollment below the cutoff. Such manipulation could generate bunching at the threshold. We find no evidence of this pattern: density tests (Cattaneo et al., 2018; McCrary, 2008) show the distribution of enrolled lives passing smoothly through 500 (Figure A3), and the eligible-employee count, which is not mechanically tied to carrier choice, exhibits no kink of its own (Figure A5(f)). Second, predetermined characteristics must be smooth at the threshold. Across seven broker and employer covariates, including network breadth, individual-agent status, firm size, and industry composition, we find no systematic kink at the cutoff (Figure A5).[18]

We next confirm that the estimate is robust to the specification choices that govern an RKD. First, we guard against spurious curvature: because a kink estimator can misread ordinary curvature in the choice probability as a break in slope (Card et al., 2015), we re-estimate the kink at placebo cutoffs from 300 to 700 (Figure A4). The true cutoff, 500, yields the largest estimate in magnitude and is the only significant cutoff outside mechanically contaminated neighborhoods. The nearby placebo cutoffs at 450 and 550 fall inside the estimation bandwidth and are therefore affected by the true kink; placebo cutoffs outside the bandwidth are indistinguishable from zero. Second, we vary the remaining tuning choices, including CER-

[18]The single exception, the education-and-health share, is marginally significant, is driven by a small cluster of firms just above the cutoff, and vanishes under a ± 10 donut; with seven tests, one marginal estimate is well within chance.

and MSE-optimal bandwidths, quadratic polynomials, a uniform kernel, alternative covariate adjustments, and donut-hole exclusions (Table A1). The estimate is stable in sign, size, and significance across these choices. Notably, the ± 10 donut, which removes the very firms responsible for the lone covariate imbalance above, increases the kink rather than attenuating it.

4.2. Effects of Commissions on Insurer-Broker Contracts

Empirical Strategy. Next, we examine whether commissions influence brokers' contracting decisions with insurers, exploiting plausibly exogenous, time-varying changes in broker commissions. Insurer 2 is the only insurer that reimbursed brokers on a per-member-per-month (PMPM) flat fee basis rather than a percentage of premiums, and it changed the level of its flat fee twice throughout our sample period: from \$5 to \$18 PMPM in 2014, then from \$18 to \$12 in 2016. No other insurers changed their commission schedules during this period, so these two changes isolate movements in broker incentives, holding the rest of the market fixed.

We estimate the event study:

$$Y_{jct} = \sum_{k \neq -1} \beta_k \cdot \mathbb{1}_{j=2} \cdot \mathbb{1}[t - t_j^* = k] + \alpha_j + \alpha_{ct} + \varepsilon_{jct}, \tag{5}$$

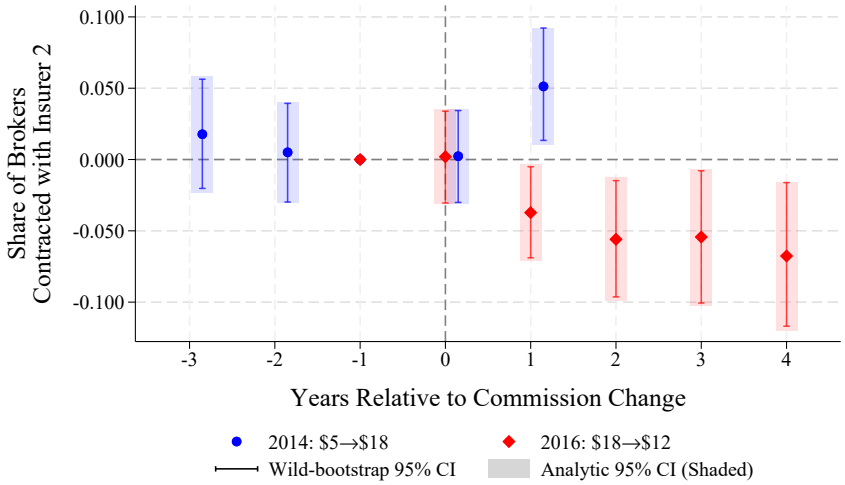
where Y_{jct} is the share of brokers contracting with insurer j in county c at time t, $\mathbb{1}_{j=2}$ is an indicator for Insurer 2, t_j^* is the event year (2014 or 2016), and α_j, α_{ct} are insurer and county-time fixed effects. The coefficient of interest, β_k, captures Insurer 2's contracting share relative to other carriers, normalized to the year before each event.[19] Identification requires that, absent the commission changes, Insurer 2's contracting share would have evolved similarly to those of the comparison insurers, conditional on the fixed effects.

Results. Figure 8 displays the results: when Insurer 2 increased commissions in 2014 (blue dots), its share of contracting brokers rose relative to other insurers that experienced no schedule changes. Conversely, when Insurer 2 decreased its commissions in 2016 (red dots), that share fell. These estimates indicate that the 2016 reduction in commissions lowered Insurer 2's share of contracting brokers significantly relative to carriers whose schedules did not change, and that the decline persists across the post-event years. Figure A8 further decomposes the contracting response into entry and exit. The 2014 increase raises entry into Insurer 2's network, while the 2016 cut raises exit. We interpret these estimates as market-level evidence that broker-insurer contracting responds to commission incentives.

Validity Checks and Mechanism Evidence. A natural concern is that Insurer 2 changed other aspects of its business alongside its commissions, such that the contracting response reflects shifts in its underlying attractiveness rather than compensation. To address this concern, we re-estimate equation (5) using Insurer 2's own characteristics as outcomes. If the fee changes were isolated adjustments to broker pay, Insurer 2's characteristics should not move at the event dates. We examine three categories. First, financial performance: Insurer 2's large-group medical loss ratio relative to other carriers is flat through both events (Figure A9), showing no evidence of contemporaneous changes in pricing, underwriting, or risk selection. Second,

[19] The 2014 event study uses 2011-2015, ending before the subsequent commission cut. The 2016 event study uses 2015-2020, with 2015 providing a baseline under the higher commission schedule. Figure A7 reports the full 2011–2020 trajectory relative to a 2013 baseline.

Figure 8. Effect of commission changes on broker networks



Notes: This figure plots the point estimates and 95% confidence interval from the estimation of equation (5). The outcome is the county-year level share of brokers contracting with each insurer, and estimates compare Insurer 2 with other insurers. We report estimates from two separate event studies: blue markers correspond to Insurer 2's 2014 commission increase from $5 to $18 PMPM, while red markers correspond to its 2016 commission decrease from $18 to $12 PMPM. We normalize the estimates to the year preceding each commission change. Solid bars denote wild-bootstrap 95% confidence intervals, while shaded bands denote analytic 95% confidence intervals; both are clustered by county.

plan generosity (deductibles and coinsurance) and provider access (the number and share of in-network NY hospitals) (Figure A10) stay flat relative to other carriers, except for coinsurance at the 2016 cut.[20] Third, employee take-up: the share of eligible employees who enroll in Insurer 2's plans when offered is unchanged at both events (Figure A11). Take-up provides a useful revealed-preference summary of plan value, since it should respond to salient changes in price, plan characteristics, or unobserved quality. Its stability indicates that the overall value of Insurer 2's coverage did not change discretely at the commission events. Taken together, these results make it unlikely that an unobserved improvement in Insurer 2's product offering, rather than changes in broker compensation, explains the contracting response.

As an additional placebo test, Figure A12 compares Insurer 2's contracting responses with placebo estimates obtained using other insurers and alternative event years. The decline following the 2016 commission cut lies in the lower tail of the placebo distribution, whereas the increase following the 2014 commission change is not statistically distinguishable from the placebo estimates.

4.3. Premium Responses to Commission-Induced Network Changes

Empirical Strategy. Finally, we examine how broker contracting networks affect competition among insurers. Because contracting is selective, employers are not exposed to every carrier in the market, and insurers do not all compete head-to-head within a given employer's choice set. Broker networks therefore determine the scope of competition each insurer faces.

[20]The plan-characteristic measures are built from claims data beginning in 2013, so the 2014 event is estimated against a single pre-period year. This lone coinsurance response need not imply that plan generosity, rather than commissions, drives contracting: the contracting response is present at both events (Figure 8), whereas the coinsurance change appears only at 2016, and we find no evidence of a plan-characteristic change at the 2014 event despite its shorter window. The absence of a corresponding change in 2014 makes plan generosity a less compelling explanation for the responses across both events.

We examine whether broker cohorts exposed to commission-driven network expansion or contraction face different prices from other (non-Insurer 2) insurers. The treatment is assigned at the broker level, based on their pre-period exposure to Insurer 2. For the 2014 commission increase, treated brokers are those who did not contract with Insurer 2 before the event and therefore could newly add it to their menus; brokers already contracting with Insurer 2 beforehand serve as the control group. For the 2016 cut, we flip the groups since the exposure is flipped: brokers already contracting with Insurer 2 are exposed to dropping it, while brokers without Insurer 2 are relatively unaffected. We estimate the response of premiums charged by other insurers, excluding Insurer 2:

$$P_{jfbct} = \sum_{k \neq -1} \beta_k \cdot \text{Exposed}_b \cdot \mathbb{1}[t - t^* = k] + \alpha_f + \alpha_b + \alpha_j + \alpha_{ct} + \varepsilon_{jfbct}, \tag{6}$$

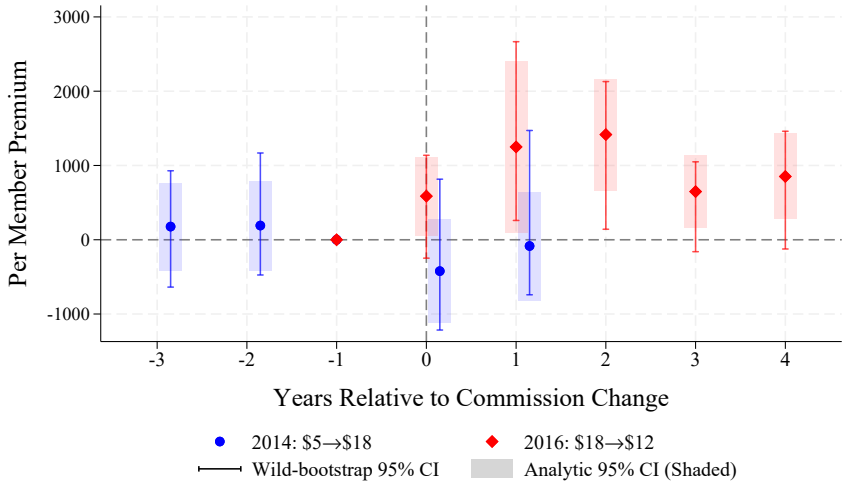
where P_{jfbct} is the premium per employee of employer f by insurer $j \neq 2$ offered through broker b in county c at time t; Exposed_b identifies the treated broker group for each event; α_f, α_b, α_f, and α_{ct} are employer, broker, insurer, and county-year fixed effects that control for time-invariant employer risk, time-invariant broker and insurer characteristics, and local market shocks.

Equation (6) estimates the effect of predetermined exposure to Insurer 2's commission changes on rival premiums. Identification requires that, absent these changes, premiums would have evolved similarly for exposed and comparison brokers, conditional on the fixed effects. Interpreting this response as operating through broker networks additionally requires that the commission changes affect rival premiums through changes in broker–insurer contracting, rather than through other channels that differentially affect exposed and comparison brokers. Under these assumptions, the comparison provides evidence on how commission-induced changes in broker networks affect insurer competition.

Results. Estimates in Figure 9 are consistent with the competition mechanism: rival premiums rise for brokers exposed to the 2016 commission cut, when Insurer 2's broker network contracts. The 2014 commission increase yields a smaller and less precise response in the opposite direction, consistent with slower relationship formation and downward price rigidity. The post-2016 estimates imply that brokers exposed to a commission-induced network contraction subsequently face higher premiums from rival carriers than comparison brokers do, and that the gap remains positive through 2020. This finding indicates that broker networks set the scope of rivalry by expanding or contracting the set of insurers employers can reach.

Validity and Robustness Checks. The design would fail if Insurer 2 simultaneously altered other dimensions of its strategy, such as plan design, provider networks, or pricing, alongside its commission changes, causing rival insurers to adjust their pricing in response to Insurer 2 rather than to changes in their own broker exposure. Several facts mitigate this concern. First, industry documents indicate that the 2014 increase and the 2016 reduction were introduced as a broker-compensation adjustment rather than as part of a product redesign or geographic expansion. Second, the placebo tests in Section 4.2 show no evidence of contemporaneous changes in Insurer 2's underlying product offering. Third, treated brokers adjust little along other margins: around the 2016 change, the number of non-Insurer-2 carriers in their networks and the number, size, and composition of their employer clients all remain stable (Figure A13). So the premium response does not appear to reflect a broader re-optimization of broker portfolios. We cannot rule out small,

Figure 9. Effects of exposure to Insurer 2 commission events on rival premiums



Notes: This figure plots the point estimates and 95% confidence interval from the estimation of equation (6). The outcome is the annual premium per enrollee charged by other insurers, excluding Insurer 2. For the 2014 commission increase (blue markers), treated brokers are those not contracting with Insurer 2 before the event; controls are brokers already contracting with Insurer 2. For the 2016 commission decrease (red markers), treated brokers are those contracting with Insurer 2 before the event; controls are brokers not contracting with Insurer 2. Estimates are normalized to the year before each event. The specifications include employer, broker, insurer, and county-by-year fixed effects. Solid bars denote wild-bootstrap 95% confidence intervals, while shaded bands denote analytic 95% confidence intervals; both are clustered at the county level.

contemporaneous changes in Insurer 2's conduct, but given the stability of its core product and of brokers' other margins, any such residual changes would likely have only a limited effect on our outcome measures and are therefore unlikely to be the primary driver of the estimated effects.

We next separate the competition interpretation from two alternatives: changes in employer composition and market-wide pricing shocks. First, we restrict to employer-insurer pairs observed both before and after the event (Figure A14) and hold the matched employer and rival insurer fixed. The response survives, indicating that rivals reprice existing relationships when choice sets narrow, rather than reflecting a change in which employers appear in the sample. Second, we split rival insurers by their pre-period broker-network overlap with Insurer 2 (Figure A15). The competition mechanism predicts larger responses among high-overlap rivals, whereas a market-wide trend or a change in Insurer 2's own conduct would move high- and low-overlap rivals alike. The estimates are consistent with the competition mechanism: following the 2016 commission cut, the largest premium increases occur among high-olver rivals. These estimates are less precise, reflecting the smaller sample. We therefore view the split as supporting evidence for the competition channel.

The 2014 comparison provides additional information about the mechanism. Commission changes may affect rival premiums through steering even when broker networks remain fixed. Before the 2014 increase, however, only control brokers carried Insurer 2. Holding networks fixed, stronger steering toward Insurer 2 would therefore operate among controls. If this steering lowers rival premiums, then, absent differential spillovers across broker groups, it would push the treatment-minus-control estimate upward. This channel would thus work against the negative point estimates we observe. The direction of the 2014 response supports a role for network expansion, although its imprecision limits the strength of this evidence.

Finally, Figure A16 compares the estimated premium responses with placebo estimates obtained by reassigning event years and broker-county exposure. Both responses lie in the predicted tails of the placebo distributions, with stronger evidence for the 2016 commission cut.

4.4. Summary and the Need for a Model

Our reduced-form exercise establishes three empirical patterns. First, commissions steer recommendations toward the carrier that pays brokers more, raising premiums when that carrier is the costlier option. Second, commissions shape contracting, as brokers add or drop carriers when commissions change, widening or narrowing the choice sets they offer. Third, network breadth governs competition: narrower broker-mediated choice sets soften insurer competition and raise premiums.

Together, these facts reveal the paper's central tradeoff. Lowering commissions can weaken steering toward high-cost carriers and lower premiums, but it can also shrink broker-insurer networks, softening competition and raising premiums. Because the two forces are of opposite sign, the net effect of a commission cap is a quantitative question that the reduced form cannot answer. We thus turn to the structural model to quantify their relative magnitudes and evaluate counterfactual commission regulation.

5. Model

Our empirical model quantifies how broker compensation affects employer choices, insurer pricing, and broker-insurer networks in equilibrium. The game proceeds in four stages. First, given commission schedules, brokers choose insurers to include in their menus and incur the associated contracting costs. Second, employers randomly meet with a broker and decide whether to use the broker's services. Third, insurers simultaneously set premiums conditional on the contracting structure. Fourth, employers choose one insurer for their workforce, either from the broker's menu or from the full market if they do not use a broker. These stages map onto our focal distortions: the menu stage generates competitive-access narrowing, while the insurer-choice stage captures commission-driven steering. Commission schedules are set by insurers, filed with DFS, and sticky within the sample period; we therefore take them as given and exogenous to the contracting and pricing decisions modeled here. In what follows, we present the details of each step in reverse order.

5.1. Employers' Choice of Insurer

Employers choose one insurer to maximize a weighted average of employee welfare and costs.[21] Let J denote the set of insurers available in the employer's geographic market, and let U denote the outside option of not offering employer-sponsored insurance, so that an employer's broadest possible choice set is $J \cup U$. We index employer types by f, where a type is defined by geographic market, industry, and firm-size categories. Because Form 5500 filings identify the insurer chosen by the employer but not the full set of

[21] As Form 5500 data only report insurer identity but not insurance plans or their detailed characteristics, we make the simplifying assumption that insurers are single-product oligopolies in a given geographic market, where financial characteristics are held fixed but product pricing could vary by market, employer type, and the competitive environment determined by contracting arrangements.

within-insurer plan designs, we interpret insurer j in market-year as a composite product. The flow utility of an employer of type f from choosing product j under broker menu m is:

$$u_{fjm} = \overbrace{-\alpha_{fj} \cdot \underbrace{(1-\iota_f) \cdot p_{fjm} \cdot \text{takeup}_{fjm}}_{\substack{\text{employer paid} \\ \text{premium}}}}^{\text{weight on own expenses}} + \overbrace{\rho \cdot \mathcal{W}_{fjm}(p_{fjm}, \text{takeup}_{fjm})}^{\text{weight on employee welfare}} + \delta_j^f + \xi_{fj} + \epsilon_{fj}, \tag{7}$$

where δ_j^f is insurer-specific brand effects perceived by the employer, such as brand, administrative quality, account service, or stable preferences of employer type f for insurer identity; ξ_{fj} is unobserved preference shocks; ϵ_{fj} is independent and identically distributed extreme-value type-I error.

The first term is the employer's premium contribution per eligible employee, weighted by its price sensitivity α_{fj}. The employer pays share $1-\iota_f$ of the premium p_{fjm} for each enrolled employee; and takeup_{fjm} is the share of eligible employees who enroll.

The second term, $\mathcal{W}_{fjm}(p_{fjm}, \text{takeup}_{fjm})$, measures the average ex-ante surplus employees receive from the observed and imputed plan attributes, associated with employer f, insurer j under broker menu m, evaluated at the equilibrium premium p_{fjm} and take-up rate takeup_{fjm}. The more that the plan of insurer j caters to employees' preferences, the higher the share of employees who take up the plan, and the larger the ex-ante surplus is. Both takeup_{fjm} and $\mathcal{W}_{fjm}$ are equilibrium objects derived from the employee demand system in Appendix D1.

We write b for an individual broker, and $b(f)$ for the broker type matched to employer f, e.g., an individual agent or a brokerage firm. $\mathbb{1}_{fb}$ is an indicator of whether the employer works with a broker to offer insurance. Let γ_{fj} denote the commission paid for placing employer type f with insurer j, expressed as an ad valorem equivalent. Percentage commissions enter directly; flat commissions are converted to premium-equivalent rates. We assume brokers affect employers' choices by adjusting how sensitive employers are to premiums:

$$\alpha_{fj} = \alpha_0 + \underbrace{\mathbb{1}_{fb} \cdot \alpha_{1,b(f)}}_{\text{broker status wedge}} + \underbrace{\mathbb{1}_{fb} \cdot \alpha_{2,b(f)} \cdot \gamma_{fj}}_{\text{steering}}. \tag{8}$$

$\alpha_{1,b(f)}$ captures non-commission differences in the way brokered employers' insurer choices load on premiums relative to non-brokered employers. We refer to this term as the broker-status wedge. It may reflect several forces: selection of employers into broker use, brokers' information provision or advice, administrative support, or other unobserved differences between brokered and non-brokered firms. We do not interpret $\alpha_{1,b(f)}$ as a causal effect of broker use or assign it a welfare.

$\alpha_{2,b(f)} \cdot \gamma_{fj}$ is the component of the choice rule that we interpret as commission-driven steering. $\alpha_{2,b(f)}$ captures brokers' steering ability, or the degree to which commissions distort brokers' recommendations. The linear-in-commission form $\alpha_{2,b(f)} \cdot \gamma_{fj}$ is a reduced-form decision rule for tractability, rather than a mechanical assumption that brokers steer in exact proportion to commissions. Appendix C6 shows that equations (7)-(8) are the reduced form of a model in which a broker maximizes commission income subject to a weight on the employer's objective surplus. We allow $\alpha_{1,b(f)}$ and $\alpha_{2,b(f)}$ to depend on $b(f)$ only through

the broker's type.[22]

5.2. Insurer Pricing

Insurers compete à la Nash-Bertrand. Conditional on the broker menu m that decides its competitors, insurer j sets premiums p_{fjm} for employer type f to maximize profits:

$$p_{fjm} = \text{argmax}_{p_{fjm}} \pi^I_{fjm} = \text{argmax}_{p_{fjm}} \left(s_{fjm} \cdot \text{takeup}_{fjm} \cdot (p_{fjm} \cdot (1 - \underbrace{\mathbb{1}_{fb}}_{\substack{\text{whether} \\ \text{use broker}}} \cdot \underbrace{\gamma_{fj}}_{\substack{\text{percentage} \\ \text{commission}}}) - mc_{fj})\right), \tag{9}$$

where s_{fjm} is the share of employers in type f choosing plan j given menu m, according to the decision rule outlined in equation (7); takeup_{fjm} is the share of employees taking up plan j. mc_{fj} denotes marginal costs of insurer j for insuring employer type f.

The marginal cost mc_{fj} is itself shaped by which employees enroll. Because employees decide whether to take up coverage, and because take-up responds to the premium, a change in p_{fjm} alters the demographic composition of the enrolled pool, and hence the insurer's average cost of serving employer type f. We write

$$mc_{fj} = \bar{c}_{fj}(p_{fjm}) + \omega_{fj}, \tag{10}$$

where $\bar{c}_{fj}(p_{fjm})$ is the enrollment-weighted average cost across employee demographics within employer type f, and ω_{fj} is an employer–insurer cost shock orthogonal to enrollee composition. The dependence of $\bar{c}_{fj}$ on price captures adverse selection: a higher premium disproportionately deters lower-cost employees, raising the average cost of those who remain. Appendix D2 derives $\bar{c}_{fj}(\cdot)$ from the employee take-up model and describes its calibration. We assume insurers observe the realizations of demand and marginal cost shocks (ξ, ω) when setting prices, whereas realizations of these shocks are unobservable to econometricians.

Insurers' first-order condition on p_{fjm} is

$$\left[\frac{\partial s_{fjm}}{\partial p_{fjm}} \text{takeup}_{fjm} + s_{fjm} \frac{\partial \text{takeup}_{fjm}}{\partial p_{fjm}}\right] \Big((1 - \mathbb{1}_{fb}\gamma_{fj})p_{fjm} - mc_{fj}\Big) + s_{fjm} \text{takeup}_{fjm} \Big(1 - \mathbb{1}_{fb}\gamma_{fj} - \frac{\partial \bar{c}_{fj}}{\partial p_{fjm}}\Big) = 0. \tag{11}$$

The first bracket captures losses from a higher premium due to a smaller share of employers choosing the insurer, and a smaller share of employees taking up the plan. The last term captures gains per-enrollee revenue, net of the commission, and the change in average cost induced by adverse selection.

5.3. Employers' Choice of Whether to Use a Broker

Given the set of brokers operating in their market, employers randomly meet a broker through a local matching process and decide whether to use the broker's services. This decision affects the employer on two margins. First, if the employer uses a broker, its choice set will only consist of the menu of insurers with which the broker has contracted $j \in m$. Instead, if it does not use a broker, its choice set will include all available insurers on the market, plus the option of not offering insurance $j \in J \cup U$. Second, the employer

[22] We assume that the intensity of steering is homogeneous across brokers of the same type. This is because we do not observe claims mapped to the employer-broker-insurer level, and thus cannot compute the extent to which brokers differentially leave money on the table. While this precludes heterogeneity in steering intensity, it offers a tractable approximation of average steering effects.

that forgoes broker services incurs hassle costs of insurer search, enrollment and renewal paperwork, and ongoing coordination with carriers.

Let V_{fb}^{wb} and V_f^{nb} denote employer f's ex-ante decision values with broker b and without brokers, respectively, calculated as the usual logsum surplus terms. The brokered value reflects the broker's insurer menu, equilibrium premiums, and the broker-mediated choice rule. We assume that employers evaluate broker use naively: they take the broker-mediated choice rule, including the broker-status wedge and the steering term, as their own preferences, and do not recognize that decision utility may differ from experienced utility. In our welfare analysis, we account for this difference following Allcott (2013).

Employer f uses a broker if the value of the broker's menu exceeds the value of purchasing without a broker, net of the hassle cost H_f:

$$\mathbb{1}_{fb} = 1 \quad \text{if} \quad V_{fb}^{wb} > V_f^{nb} - \underbrace{H_f}_{\text{hassle cost}}, H_f = \mu_f + \sigma\eta_f, \eta_f \sim N(0,1). \tag{12}$$

where μ_f is a function of observed employer characteristics and η_f an idiosyncratic shock. We refer to H_f as the hassle cost throughout, but interpret it broadly: it captures all benefits to the employer of using a broker not included in ΔV_{fb}, such as avoided search costs, administrative and claims-related assistance, and stable preferences for broker services; see Section 6.3 for more discussion.

We do not model brokers competing directly against one another, which amounts to assuming that employers do not actively shop across brokers on menu breadth. We adopt this assumption for two reasons. First, it is consistent with the stylized facts in Section 2.3: employer-broker relationships are persistent, switching rates are similar in thick and thin broker markets, and broker switches rarely expand access to new insurers. Appendix D3 also shows that geographic proximity is far more predictive of employer-broker matches than broker network breadth. Employers thus do not appear to shop across brokers on the very margin on which brokers would compete, so the discipline brokers face comes primarily from the outside option rather than from one another. Second, the assumption keeps the model tractable. Because each employer meets at most one broker, a broker's profit depends on its own menu and potential clients, but not on any other broker's menu. So the menu game is separable across brokers (Section 5.5). This separability lets us solve each broker's contracting problem independently, rather than as a fixed point across all brokers' menus, and removes strategic multiplicity from the menu stage.

The outside option still disciplines broker behavior. A broker that steers too aggressively or offers too narrow a menu lowers the employer's expected value from broker use, making the employer more likely to forgo broker services and purchase directly. Thus, the model preserves the main economic discipline relevant for steering and menu breadth while abstracting from direct broker-to-broker competition. The cost of this abstraction is that the model may understate the disciplining forces on steering and network design if direct broker-to-broker competition is stronger than the evidence above suggests. In this case, our counterfactuals may overstate the welfare losses from these frictions. Our findings are robust to alternative specifications where we allow the attractiveness of offering insurance without a broker to increase exogenously or to rise with the number of brokers available in the market.

5.4. Brokers' Choice of Insurer Menu

Brokers form expectations over the distribution of employers' demand shocks ξ, insurers' variable cost shocks ω, and employers' hassle cost shocks H. Brokers choose menus at the same stage, and the resulting collection of broker-insurer links determines the network structure of the industry.

The expected commission income from menu m is

$$\Pi_{bm} = \mathbb{E}_{\xi,\omega,H}\Big[\sum_{f \in \mathcal{F}_b} \underbrace{\mathbb{1}_{fb}}_{\substack{\text{whether employer} \\ \text{uses broker}}} \cdot \sum_{j \in m} \underbrace{s_{fjm} \cdot \text{takeup}_{fjm}}_{\substack{\text{demand from} \\ \text{employer } f\text{, plan } j}} \cdot \underbrace{p_{fjm} \cdot \gamma_{fj}}_{\substack{\text{commission} \\ \text{income}}} \Big], \tag{13}$$

where $\mathcal{F}_b$ is the set of employers in broker b's market.

Brokers simultaneously choose a menu to maximize expected profit, which equals their commission income minus contracting costs.

$$m = \text{argmax}\Big(\underbrace{\Pi_{bm}}_{\text{commission income}} - \underbrace{K_{bm}}_{\text{contracting costs}} \Big) \tag{14}$$

Equation (13) and (14) reveal several factors that affect the brokers' contracting decisions. First, higher commissions from an insurer, γ_{fj}, increase the benefit of including that insurer. Second, a broader menu may attract more employers to use the broker, thereby increasing the probability $\mathbb{1}_{fb} = 1$ and expected income. However, a broader menu also incurs higher contracting costs, namely, a larger K_{bm}. Third, brokers take into account that their menu decision affects insurer pricing p_{fjm} and consumer choices $s_{fjm}, \text{takeup}_{fjm}$. Equation (14) thus yields broker-specific menus m, which are functions of insurers' commission rates, pricing power, product characteristics, and the preferences of employees and employers on the market.

We focus on pure-strategy broker menus. In counterfactuals, brokers choose the menu that maximizes expected profit given commission schedules and the pricing equilibrium induced by each candidate menu.

5.5. Equilibrium

We solve the finite-stage game by backward induction. Because the Nash-Bertrand pricing subgame can admit multiple equilibria, the object we compute is a selected subgame-perfect Nash equilibrium. Formally, a selected equilibrium of this model consists of broker menus m_b^*, and, for each realization of the residual demand, marginal costs, and hassle costs (ξ, ω, H), insurer premiums p_{fjm}^*, employer broker-use decisions $\mathbb{1}_{fb}^*$, employer insurer-choice probabilities s_{fjm}^*, and employee take-up rates takeup_{fjm}^* such that: each broker's menu maximizes expected profit in equation (14); each insurer's premium satisfies the first-order condition in equation (9); each employers' broker-use decision satisfies the threshold rule in equation (12); each employer's insurer choice satisfies the choice rule in equation (7); and each employee's take-up satisfies the choice rule in equation (A7). Moreover, brokers choose menus ex ante, forming expectations over shocks and the continuation equilibrium each menu induces, whereas premiums, broker use, and choices are determined ex post.

The continuation game has a unique outcome at every stage but pricing. The choice and broker-use

stages are unique by standard arguments.[23] The menu stage also does not create strategic multiplicity, because it is separable across brokers, which follows from the employer-broker matching assumption in Section 5.3: premiums are menu-specific, and each employer meets at most one broker, so a broker's profit does not depend on any other broker's menu, reducing its problem to an independent single-agent maximization over a finite set of menus. A maximizer always exists and is generically unique; we resolve the knife-edge case in which two menus yield equal expected profit with a pre-specified tie-breaking rule.[24] Genuine multiplicity is confined to pricing, where the premium feeds back into both employer demand and insurer cost: it enters employer utility directly and through employee surplus W and take-up, and it shifts marginal cost through the effect of take-up on the enrolled pool. These make insurers' best responses nonlinear, so the pricing first-order conditions need not have a unique solution. Appendix D4 describes how we map the equilibrium set and select one equilibrium per market.

6. Estimation, Identification and Results

Our estimation proceeds in four stages, each recovering a distinct set of model primitives: employers' preferences over prices and employee welfare, insurers' marginal costs, employers' hassle costs, and brokers' contracting costs.

6.1. Employers' Choice of Insurer

Key primitives include employers'weights for their own expenses α_0 (non-brokered baseline), $\alpha_{1,b(f)}$ (broker-status wedge), and $\alpha_{2,b(f)}$ (broker steering response to commissions); weights for employee surplus, ρ; and brand perceptions, δ_j^f.

Estimation. We estimate the conditional logit model in equation (7) by maximum likelihood. This is equivalent to including employer fixed effects in flow utility, which implicitly controls for time-invariant preference shocks by market, industry, and firm size.

We draw on two additional data inputs. First, because the Form 5500 data reports premiums only for the realized insurance choice, we assume that employers of the same type in the same market, facing the same insurer and broker menu, will get the same price quote. Appendix B1 describes how we compute premiums for each alternative insurer. Second, we compute the average welfare of employees for each alternative insurer $\hat{\mathcal{W}}_{fjm}$ in the product menu offered by their observed broker, the empirical analog of $\mathcal{W}_{fjm}$ constructed from the estimated worker demand model, the imputed premium, and product characteristics. See Appendix D1 for details.

Identification. All three components of the premium weight α_{fj} in equation (8) multiply the employer-paid premium and share a common endogeneity concern: insurers' knowledge of employers' unobserved preferences when setting prices creates a correlation between residuals and prices. To address this concern,

[23] Given premiums and menus, employer and employee choices are logit and broker use is a threshold in the normally distributed hassle cost.

[24] We select the menu closest to the observed status-quo menu, measured by Hamming distance over broker-insurer links.

we use a control function with Hausman instruments (Nevo, 2001): cost changes move an insurer's premiums across all markets, while demand shocks are local. The first stage regresses premiums on leave-out prices in other markets and controls; the estimated residual is included in the second stage choice equation. The identifying assumption is that residual demand shocks in other markets are mean-independent of those in the focal market, conditional on employer fixed effects that control for time-invariant shocks by market, industry, and firm size, and insurer fixed effects. Given this correction, the three components of the premium weight are separated by three distinct sources of variation.

The baseline weight α_0 is identified by non-brokered employers' choices: conditional on employee WTP, insurer effects, and other controls, how strongly their choices shift away from an insurer as its premium rises pins down α_0. We interpret α_0 as the revealed weight of non-brokered employers, not as the counterfactual preference of all employers absent brokers.

The broker-status wedge $\alpha_{1,b(f)}$ is identified by the difference in premium responsiveness between brokered and non-brokered employers. We interpret $\alpha_{1,b(f)}$ as a reduced-form wedge that absorbs non-commission differences in choice behavior, including selection into broker use, information provision, and administrative support; but not the causal effect of brokers on choice quality.

The identification of the steering coefficient, $\alpha_{2,b(f)}$, exploits discontinuities in commission schedules within insurers, similar to the intuition in Section 4. Consider two brokered employers just below and above a commission cutoff. Both face the same insurer at the same premium; the only difference is the commission the broker earns from placing them there. If commissions did not enter the choice rule, the two employers would choose that insurer at the same rate; if brokers steer, their choice rates differ. As commissions enter through the premium term, this difference reflects a shift in how employers weigh premiums across the cutoff when commissions change, precisely what $\alpha_{2,b(f)}$ measures.

The correlation between predicted employee WTP and choice patterns identifies the weight on employee welfare, ρ. Conditional on premiums and insurer fixed effects, the remaining variation in $\mathcal{W}_{fjm}$ comes from differences in plan attributes and from how well each plan fits the employer's workforce (Appendix D1). The identifying assumption is that this remaining variation is uncorrelated with the unobserved employer–insurer match quality ξ_{fj}.[25]

Finally, δ_j^f are recovered as insurer fixed effects, identified by insurers' average choice shares conditional on premiums and employee WTP, with utility levels normalized against the no-offer option.

Results. Table 4 reports estimates of employers' preferences and Table A2 reports the first stage. Employers draw utility from higher employee surplus and incur disutility from their share of the total premium paid. Brokered employers weigh premiums less than non-brokered employers, and higher commissions further attenuate that weight. The effective premium coefficient nonetheless remains negative throughout the observed range of average commissions. The coefficient on the control-function residual is positive, indicating that the endogenous component of premiums is positively correlated with unobserved utility: insurers tend to charge higher premiums in employer-type or market cells where they are also more attractive for reasons not fully captured by $\mathcal{W}_{fjm}$, δ_j^f, and observed controls.

[25] We instrument only for premiums. $\mathcal{W}_{fjm}$ integrates out employees' idiosyncratic preference shocks, so it introduces no employee-level unobservable into the employer choice equation, and its dependence on premiums is addressed by the control

Table 4. Estimates of Employers' Preferences

Realized Choice	Coefficient	Standard Error
WTP per employee, ρ	0.0175	0.0033
Premium per employee, α_0	-0.0133	0.0023
has broker $\times$ premium, $\alpha_{1,b(f)}$	0.0020	0.0010
has broker $\times$ premium $\times$ effective commission, $\alpha_{2,b(f)}$	0.0198	0.0062
Control function for premium	0.0133	0.0023

Notes: This table reports control-function conditional-logit estimates of employers' insurer choice, equation (7) in a conditional logit model. The sample includes fully-insured New York employers with at least 100 employees, 2013–2019. Each observation is an insurer alternative in an employer's annual choice set; we include a no-offer option for employers without a broker, and the dependent variable indicates the realized choice. WTP per employee is predicted employee willingness to pay, constructed in Appendix D1; premium per employee is the employer-paid premium; both are in annual dollars per employee. Effective commission is the insurer's effective average commission rate. All specifications include insurer fixed effects. The control function is the estimated residual from the first-stage regression of premiums on the leave-out (Hausman) premium instrument and controls.

The implied price elasticities are sensible. On the offer margin, for firms not using a broker and hence given the option not to offer insurance, a 10% increase in all premiums reduces the probability of offering insurance by 15.2 percentage points (pp), from a baseline of 52%. At the insurer level, we report two own-price elasticities. The elasticity of employers' insurer choice averages -4.64 for brokered and -6.03 for non-brokered employers. Combining workers' take-up, the elasticity of the insurer's covered enrollment, the object relevant for insurer pricing, averages -5.13 and -6.47; for non-brokered employers, both objects reflect movement to the no-offer option. The small gap between the two indicates that the insurer-level price response operates mainly through employers' choice of insurer, the margin on which brokers act. Figure A17 reports both distributions at the status quo equilibrium.

Turning to broker intermediation, the positive broker–premium interaction implies that brokered employers are modestly less responsive to premiums, even before accounting for commissions. This wedge, $\alpha_{1,b(f)}$, is descriptive, as it may reflect selection into broker use or non-commission services brokers provide.

The positive commission–premium interaction, which identifies the steering coefficient $\alpha_{2,b(f)}$, implies that higher average commissions make brokered employers less responsive to the commission-paying insurer's premium. The implied steering effect is economically meaningful. Holding premiums and broker menus fixed, a 1 pp reduction in an insurer's effective commission rate lowers its choice probability among brokered employers by 1.77pp on average, and by 3.03pp for a prominent insurer. This provides a structural counterpart to the reduced-form evidence in Section 4: both show that lower commissions reduce the probability that a brokered employer selects the affected insurer.

We further compare the weight on employee WTP ρ with the effective premium weight α_{fj}, expressed as the premium an employer pays per additional dollar of employee WTP. Under a benchmark with full wage offsets and no tax wedge, equal marginal weights would be efficient: if the employer offers more generous plans than the average employee values, it could reduce coverage levels and compensate employees with higher wages, thereby improving employees' total surplus. We report deviations as revealed weights, not necessarily distortions.

Non-brokered employers pay, on average, \$1.32 in insurance premiums for \$1 in value of employee surplus. In other words, employers "overweight" health insurance relative to average employee valuations.

function: conditional on the first-stage residual, premiums and any function of them are exogenous.

This could be rationalized by the tax treatment of ESI (Gruber, 2010) or other employer objectives; a ratio of 1.32 implies an effective subsidy of roughly 24 percent, in line with standard estimates of the marginal tax wedge on premiums. For brokered employers, the ratio is \$1.56 at zero commission and \$1.65 at observed commission rates, consistent with estimates in Tilipman (2022). The wedge implies that firms with brokers purchase more expensive coverage than average employee valuations alone would justify: they pay about 25% more in premiums to generate the same dollar of employee value as non-broker firms, and 27% more at the highest observed commissions. The estimated commission channel accounts for approximately 28% of the difference between brokered and non-brokered employers.

6.2. Insurer Pricing

The key primitive in this stage is the marginal cost, mc_{fj}, by insurer, employer type, and geographic market.

Estimation and Identification. We recover marginal costs by inverting insurers' pricing first-order conditions (equation (11)). Each premium is associated with a single first-order condition, so the system is just-identified: given the estimated demand elasticities and the implied markups, the first-order conditions define a one-to-one mapping from observed premiums to marginal costs, which we invert directly (Berry, 1994; Berry et al., 1995). We then project the inverted marginal costs onto insurer, year, region, industry, and firm-size fixed effects by ordinary least squares to obtain smoothed cost estimates.

For the equilibrium simulations, we map these employer-type cost estimates into insurer-by-worker-type cost primitives. Appendix D2 describes this mapping and its calibration. The insurer's average marginal cost is the take-up-weighted average cost of the workers who enroll. Consequently, changes in premiums can change average marginal costs by shifting enrollment across worker types.

Results. Table 5 reports estimated marginal costs and premiums per enrollee, the unit at which insurers set prices. Because commissions are netted from insurer revenue in equation (9), these are claims costs alone.

Recovered marginal costs average \$383 per enrollee per month. This level lines up with insurers' filed claims: New York large-group Medical Loss Ratio (MLR) filings for 2013–2019 report incurred claims of \$375 per covered life per month. The model-implied marginal cost-to-premium ratio, which does not depend on the unit, is approximately 79 percent net of commissions and 82 percent including them, consistent with MLR requirements.[26] To facilitate comparison with other estimates in the paper, the final two columns express marginal costs and premiums per eligible employee. These amounts include coverage for any enrolled dependents and are multiplied by employee take-up. On this basis, marginal costs average \$607 per eligible employee per month.

The heterogeneity of the estimates is sensible. First, costs vary nearly twofold across insurers. Margins are highest for carriers with dominant local positions. Second, downstate markets are more expensive: New York City and the Mid-Hudson region incur the highest claims costs, and Upstate the lowest. Third, costs vary with workforce composition across industries. Fourth, larger employers have lower claims costs than smaller employers.

[26]The statutory MLR adds quality-improvement expenses to the numerator and excludes taxes and fees from the denominator, so raw claims shares run below the 85% large-group floor; incurred claims average 83% of premium in the filings themselves.

Table 5. Marginal Cost Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
	Per enrolled member per month				Per eligible employee per month	
	MC	P	MC/P	$(MC+\gamma P)/P$	MC	P
All	383	481	79.33	81.88	607	762
(a). By Insurer						
Insurer A	405	531	76.32	79.60	652	854
Insurer B	274	353	77.74	77.89	439	564
Insurer C	434	535	81.08	83.09	687	847
Insurer D	398	485	81.73	83.17	629	768
Insurer E	460	560	82.07	85.37	747	909
Insurer F	300	387	77.46	79.42	483	623
Insurer G	254	333	75.84	78.94	399	525
Insurer H	339	433	78.30	81.49	504	643
Insurer I	320	417	76.44	79.13	487	635
Insurer J	282	367	76.88	77.94	446	581
(b). By Region						
Capital	277	364	75.91	78.18	443	583
Central, Finger, Mohawk, Southern	266	348	76.09	78.69	424	556
Mid Hudson	393	489	80.33	82.76	625	777
Long Island	376	476	78.88	81.42	586	743
NYC	421	524	80.26	82.93	668	830
Western	283	364	77.75	79.09	449	578
North	229	311	73.43	73.43	358	488
(c). By Industry						
Natural resources	253	311	81.03	82.85	510	627
Mining	269	351	76.39	79.32	452	589
Construction	374	470	79.13	81.85	606	762
Manufacturing	351	444	78.72	81.34	566	715
Wholesale and retail trade	347	434	79.62	81.94	570	714
Transportation and utilities	394	491	80.04	82.02	617	768
Information	417	518	80.26	82.80	637	792
Financial activities	472	586	80.42	83.10	737	914
Professional and business services	397	502	78.95	81.56	612	773
Education, health, and social services	348	441	78.72	81.31	560	708
Leisure and hospitality	356	448	79.36	81.78	570	717
Other services	403	502	80.00	82.55	641	799
(d). By Employer Size						
100-499 employees	388	487	79.38	81.94	615	771
500+ employees	346	437	78.95	81.36	549	692

Notes: This table reports estimated marginal costs (MC) and average observed premiums (P) in dollars, and cost-to-premium ratios in percent, by insurer, rating region, industry, and employer size, pooled over 2013–2019. Commission payments are netted from insurer revenue in the inversion, so marginal costs in Column (1) measure claims costs exclusive of distribution expenses. Column (3) reports marginal cost as a share of premium; Column (4) adds commission payments, i.e., the commission rate times the premium, plus flat fees to the numerator. Costs are evaluated at each cell's average demographic composition. Entries are means across insurer–market–employer-type cells, weighted by employer mass; because the ratio columns average cell-level ratios, they need not equal the ratio of the reported level columns. Cells implying cost-to-premium ratios below 0.7 (about 10 percent of observations) are excluded. The Central New York row pools the Central New York, Finger Lakes, Mohawk Valley, and Southern Tier rating regions.

The model further implies that broader broker menus reduce insurer markups. Within the model's solved menu space, the take-up-weighted price-cost ratio declines from 1.55 for menus with one insurer to 1.26 for menus with six insurers (Table A3). The pattern is nearly unchanged when worker composition is held fixed,

indicating that it reflects stronger price competition rather than changes in enrollee selection. This pattern is consistent with the reduced-form evidence in Section 4: broader menus strengthen insurer competition.

6.3. Employers' Choice of Whether to Use a Broker

The key primitives in this stage are the mean and dispersion of H_f in equation (12).

Estimation. We parameterize $\mu_f = \theta X_f$, where X_f includes firm-size, industry, region, and year fixed effects.[27] Let $\Delta V_f = V_{fb}^{wb} - V_f^{nb}$ denote the expected value difference between using a broker and purchasing directly, before accounting for H_f. For broker users, we compute V_f^{wb} under the menu of their observed broker. For nonusers, the broker they would have met is unobserved. Consistent with the local matching process, we approximate V_f^{wb} using the equally weighted average value across brokers operating in the employer's county in the specific year. We use the same aggregation of candidate-broker instruments for nonusers. For both groups, V_f^{nb} is evaluated under the full set of insurers available in the market, including the option of not offering insurance. An employer uses a broker when $\Delta V_f + H_f > 0$.

We estimate this decision rule by a maximum likelihood IV-probit model, with a linear first stage for ΔV_f on the instrument described below and X_f. In the normalized choice equation, the coefficient on ΔV_f is $\beta_V = 1/\sigma$, and the coefficients on X_f are $\beta_X = \theta/\sigma$. Their ratios therefore recover θ, the parameters governing the mean of H_f. We report the estimated conditional mean, $\widehat{\mu}_f = \hat{\theta} X_f$.

Identification. ΔV_f may be endogenous under the model's local matching assumption: brokers choose menus in response to the demand of prospective clients. Selective matching with a particular broker would create an additional concern if random matching is only an approximation to the data. Either force can generate a correlation between ΔV_f and the unobserved component of H_f.

We instrument for ΔV_f using the value that the relevant broker generates for its other clients in other counties, leaving out the focal county. A broker's menu serves clients across counties, so its value elsewhere predicts the value of the brokered choice for the focal employer. The identifying assumption is that, conditional on X_f, the menu value that a broker generates for its other clients is uncorrelated with the focal employer's unobserved propensity to use a broker. The instrument may affect the focal employer's broker-use decision through the employer's own relative menu value, ΔV_f, but not through other determinants of H_f.

Under this restriction, differences in broker use across employer characteristics, conditional on ΔV_f, identify θ/σ. The response to instrument-induced variation in ΔV_f identifies $1/\sigma$. Together, these identify the mean and dispersion of H_f.

Results. Table 6 reports the estimates in dollars per eligible employee per month. The average estimate is \$68: employers behave as if buying insurance with a broker provides \$68 per employee per month beyond any difference in the value of the menu.

The mean hassle cost equals 12.5% of the average employer-paid premium, and 2.8 times the average broker commission. The value employers place on avoiding the hassle of purchasing directly substantially exceeds what brokers are paid for placing their coverage. For a firm with 200 eligible employees, the hassle

[27] For notational simplicity, we drop the time-dimension subscript t throughout.

Table 6. Hassle Cost Estimates, Per Eligible Employee Per Month

	Hassle Cost ($ PMPM)	SE
All employers	68.47	33.52
(a). By Industry		
Construction	119.23	48.35
Manufacturing	111.18	34.64
Wholesale and retail trade	48.83	24.19
Transportation and utilities	28.44	21.87
Information	37.75	40.75
Financial activities	41.19	43.64
Professional and business services	60.73	37.41
Education, health, and social services	86.13	32.12
Leisure and hospitality	45.27	33.03
Other services	45.03	35.08
(b). By Employer Size		
100–499 employees	68.65	33.37
500+ employees	67.02	35.21
(c). By Region		
Capital	105.42	34.41
Central/Finger/Mohawk/Southern	146.03	38.21
Mid Hudson	38.79	28.19
Long Island	39.59	35.48
NYC	31.86	46.65
Western	82.09	37.22
North Country	112.38	41.19

Notes: Estimates are hassle costs per eligible employee per month: each coefficient divided by the coefficient on V_{diff}. Standard errors are derived from the delta method.

cost amounts to about $165,000 per year. This is roughly the salary of one benefits manager, which is about $150,000 in New York City, or two benefits specialists, about $80,000 each.

Estimated hassle costs vary with employer characteristics. They decrease slightly with employer size: $69 for employers with 100–499 employees and $67 for those with 500 or more. Across industries, they range from $28 in transportation and utilities and $38 in information to $111 in manufacturing and $119 in construction. Across regions, they are lower in areas around New York City and higher in upstate New York, consistent with scarcer benefits-administration expertise outside the New York City area.

We interpret the estimated H_f broadly. It captures all benefits of using a broker other than its effect on the insurer menu, including avoided search costs, administrative cost savings, stable preferences for broker services, and the benefits of established relationships. Microfounding these components separately is beyond the scope of this paper, and the estimation does not distinguish among them. To assess how much of the estimate reflects persistence in broker use rather than the cost of arranging insurance without a broker, Table A4 adds lagged broker use, allows for persistent unobserved employer preferences, and restricts to employers observed both with and without a broker.

We further compare the value of broker services with steering distortions brokers induce. Using demand estimates, we compare the plans chosen by brokered employers with those they would choose absent broker influence, holding their menus fixed and allowing insurers to reprice. At the average commission rate, steering raises premiums by 2.9% for the same employee surplus level. This is equivalent to a $17.4 PMPM loss in employer surplus. The avoided hassle cost is roughly four times this loss. For a firm with 200 eligible

employees, the value of broker services exceeds their cost by approximately $112,000 per year.

6.4. Brokers' Choice of Insurer Menu

The key primitives in this stage are insurer-specific contracting costs.

Estimation and Identification. Estimating broker network formation presents two challenges. First, the number of possible network configurations increases exponentially with the number of insurers. Second, insurers are interdependent. That is, the revenue from adding an extra insurer depends on the identity of insurers in the existing portfolio. Instead of solving the industry equilibrium directly, we exploit the inequality restrictions implied by brokers' revealed preferences. A broker's observed menu must be at least as profitable as any menu that adds or drops one insurer.

We assume the total contracting costs of the broker menu m is additive across the contracting costs of each included insurer.

$$K_{bm} = \sum_{j \in m} K_{bmj} = \sum_{j \in m} (K_j + \nu^1_{bjm} + \nu^2_{bjm}).$$

We interpret K_j as the mean annual fixed cost of maintaining a relationship with insurer j, incurred for both new and continuing contracts. These costs can reflect recurring tasks such as learning about changes in insurers' products and procedures. We do not separately identify underlying cost components.

The two error terms capture different sources of unobserved variation. The expectational error ν^1_{bjm} is the difference between model-computed incremental commission income and the income the broker expects when choosing its menu (Dickstein and Morales, 2013; Prager and Tilipman, 2025). Because the revenue comparisons use model-computed income, we absorb this difference into K_{bmj}. $\nu^1_{bjm}$is unknown to brokers when they make their network decisions and satisfies $\mathbb{E}[\nu^1_{bjm} \mid \mathcal{I}_{bm}] = 0$, where $\mathcal{I}_{bm}$ is the broker's information set when choosing its menu. This error rationalizes cases in which the model predicts that a broker should carry an insurer but the broker does not.

Following Pakes et al. (2015), ν^2_{bjm} is a structural error term known to brokers at menu decisions but unknown to the econometrician. ν^2_{bjm} allows contracting costs to vary across markets and brokers in ways not otherwise captured by this parameterization.

We form revealed preference inequalities from the brokers' observed menu decisions that provide upper and lower bounds on contracting costs. If a broker contracts with an insurer, its contracting cost must be weakly lower than the commission income from including that insurer in the menu. This delivers an upper bound on the contracting cost,

$$\Delta\Pi_{bm,j^-} = \Pi_b(m) - \Pi_b(m \setminus j) \geq K_{bmj}.$$

Conversely, if a broker does not contract with an insurer, the contracting cost must be weakly higher than the counterfactual commission income from including that insurer, $m^* \cup j$. This generates a lower bound for the contracting cost.

$$\Delta\Pi_{bm,j^+} = \Pi_b(m \cup j) - \Pi_b(m) \leq K_{bmj}.$$

The presence of ν^2_{bjm} introduces selection: brokers that receive high draws of ν^2_{bjm} will be less likely

to contract with insurer j. We make two assumptions to address this selection problem (Eizenberg, 2014; Canay et al., 2023; Geddes, 2022).

First, contracting costs are nonnegative and have finite insurer-specific bounds: $0 \leq \underline{K}_j \leq K_{bmj} \leq \overline{K}_j < \infty$. Second, these bounds lie within the support of expected incremental commission income associated with adding or retaining insurer j: $[\underline{K}_j, \overline{K}_j] \subseteq [\underline{\Delta\Pi}_j, \overline{\Delta\Pi}_j]$. The support is taken over the income lost by dropping j, for brokers that carry it, and the income gained by adding j, for brokers that do not, i.e., $\{\Delta\Pi_{bm,j^-} : j \in m\} \cup \{\Delta\Pi_{bm,j^+} : j \notin m\}$, and $\underline{\Delta\Pi}_j$ and $\overline{\Delta\Pi}_j$ are its lower and upper endpoints. This requires that any broker would weakly contract with insurer j if that contract were as profitable as it is for the most profitable broker in the most profitable market and year. Unlike the cost support, these revenue endpoints are identified from the demand model.

Under these assumptions, the completed bounds become:

$$L_{bjm} = \begin{cases} \max\{\underline{\Delta\Pi}_j, 0\}, & j \in m, \\ \max\{\Delta\Pi_{bm,j^+}, 0\}, & j \notin m, \end{cases} \qquad U_{bjm} = \begin{cases} \Delta\Pi_{bm,j^-}, & j \in m, \\ \overline{\Delta\Pi}_j, & j \notin m. \end{cases}$$

For a contracted pair, the deletion comparison supplies the upper bound and the support restriction the lower bound. For an uncontracted pair, the addition comparison supplies the lower bound and the support restriction the upper bound.

We average these completed restrictions within insurer–region (jr) cells. We assume the conditional mean-zero assumption on expectational errors applies within these cells. Together with the assumptions on structural errors, this yields

$$\mathbb{E}[L_{bjm} - K_j \mid j, r] \leq 0, \qquad \mathbb{E}[K_j - U_{bjm} \mid j, r] \leq 0.$$

Because each expectation includes both contracted and uncontracted pairs, we do not require unobserved costs to have mean zero conditional on contracting status.

A single K_j must satisfy inequalities in every region. Regional variation in incremental commission income therefore supplies distinct bounds on the same cost parameter. Keeping these restrictions separate prevents slack in one region from offsetting a violation in another. The identified set is

$$\max_{r \in \mathcal{R}_j} \mathbb{E}[L_{bjm} \mid j, r] \leq K_j \leq \min_{r \in \mathcal{R}_j} \mathbb{E}[U_{bjm} \mid j, r], \tag{15}$$

where $\mathcal{R}_j$ denotes the eligible regions for insurer j. We estimate the regional expectations using the corresponding sample averages. We construct marginal 95% confidence intervals for each K_j with the procedure of Andrews and Soares (2010).

Results. Table 7 reports the estimated bounds on contracting costs, their midpoints, and marginal 95% confidence intervals.

For a single insurer, the midpoint estimates range from \$387 to \$2,296 per potential client firm per month, equivalent to \$1.36–\$8.07 per employee per month. The estimates display substantial heterogeneity across insurers: four of the ten midpoints fall below \$500, while three exceed \$1,500. Thus, contracting costs depend substantially on the composition of a broker's menu, in addition to the number of insurers it

Table 7. Contracting Cost Estimates by Insurer

Insurer	Midpoint	[lb, ub]	95% CI
Insurer A	1,642	[945, 2,339]	[876, 2,500]
Insurer B	2,220	[1,897, 2,542]	[1,826, 2,605]
Insurer C	857	[430, 1,284]	[320, 1,508]
Insurer D	487	[250, 724]	[145, 990]
Insurer E	983	[462, 1,504]	[355, 1,812]
Insurer F	486	[354, 618]	[204, 750]
Insurer G	387	[260, 514]	[180, 718]
Insurer H	390	[179, 602]	[139, 762]
Insurer I	1,231	[678, 1,784]	[561, 2,188]
Insurer J	2,296	[1,738, 2,853]	[1,616, 2,971]

Notes: This table reports estimates of brokers' mean contracting costs, K_j, in dollars per potential client firm per month, by insurer, pooled over 2013-2019. The sample includes brokers serving fully insured New York employers with at least 100 employees. The identified set is estimated from the insurer-by-region moment inequalities in equation (15). The 95% CI is the marginal confidence interval from the generalized moment selection procedure of Andrews and Soares (2010).

contains. We use the midpoint costs in the baseline counterfactuals and examine sensitivity to the lower and upper cost endpoints and to cost heterogeneity across brokers.

Summed across the insurers in brokers' observed menus, contracting costs average $3,569 per potential client firm per month, or $146,000 per year for a broker. Expected commission revenue averages $215,264 per broker per year, and profit averages $69,646. The implied margin, 32.4% of commission revenue, is consistent with a published retail brokerage margin of 28.4% (Brown and Brown, 2020).

6.5. Model Fit

We assess model fit in two steps in Table A5, reflecting the model's identification structure. Demand parameters, insurers' marginal costs, and hassle costs are point identified conditional on observed broker menus, whereas brokers' contracting costs are set identified. We first hold menus at their observed values and simulate employer choices and equilibrium premiums. This evaluates fit without selecting values for contracting costs. We then allow brokers to choose their menus, using midpoint contracting costs and broker-specific cost draws. This second exercise assesses the full equilibrium under the contracting-cost specification used in the counterfactuals.

At observed menus, the model fits aggregate employer choices and premiums reasonably well. Predicted broker use is 81.3%, compared with 85.1% in the data. Coverage offering and take-up among offering employers also closely match their observed levels. The model also closely matches observed choice patterns for both brokered and non-brokered employers. Average monthly premiums are $532 per enrollee, compared with $507 in the data. Simulated prices correlate with observed prices at 0.759 for employers with brokers and 0.984 for those without.

The full equilibrium also fits aggregate premiums, take-up, and insurer market shares closely. Take-up among offering employers is 75.6%, compared with 75.5% in the data, and each insurer's aggregate market share is within 1.3 percentage points of its observed value. Simulated menus are narrower, and broker use is lower than in the data. The simulated menus nonetheless carry the right insurers: the share of brokered employers whose broker carries each insurer correlates with its observed value at 0.95. We use this simulated status quo as the benchmark for the full-equilibrium counterfactuals in Section 7. Policy effects therefore

compare equilibria generated under the same contracting-cost specification and solution procedure.

We also examine whether the demand model captures the reduced-form response to commission incentives. Reestimating the regression kink design in Section 4 on the demand-estimation sample yields a kink of 2.60 percentage points per ten enrolled employees. In a specification that additionally allows an Insurer 1-specific steering response, the model-implied kink is 1.41. The model thus captures the direction and a substantial share of the reduced-form response magnitude. The smaller magnitude reflects differences in how the two approaches use the data. Whereas the RKD uses brokered employers near the commission threshold, the demand model estimates common insurer effects and price sensitivity using both brokered and nonbrokered employers over a broader range of firm sizes.

7. Counterfactuals

This section uses the estimated model to quantify the steering and competitive-access distortions and to evaluate policies that regulate broker compensation. Section 7.1 examines commission caps, and Section 7.2 compares caps with alternative policies.

All simulations use the year 2018. We report dollar amounts per eligible employee per month unless noted. We evaluate employer surplus at employers' preferences net of the commission term: for brokered employers, we consider the broker-status shift α_1 as a welfare-relevant component of preferences, whereas we consider the commission term $\alpha_2\gamma$ as a distortion. Employer surplus is inclusive of hassle costs and is measured relative to not offering coverage, so its baseline level can be close to zero or negative; we therefore report changes in dollars rather than in percentage terms. Insurer pricing stays on the lowest-price equilibrium in every simulation (Appendix D4).

7.1. Commission Caps

Commission caps are a common regulatory response to conflicted advice. Before 2010, New York capped broker commissions at 4 percent for HMO products and the in-network portion of point-of-service products, and CMS caps broker compensation in Medicare Advantage (see Section 2). As the highest observed commission is 4% of premiums in the data, we simulate equilibria under caps of 1, 2, and 3% on every insurer's commission schedules.

A cap affects equilibrium outcomes through three channels. First, it corrects distorted choices: with smaller commissions, brokered employers' insurer choices tilt less toward the insurers that pay the most. Second, it changes insurers' pricing incentives. Insurers pay a smaller commission on each brokered sale, and brokered demand becomes more price-elastic as steering weakens; both forces push premiums down. Third, it changes the commission income that sustains broker networks. Brokers drop insurers whose commissions no longer cover their contracting costs, and narrower menus reduce variety and soften price competition, which pushes premiums up. Because a cap also compresses commission differences across insurers, it can instead lead brokers to add insurers they previously excluded, as in the menu-expansion case in Section 3. The first two mechanisms raise employer surplus, whereas the third is ambiguous and depends on the broker menu response to the new commission rate. Thus, the effect of a cap on employer surplus is

ambiguous in sign. Employers' decisions to use a broker respond to all three mechanisms.

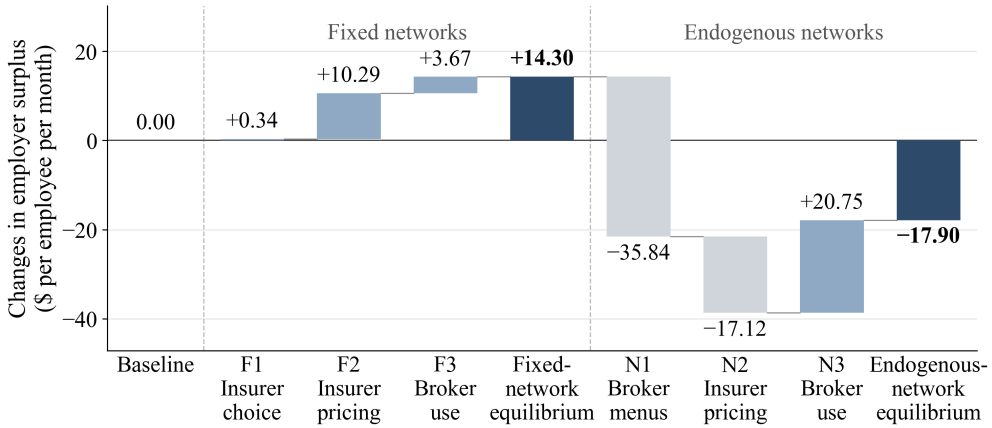
7.1.1. Decomposition.

To quantify each mechanism, we run a series of counterfactuals, sequentially introducing an additional margin of adjustment. The first three hold broker networks fixed. Counterfactual F1 lets brokered employers' insurer choices respond to the capped commissions, holding premiums, broker use, and broker menus at the status quo. F2 further lets insurers reset premiums, and F3 lets employers decide whether to use a broker. F3 is the fixed-network equilibrium. Moving from the status quo to the fixed-network equilibrium captures the gains from relieved steering and lower commission costs.

The next three counterfactuals, which we call endogenous networks, let brokers choose their menus anticipating the premiums and broker use changes that follow network changes. We add the following responses in order: N1 lets brokers change the identity of insurers included in the menu, but hold insurer premiums and employers' broker use at the fixed-network equilibrium level (F3). N2 further lets insurers reprice to the new menus, and N3 lets employers re-decide whether to use a broker. N3 is the endogenous-network equilibrium outlined in Section 5. Moving from the fixed newtork equilibrium to the endogenous-network equilibrium captures the losses from a worsened competitive-access distortion.

In order to highlight the roles of steering and competitive access, Figure 10 illustrates the decomposition of the effects of a commission cap set at 1% of premiums across different channels. Table 8 reports key equilibrium statistics underlying the decomposition.

Figure 10. Effect of a 1% Commission Cap on Employer Surplus



Notes: This figure decomposes the change in employer surplus under a 1% cap on commissions, in dollars per eligible employee per month. Steps F1–F3 add one margin of adjustment at a time, holding broker networks fixed; steps N1–N3 let brokers choose their networks and add the responses that follow, holding the set of active brokers fixed. Light bars report the change at each step, and dark bars report the cumulative change from the status quo in the fixed-network and endogenous-network equilibria. Simulations cover all markets in 2018.

Steering. Holding broker networks fixed, employer surplus rises by $14.30 per eligible employee per month. Counterfactual F1 isolates the cap's effect strictly through employers' choices. The cap lowers the average commission on brokered placements from 3.33% to 1% of premiums. Moving from the status quo to F1, employers' insurer choices align more with their preferences and employer surplus rises by $0.34, accounting for 2.5% of the fixed-network gains. We then compute F2, quantifying how insurers' price re-

Table 8. Equilibrium Objects under a 1% Commission Cap: Decomposition

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		Fixed networks			Endogenous networks			
	S0 Status quo	F1 Insurer choice	F2 Insurer pricing	F3 Broker use	N1 Broker menus	N2 Insurer pricing	N3 Broker use	Broker exit
Employers								
Employer surplus	-12.15	-11.81	-1.52	2.15	-33.69	-50.80	-30.05	-32.62
Choice value	17.41	17.75	28.04	26.12	-9.72	-26.83	18.87	29.42
Hassle costs	29.55	29.55	29.55	23.97	23.97	23.97	48.92	62.04
Broker use (%)	70.3	70.3	70.3	76.5	76.5	76.5	48.9	35.4
Not offering (%)	14.4	14.4	14.4	11.6	11.6	11.6	24.4	30.7
Employees								
Employee WTP	165.86	165.56	171.34	176.44	175.09	161.16	148.37	140.93
Take-up (%)	63.5	63.5	65.4	67.3	67.4	62.6	56.7	53.6
Insurers								
Premium per enrollee	540.0	538.9	521.3	523.7	512.5	563.6	539.6	516.5
Brokered	555.7	554.4	533.0	533.5	520.5	579.7	576.4	555.5
Markup (p/mc)	1.35	1.35	1.31	1.31	1.31	1.43	1.35	1.30
Insurer profit	150.78	162.87	147.49	154.86	149.74	182.63	132.60	111.59
Brokers								
Insurers per menu	3.13	3.13	3.13	3.15	2.15	2.15	2.43	2.44
Commission (%)	3.33	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Broker commissions	19.44	5.79	5.76	6.28	6.14	6.11	3.92	2.85
Broker contracting cost	8.03	8.03	8.03	8.03	2.98	2.98	2.98	1.83
Broker profit	11.40	-2.24	-2.28	-1.76	3.16	3.13	0.94	1.01
Share active (%)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	37.1

Notes: This table decomposes the effects of a 1% cap on commissions for all employers in 2018. Column (1) reports status-quo levels, and each remaining column reports levels after adding one margin of adjustment, in the order listed: F1–F3 hold broker menus at the status quo, and N1–N3 let brokers choose their networks, holding the set of active brokers fixed. Premium per enrollee is in dollars per covered member per month. Employer surplus, employee WTP, broker commissions, and insurer profit are in dollars per eligible employee per month. Premium per enrollee and markup are averages over all employers, over firms that use a broker, and over firms that do not; insurers per menu and the commission rate are averages over firms that use a broker; the other rows are averages over all employers. Share active is the share of brokers that remain in the market; brokers exit only in Column (8). Employer surplus is net of hassle costs.

sponses amplify the effect of the cap. Insurers pay a smaller commission on each brokered sale and face brokered employers who are more elastic with respect to premiums. As such, premiums for brokered employers fall by $22.7 per enrollee per month, a 4% decrease. Employer surplus rises by a further $10.29, which accounts for 72% of the fixed-network gain. Part of this repricing reflects passthrough of lower commission costs, and the rest reflects more elastic demand, thus more intense price competition. Next, F3 measures how employers' broker-use decisions respond. Lower premiums and less-distorted recommendations raise the value of broker services: broker use rises by 6.2 pp, and the share of employers not offering coverage falls by 2.8 pp. Employers that switch to a broker no longer bear hassle costs, and employer surplus rises further by $3.67, or 25.5% of the fixed-network gain. Notably, employee WTP also increases, as does within-firm take-up of insurance (rising from 63.5% at baseline to 67.3% with the cap). Overall, comparing the status quo with the F3 captures the net effect of commision-based steering. Removing steering incentives increases choice value for employer and employees, primarily through lower prices and markups, while raising employee take-up and hence employer surplus. Insurer profits also increase, largely driven by

the reduction in commission requirements, which effectively reduces their marginal costs.

Competitive Access. However, we also show that network adjustment could reverse this gain, lowering employer surplus by $32.20 per employee per month relative to F3. With networks fixed, the cap lowers brokers' commission income significantly, so many broker-insurer relationships no longer cover their contracting costs. N1 shows that brokers drop insurers on net: the average brokered menu shrinks from 3.15 to 2.15 insurers. Even at fixed-network premiums, employer surplus falls by $35.84 because of lost product variety. N2 shows that insurers then reprice to the narrower menus, raising premiums for brokered employers by $46.20 per enrollee (an 8.6% increase) relative to the fixed-network equilibrium, and $23.6 relative to the status quo (a 4.4% increase). This lowers employer surplus further by $17.12. The price effect from lessened competition mirrors the rise in rival premiums for exposed brokers after Insurer 2's 2016 fee cut in Section 4. Employee WTP decreases, and within-firm take-up decreases to 62.6%, even below the take-up rate in the status quo. Note that if employers are captive to their brokers, insurer profits *increase*, as employers have no counterveiling mechanism to switch carriers. However, In N3, employers re-decide whether to use a broker. The share of firms using brokers falls as employers become more willing to bear the hassle cost of purchasing insurance directly rather than pay the inflated premiums of brokered plans. Insurer profits decline to well below the fixed-network equilibrium and status quo. Although this recovers about $20.75 in employer surplus, employees ending up bearing the cost of the narrower networks and and higher premiums: the share of employers not offering insurance rises significantly, employee WTP declines by about $12.79 (8%) relative to N2 and $28.07 (16%) relative to the fixed-newtork equilibrium. Employee take-up rates also fall to 56.7%. Taken together, the difference between F3 and N3 highlights the effects of the competitive-access channel, which more than eliminate all the gains from steering reductions and impose significant penalties to employees directly.

While our main specification holds the set of active brokers fixed, the last column of Table 8 applies an exit rule to the endogenous-network equilibrium. We allow a broker to exit the market when its profit under the cap falls below the first percentile of simulated status-quo profits in the market. Their clients change to purchase without a broker or stop offering coverage; menus and premiums stay at their N3 values.[28] Under the 1% cap, 63% of brokers exit, and average employer surplus falls further by $2.57. Though this surplus change is relatively mild in aggregate, the loss is concentrated: the clients of exiting brokers end up $19.19 below the status quo, $6.05 of it from losing their broker and the rest from the narrower menus their brokers chose under the cap before exiting. Since this outcome depends critically on how such clients would re-match, we report the endogenous-network equilibrium without exit as our main outcome.

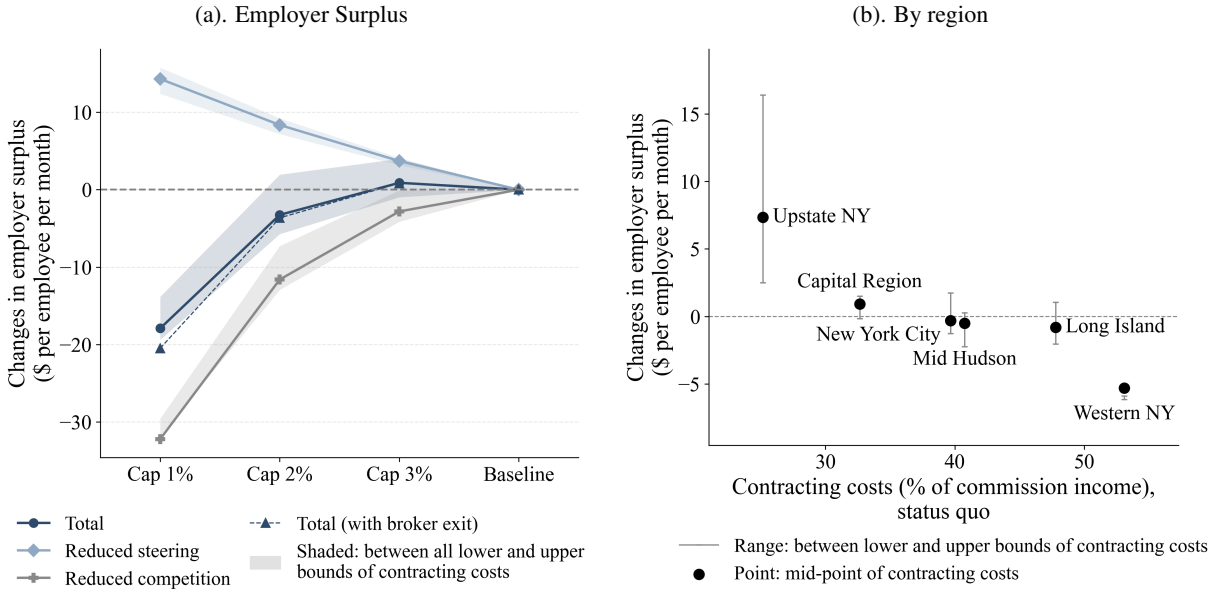
We note again that a 1% commission cap is on the extreme side of the regulatory spectrum and the effects presented above as extremely non-linear. For instance, at a cap of 2%, our model predicts that only 11% of brokers would exit the market, and more modest surplus and pricing changes overall. Below we turn to comparing changes in employer surplus at different levels of regulation.

[28] As broker-employer matches, and thus profits, are independent of one another in our model, the removal of a broker under this exit scenario does not pose any multiplicity issues that might otherwise arise if the broker's former clients were reassigned to another broker.

7.1.2. Optimal Cap.

Figure 11a plots the change in employer surplus under caps of 1, 2, and 3% of premiums, decomposed into the gain from reduced steering with networks fixed, and the loss from reduced competition as networks adjust. Table A6 reports the full set of equilibrium outcomes under alternative commission caps. We highlight three patterns.

Figure 11. Employer Surplus under Alternative Commission Caps



Notes: Panel (a) plots the change in employer surplus relative to the status quo under caps of 1, 2, and 3% of premiums, in dollars per eligible employee per month, averaged over all markets. Reduced steering is the change from the status quo to the fixed-network equilibrium; reduced competition is the change from the fixed-network equilibrium to the endogenous-network equilibrium. The total is their sum. The dashed line applies the broker exit rule described in the text. Panel (b) plots, for each market, the total change in employer surplus under the 3% cap against the share of brokers' commission income that contracting costs absorb in the status quo. Points and lines use the midpoint of each insurer's contracting-cost bounds in Table 7, with broker-insurer-specific draws around the midpoint. The shaded area in panel (a) and the range in panel (b) re-solve the equilibrium with every contracting cost at its lower bound and at its upper bound. They show the sensitivity of the results to the contracting-cost estimates and are not confidence intervals. Simulations cover all markets in 2018.

First, the gain from reduced steering rises steadily as the cap tightens, nearly quadrupling from $3.70 under the 3% cap to $14.30 under the 1% cap. A tighter cap compresses more of the commission dispersion that distorts choices and removes more commission cost from premiums, so all three fixed-network channels, including choices, prices, and broker use, strengthen. The effect of steering is fairly linear in regulation strength.

Second, the loss from reduced competition rises faster: it offsets three-quarters of the fixed-network gain under the 3% cap, more than the full gain under the 2% cap, and more than twice the gain under the 1% cap. The asymmetry reflects what a cap acts on. It curbs steering only on the placements that commissions distort, but it lowers broker income on every placement whose commission exceeds the cap. Under the 3% cap, broker commissions fall by 17% with networks fixed and the average brokered menu loses 0.2% of total insurers; under the 1% cap, commissions fall by 68%, and the menu loses 33% of total insurers.

Third, the total effect is non-monotonic in cap stringency: the 3% cap raises employer surplus by $0.89,

whereas tighter caps lower it, by $3.25 under the 2% cap and $17.90 under the 1% cap. Allowing brokers to exit lowers employer surplus further under every cap, from $0.06 under the 3% cap to $2.57 under the 1% cap. Employee willingness to pay follows the same pattern as employer surplus (Table A6), in part because tighter caps raise the share of employers not offering coverage.

Fourth, the robust lesson is the non-monotonicity itself rather than the location of the peak. The shaded bands in Figure 11a span simulations at the lower and upper bounds of the identified set of contracting costs, and across the set, employer surplus first rises and then falls as the cap tightens, with the 1% cap lowering it for every draw. Among the caps we consider, the 3% cap yields the highest employer surplus with endogenous networks and midpoints of contracting costs; because we evaluate only three caps, we do not interpret it as optimal. Within the range of commissions observed in the data, a cap can be loose enough to help employers and tight enough to result in adverse consequences.

Heterogeneity. The average effect of a cap masks heterogeneity in its sign. Table A7 reports changes by geographic market. For example, under the 3% cap, our model predicts that employer surplus slightly rises in aggregate. This is driven by increases in employer surplus in the Upstate and Capital Region. However, surplus actually falls in the remaining four markets, including New York City. This implies there is a different surplus-maximizing cap in different regions of the state. The employer-surplus-maximizing cap is 3% Upstate, 2% in the Capital Region, and status quo (no cap) in the other four markets, where any restriction yields decreases in surplus. Indeed, the gain from relieved steering is positive in every market; the differences come from how broker networks respond.

The structure of broker networks before the cap explains the differences. In Western New York, the average brokered menu already contains 83% of all available insurers, and contracting costs absorb 53% of brokers' commission income, so many broker-insurer relationships are close to the participation margin. With networks fixed, the cap lowers broker profits by 39%. With endogenous networks, brokers drop insurers, and insurers raise premiums for brokered employers by $17 per enrollee. In contrast, in the Upstate market, the average menu includes only 40% of all available insurers, and contracting costs acount for 25% of commission income. The cap narrows the commissions gap between insurers that brokers carry and those they exclude, lowering the commission income brokers forgo when clients switch to newly added insurers. As in the menu-expansion case in Section 3, upstate brokers add insurers on net: the average menu widens to 51% of all available insurers, brokered employers shift toward lower-premium insurers, and employer surplus rises with network adjustment.

Figure 11b summarizes the pattern. The effect of the 3% cap on employer surplus falls with the share of commission income that contracting costs absorb. Broker menus widen when that share is low and narrow when it is high. The pattern is robust to where contracting costs lie within their identified set.

These results underscore the welfare ambiguity. A cap trades the steering distortion against the competitive-access distortion, and which one dominates depends on the networks the cap acts on: the same cap widens networks in one market and narrows them in another. Whether a cap benefits employers is therefore an empirical question that requires modeling brokers' network choices.

7.2. Alternative Policies

The commission-cap results show that reducing steering can come at the cost of narrower broker networks. We next examine alternative policies that target three distinct margins. Uniform commission rates regulate compensation by requiring all insurers to pay the same percentage of premiums. A fiduciary duty regulates conduct by requiring brokers to recommend plans in employers' interests, similar in spirit to fiduciary obligations for investment advisers. Minimum menu requirements regulate access by specifying how many insurers brokers must carry. This intervention is motivated by the requirement that web-brokers in the ACA Marketplaces allow consumers to view all qualified health plans offered through the Exchange. We also combine a fiduciary duty with a full-network requirement to assess the gains from jointly addressing steering and restricted access. Finally, a market without brokers provides a benchmark for evaluating the value of intermediation. For all these alternative policies, we simulate the endogenous network equilibrium.

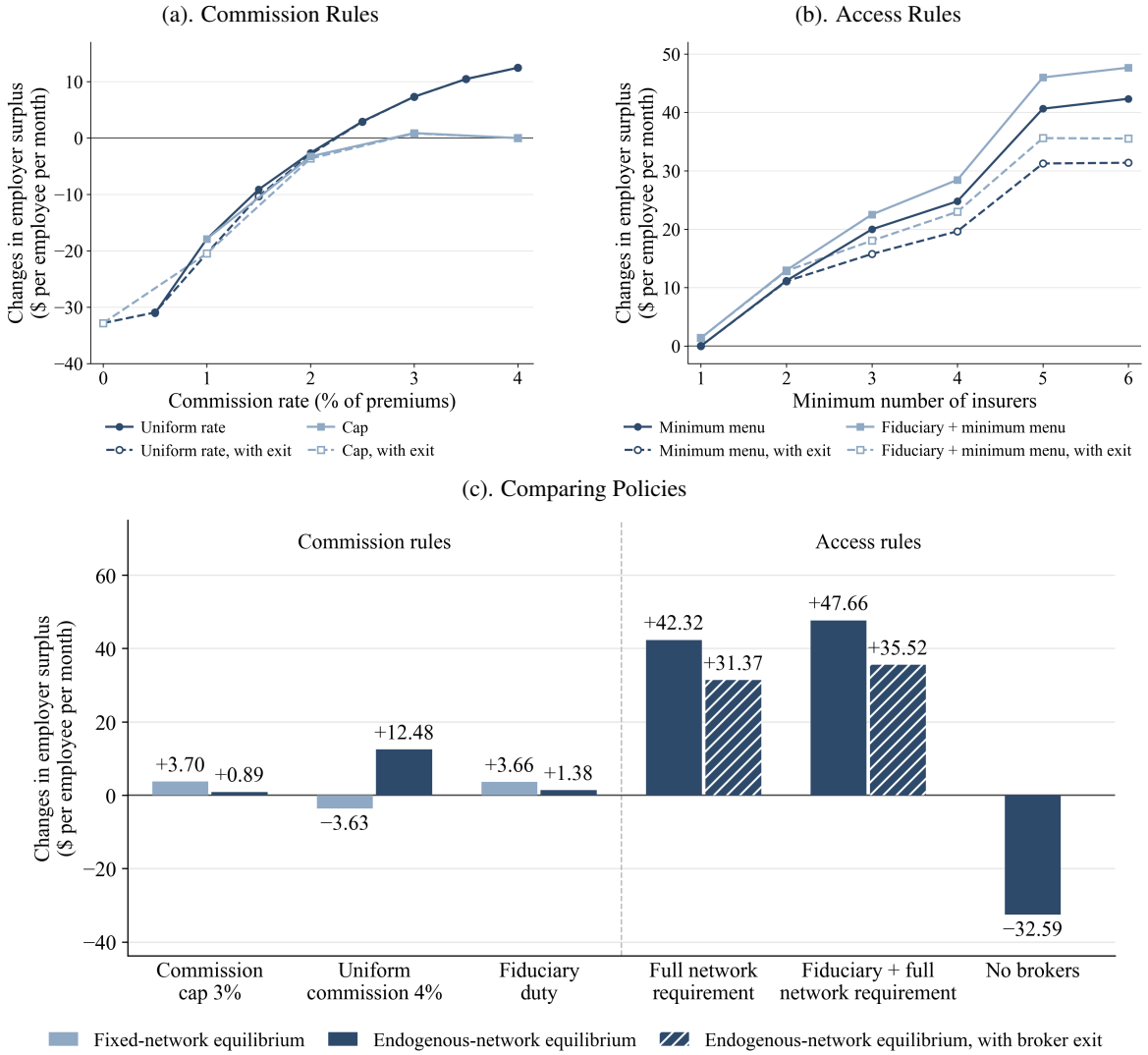
Uniform commission rate. We require all insurers to pay the same percentage commission and vary the common rate over the policy grid. With endogenous networks, the highest evaluated rate delivers the largest employer-surplus gain among the commission rules considered. Figure 12a and Table A8 report results.

First, higher common rates support broader networks and greater broker use. At low rates, commission revenue cannot cover the contracting costs of many insurer relationships, and brokers offer narrow menus. As the rate rises, more relationships become profitable: brokers carry more insurers, and broader choice attracts more employers to their services. Commission income and broker profits rise, and fewer employers forgo coverage. When we allow brokers to exit, exit is concentrated at low rates and becomes negligible at the upper end of the grid.

Second, broader networks can offset the price effects of higher commissions. Higher rates increase insurers' commission expenses and weaken brokered employers' sensitivity to premiums. The latter distortion persists because compensation remains proportional to premiums, even when its rate is the same across insurers. Broader menus work in the opposite direction by exposing insurers to more competitors: at low rates, narrow menus mute this force, and premiums exceed the status quo. As the rate rises, average premiums decline and markups fall below their status-quo level, consistent with competitive benefits of broader menus outweighing upward pricing pressures. Near the top of the grid, premiums turn upward, and markups return to the status quo, even as menus continue to expand. At the selected rate, premiums nonetheless remain below the status quo despite higher commission payments. The rate that minimizes premiums therefore differs from the rate that maximizes employer surplus.

Third, employer surplus rises with the uniform commission rate, with diminishing gains. Low rates reduce surplus relative to the status quo because employers lose access to insurers and to broker services. At higher rates, broader menus and greater broker use reverse these losses, and surplus continues to rise even after premiums turn upward, as more employers avoid the hassle costs of purchasing directly. The importance of network adjustment is apparent at the selected 4% rate: employer surplus falls by \$3.63 per eligible employee per month with networks fixed but rises by \$12.48 once networks adjust (Figure 12c). The gain is split roughly evenly between higher choice value and lower hassle costs; employers, employees, and brokers gain, while insurer profits decline slightly. Gains flatten but remain increasing at 4%, the highest commission observed in the data and New York's pre-2010 cap for HMO products, so we interpret the

Figure 12. Employer Surplus by Policy Instrument



Notes: This figure reports the change in employer surplus relative to the status quo under each policy instrument over a grid of levels, in dollars per eligible employee per month. The sample includes employers in the multi-insurer regions of New York in 2018. Panel (a) varies commission rules: caps of 1-3% and uniform rates of 0-4% of premiums. Panel (b) varies the minimum menu size from two to six insurers, with and without a fiduciary duty; we cap the minimum at the number of insurers active in each region. Panel (c) compares the policies of Table 9, each at the level that maximizes employer surplus with endogenous networks. All outcomes in panels (a) and (b) are endogenous-network equilibria.

selected rate as the best within the observed range rather than an unrestricted optimum.

Fourth, uniform rates can reduce commission dispersion while sustaining broker compensation, yielding larger employer gains than caps in our simulations. While caps narrow rate differences by cutting commissions above the ceiling, uniform rates also raise commissions paid by low-paying insurers, removing the rate disadvantage that discourages their inclusion in broker menus. Uniform rates thus allow broker compensation to be maintained without preserving commission differences across insurers. Aggregate income alone does not explain the network response: at a uniform rate of 3%, broker commission income remains below the status quo, yet menus are broader. A cap at the same rate instead narrows menus. Network breadth therefore depends on the returns to individual insurer relationships as well as the overall level of broker

Table 9. Alternative Policies

	(0)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		Commission rules			Access rules				
					Full network		Fiduciary + full network		
	Status quo	Cap 3%	Uniform 4%	Fiduciary duty	No exit	With exit	No exit	With exit	No brokers
Employers									
Employer surplus	-12.15	-11.26	0.33	-10.77	30.17	19.22	35.51	23.37	-44.74
Choice value	17.41	18.52	23.86	17.93	41.21	43.48	45.37	47.01	56.15
Hassle costs	29.55	29.78	23.53	28.70	11.03	24.26	9.86	23.64	100.89
Broker use (%)	70.3	69.7	74.5	71.2	87.0	73.8	88.4	74.5	0.0
Not offering (%)	14.4	14.7	12.2	14.0	6.1	12.5	5.5	12.1	48.0
Employees									
Employee WTP	165.86	166.37	170.78	167.14	186.20	175.55	188.82	177.39	114.85
Take-up (%)	63.5	63.7	65.5	64.0	71.5	67.2	72.5	67.9	43.0
Insurers									
Premium per enrollee	540.0	537.2	535.7	537.7	519.4	515.9	514.2	510.9	474.7
Markup (p/mc)	1.35	1.35	1.35	1.35	1.32	1.30	1.30	1.29	1.18
Insurer profit	150.78	150.52	149.29	149.22	151.04	137.38	146.04	132.71	59.29
Brokers									
Insurers per menu	3.13	3.06	3.56	3.06	4.39	4.38	4.39	4.38	—
Commission (%)	3.33	2.68	4.00	3.32	3.21	3.21	3.16	3.15	—
Broker commissions	19.44	15.37	24.34	19.52	22.72	19.32	22.48	18.98	—
Broker contracting cost	8.03	6.98	8.96	7.96	16.17	11.35	16.17	11.22	—
Broker profit	11.40	8.40	15.38	11.56	6.55	7.96	6.31	7.76	—
Share active (%)	100.0	100.0	100.0	100.0	100.0	81.2	100.0	80.5	—

Notes: This table reports equilibrium outcomes under alternative policy instruments, for all employers in the multi-insurer regions of New York in 2018; the North Country market, served by a single insurer, is excluded. Column (0) reports the status quo. Columns (1)-(3) report a 3% commission cap, a uniform commission rate of 4%, and a fiduciary duty; Columns (4)-(7) report the full-network requirement, a minimum menu of six insurers that equals every region's full network, alone and combined with a fiduciary duty. Columns (1)-(4) and (6) report endogenous-network levels before broker exit; Columns (5) and (7) report the access rules after broker exit. Column (8) removes brokers entirely: every employer purchases without a broker, and the rows defined among brokered firms and the broker rows are left blank. The cap, the uniform rate, and the minimum menu maximize employer surplus with endogenous networks over grids of 0–3% of premiums, 0-4% of premiums, and two to six insurers; all three lie at the edge of their grids. A broker exits when its profit falls below the first percentile of simulated status-quo profits in its market; the with-exit columns report expected outcomes over contracting-cost draws. Column (6) removes both distortions. Premium per enrollee is in dollars per covered member per month; the other dollar rows are in dollars per eligible employee per month. Premium per enrollee and markup are averages over all employers, over firms that use a broker, and over firms that do not; insurers per menu and the commission rate are averages among firms that use a broker; the remaining rows are averages over all employers. Share active is the share of brokers that remain active; it is 100 where no exit is applied. Employer surplus is net of hassle costs and equals choice value minus hassle costs.

compensation.

Fiduciary duty. We model fiduciary duty as removing the commission term from brokered employers' choice rule (8) while retaining commission schedules. This benchmark fully aligns recommendations with employers' preferences. Figure 12c and Table 9 report equilibrium outcomes. Network adjustment nevertheless erodes most of the gain.

First, fiduciary duty strengthens price competition within existing networks. Without the commission term, brokered employers respond more strongly to premiums, inducing insurers to lower prices. Section

7.1 documents this mechanism, with larger markup reductions among high-commission insurers. Holding networks fixed, fiduciary duty raises employer surplus by $3.66 per eligible employee per month, comparable to the gain under the selected commission cap. Fiduciary duty achieves this gain without mandating lower commission rates.

Second, network contraction offsets most of this gain, although total broker income barely changes. The employer-surplus gain falls from $3.66 per eligible employee per month with networks fixed to $1.38 once networks adjust. Steering had propped up the high-commission insurers in brokers' menus: steered placements, rather than clients' own demand, made these contracts worth their costs. Once clients follow their own preferences, those placements evaporate, and brokers drop the contracts. Thus, while fiduciary duty aligns recommendations with employer preferences within the menu, brokers still choose which insurers to include based on commissions net of contracting costs. Preserving commission rates therefore does not necessarily preserve insurer access. Fiduciary duty's modest edge over the cap comes through increased broker use and lower hassle costs, both of which move in the opposite direction under the cap. Overall, employers gain modestly at the expense of insurers, while brokers are left essentially unaffected.

Minimum menu sizes. We next vary the minimum number of insurers that each broker must carry, capped at the number available in its region. Figure 12b and Table A10 report equilibrium outcomes.

First, stricter requirements expand employers' choices and intensify insurer competition. Broker menus expand to include previously excluded insurers, reducing premiums and markups as brokered employers gain access to more competitors. The gains vary across the grid because each increase requires some brokers to expand their menu while leaving others unconstrained.

Second, aggregate broker use becomes non-monotonic in the minimum-menu requirement once exit is allowed. We use the exit rules in Section 7.1. Holding the set of active brokers fixed, broker use rises steadily with the requirement, as employers can access more insurers while retaining brokerage services. Commission income also rises, but contracting costs increase faster, reducing broker profits and inducing exit. Broader menus therefore attract employers to brokerage, while broker exit reduces its availability. These opposing forces cause broker use to peak at five insurers and decline underfull networks. Nevertheless, roughly 80% of brokers remain active even under the full-network requirement.

Third, employer surplus rises with the requirement, but broker exit substantially reduces the gains and leaves little additional benefit from the strictest requirement. Before exit, full networks raise employer surplus by $42.32 per eligible employee per month, reflecting broader meanus, lower premiums, and reduced hassle costs from purchasing through brokers. Employee take-up and willingness to pay also rise. Allowing exit reduces the employer-surplus gain to $31.37, as displaced clients must purchase directly or stop offering coverage. The resulting increase in hassle costs erodes the gains and partially reverses the expansion in coverage. Moving from five insurers to full networks adds little additional access, while additional exit offsets most of the remaining benefit. A five-insurer requirement therefore delivers nearly the same employer-surplus gain as full networks.

Finally, insurer profits remain close to the status quo before exit, as expanded coverage offsets stronger competition, but decline once broker exit reduces the coverage gains. The menu size requirement thus preserves broad insurer access through surviving brokers but may reduce 4employers' access to brokerage

services. Its overall benefits depend on both the breadth of broker menus and brokers' continued participation.

Fiduciary duty and menu size requirements. Figure 12c shows that conduct and access rules reinforce each other in raising employer surplus. A full-network requirement prevents brokers from offsetting the benefits of fiduciary duty by narrowing their menus. Within these maintained networks, removing the commission term strengthens price competition, lowering premiums and insurer profits. Accordingly, fiduciary duty adds $5.34 per eligible employee per month under full networks, compared with $1.38 when implemented alone. Its incremental benefit generally increases with the minimum-menu requirement and levels off near full networks. Combining fiduciary duty with full networks raises employer surplus by $47.66 before broker exit, exceeding the sum of the two policies' separate gains.

No brokers. Finally, we remove brokers entirely: employers that offer coverage purchase directly, choosing among all insurers in their market and free of commission-driven steering, but bearing the hassle costs we estimate. This exercise quantifies the value brokers bring. Table 9 reports the results.

First, higher hassle costs outweigh the improvement in choice value. With unrestricted menus and no steering, choice value rises by more than under any other policy, up $38.7. However, hassle costs rise by $71.3 as employers offering coverage must purchase directly. Employer surplus consequently falls by $32.59, the largest loss among the policies considered. The comparison with minimum-menu requirements is instructive: broader access raises employer surplus when brokers' services are preserved. Brokers' value thus lies not in access alone, which employers could obtain directly, but in reducing the costs of obstaining it.

Second, eliminating brokerage services sharply reduces coverage and insurer sales. The share of employers not offering coverage rises from 14% to 48% as many forgo coverage rather than bear the costs of purchasing directly. Employee take-up and willingness to pay also fall sharply, while insurer profits fall by more than half as the insured market contracts. Brokers therefore serve as an important distribution channel for insurers. The commission-rule grids reinforce this finding: as commission caps or uniform rates fall toward zero, commissions no longer cover contracting costs, and both instruments converge to the no-broker outcome (Figure 12a). Brokerage services thus remain valuable despite the distortions from steering and restricted menus.

Comparing policies. Figure 12c compares the policy instruments above. Four lessons emerge.

First, correcting steering can exacerbate restricted access, reversing the ranking of commission policies once broker networks respond. With networks fixed, commission caps and fiduciary duty increase employer surplus, while uniform rates lower it. With endogenous networks, this ranking reverses because the same compensation arrangements govern both brokers' recommendations and their insurer relationships. Caps and fiduciary duty improve choices and prices within existing menus, but the cap cuts the income every relationship generates, and the duty changes what each relationship earns even at unchanged schedules, so both induce network contraction and competition softens. In contrast, uniform rates, which look worst at fixed networks, equalize and sustain commission returns, encouraging broader networks. As Section 7.1 shows, ignoring network adjustments leads to overly stringent caps, and across policy interventions, it can even lead to the wrong policy choice.

Second, the distribution of gains determines which insurer relationships brokers keep and whether brokers remain in the market. Fiduciary duty, which eliminates steering, delivers modest employer gains at insurers' expense and leaves aggregate broker income unchanged; yet menus contract, because preserving aggregate compensation does not preserve the incremental return to every relationship. The selected uniform rate instead improves employer outcomes, raises broker profits, and broadens menus. The full-network requirement, which leaves steering in place, delivers much larger gains, nearly ninety percent of the benchmark policy's (the fiduciary duty combined with full networks), and brokers bear the mandated contracting costs through lower profits. Insurer profits remain roughly unchanged before exit, as expanded coverage offsets compressed margins. Curbing steering thus mainly redistributes surplus; broadening access creates it. That creation is financed out of broker profits, however, and the resulting exit erodes employers' hassle-cost savings and reverses part of the coverage expansion.

Third, policy design can mitigate the tradeoff. A uniform rate eliminates commission dispersion while leaving the compensation level to be chosen separately; at the selected rate, it supports broader networks, though proportional compensation retains some steering distortion. Combining fiduciary duty with the full-network requirement instead corrects recommendations while preventing active brokers from narrowing their menus. One works through compensation incentives, the other through direct regulation of recommendations and access, but both share one design: they decouple brokers' pay from their placements while leaving brokers enough income to stay in the market.

Fourth, sustaining brokerage services remains the constraint: every instrument's gains run through brokers who stay in the market. The uniform rate induces almost no exit; the access mandates do, and once exit is accounted for, their gain on a total-surplus basis is no larger than the uniform rate's. The no-broker counterfactual shows the constraint at its starkest: without brokers' services, even full, steering-free access leaves employers worse off. Together with the benchmark policy, it bounds the objects at stake: the broker-free loss measures the net value of brokers' services to employers; the benchmark's gain measures the cost of the two distortions; and the second exceeds the first. Every instrument, in the end, trades the steering distortion against the competitive-access distortion, and its net effect depends on the networks that compensation sustains, not only on the steering it corrects.

8. Conclusion

Our findings highlight the central role of intermediaries in employer-sponsored health insurance markets. Our key insights are twofold. First, brokers influence employers' plan choices and, through their contracting decisions, alter the degree of competition among insurers. Second, regulations on commissions entail an inherent tradeoff: while commission-driven steering distorts employers' choices, the expansion of broker-insurer networks can also intensify competition. This tension highlights the importance of carefully designed policies that strike a balance between the benefits of limiting distortions and the potential costs of weakening competition.

These insights have broader policy relevance. First, similar dynamics may arise in other vertical markets where intermediaries shape consumers' effective choice sets, such as the grocery retailing setting, where

store shelves determine which upstream brands are visible to consumers. Second, our results also inform ongoing policy discussions regarding the regulation of payments to intermediaries. By quantifying the welfare effects of steering and network breadth, the paper sheds light on whether commission caps improve efficiency. More generally, the evidence suggests that intermediary frictions may constitute a significant barrier to competition, with implications for explaining the persistently rising healthcare costs.

References

Abaluck, Jason and Jonathan Gruber, "Choice Inconsistencies among the Elderly: Evidence from Plan Choice in the Medicare Part D Program," *American Economic Review*, 2015.

Allcott, Hunt, "The welfare effects of misperceived product costs: Data and calibrations from the automobile market," *American Economic Journal: Economic Policy*, 2013, *5* (3), 30–66.

Andrews, Donald W. K. and Gustavo Soares, "Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection," *Econometrica*, 2010, *78* (1), 119–157.

Arnold, Jonathan, Atul Gupta, and Alexander Olssen, "Targeting Efficiencies? Reducing Costs Through Hospital Consolidation," *Unpublished Working Paper*, 2025.

Barwick, Panle Jia, Parag A Pathak, and Maisy Wong, "Conflicts of interest and steering in residential brokerage," *American Economic Journal: Applied Economics*, 2017, *9* (3), 191–222.

Berry, Steven, "Estimating Discrete-Choice Models of Product Differentiation," *RAND Journal of Economics*, 1994, pp. 242–262.

__ , James Levinsohn, and Ariel Pakes, "Automobile Prices in Market Equilibrium," *Econometrica*, 1995, *63* (4), 841–890.

Bhattacharya, Vivek, Gaston Illanes, and Manisha Padi, "Fiduciary duty and the market for financial advice," Technical Report, National Bureau of Economic Research 2019.

Böckerman, Petri, Sami Jysmä, and Ohto Kanninen, "Difference-in-Kinks Design," Discussion Paper 18313, IZA 2025.

Boehm, Eduard, "Intermediation, Choice Frictions, and Adverse Selection: Evidence from the Chilean Pension Market," Technical Report 2023.

Brown and Brown, "Investor Presentation—2020," SEC Exhibit 99.1 May 2020.

Bundorf, M Kate, Jonathan Levin, and Neale Mahoney, "Pricing and Welfare in Health Plan Choice," *American Economic Review*, 2012, *102* (7), 3214–3248.

Calonico, Sebastian, Matias D Cattaneo, and Rocio Titiunik, "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica*, 2014, *82* (6), 2295–2326.

Canay, Ivan A, Gastón Illanes, and Amilcar Velez, "A User's Guide for Inference in Models Defined by Moment Inequalities," *Journal of Econometrics*, 2023, p. 105558.

Card, David, David S. Lee, Zhuan Pei, and Andrea Weber, "Inference on Causal Effects in a Generalized Regression Kink Design," *Econometrica*, 2015, *83* (6), 2453–2483.

Cattaneo, Matis D., Michael Jansson, and Xinwei Ma, "Manipulation testing based on density discontinuity," *The Stata Journal*, 2018, *18* (1), 234–261.

CMS, "The National Health Expenditure Accounts," 2024. link.

Commonwealth Funds, "A New Rule to Protect Medicare Beneficiaries Against Inappropriate Sales Tactics Is Stuck in the Courts," 2025. link.

Conlon, Christopher T and Julie Holland Mortimer, "Efficiency and Foreclosure Effects of Vertical Rebates: Empirical Evidence," *Journal of Political Economy*, 2021, *129* (12), 3357–3404.

Craig, Stuart, "Competition in Employer-Sponsored Health Insurance: Implications for a Public Option," Technical Report, Working Paper 2022.

Cutler, David and Sarah Reber, "Paying for Health Insurance: The Trade-off Between Competition and Adverse Selection," *Quarterly Journal of Economics*, 1998, *113* (2), 433–466.

Dafny, Leemore, Kate Ho, and Maurcio Varela, "Let Them Have Choice: Gains from Shifting Away from Employer-Sponsored Health Insurance and Towards an Individual Exchange," *American Economic Journal: Economic Policy*, 2013, *5* (1), 32–58.

Dickstein, Michael J and Eduardo Morales, "Accounting for expectational and structural error in binary choice problems: A moment inequality approach," Technical Report, National Bureau of Economic Research 2013.

Eizenberg, Alon, "Upstream Innovation and Product Variety in the US Home PC Market," *Review of Economic Studies*, 2014, *81* (3), 1003–1045.

Fleitas, Sebastián, Caitlyn Fleming, Gautam Gowrisankaran, and Anthony T Lo Sasso, "Group Health Insurance Versus ACA Marketplaces: Selection, Subsidies, and Welfare," Technical Report, National Bureau of Economic Research 2025.

Gates, Susan M., Pinar Karaca-Mandic, James R. Burgdorf, and Kanika Kapur, "Health Savings Accounts for Small Business and Entrepreneurs," *Open Health Services and Policy Journal*, 2010, *3* (1), 10–20.

Geddes, Eilidh, "The Effects of Price Regulation in Markets with Strategic Entry: Evidence from Health Insurance Markets," 2022.

Grembi, Veronica, Tommaso Nannicini, and Ugo Troiano, "Do Fiscal Rules Matter?," *American Economic Journal: Applied Economics*, 2016, *8* (3), 1–30.

Grennan, Matthew, Kyle R Myers, Ashley Swanson, and Aaron Chatterji, "No free lunch? welfare analysis of firms selling through expert intermediaries," *Review of Economic Studies*, 2024, p. rdae090.

Gruber, Jonathan, "The tax exclusion for employer-sponsored heatlh insurance," *NBER Working Paper*, 2010.

__, Benjamin R Handel, Samuel H Kina, and Jonathan T Kolstad, "Managing intelligence: Skilled experts and decision support in markets for complex products," *NBER Working Paper*, 2021, (w27038).

Handel, Benjamin R, "Adverse selection and inertia in health insurance markets: When nudging hurts," *American Economic Review*, 2013, *103* (7), 2643–2682.

__ and Jonathan Kolstad, "Health Insurance for "Humans": Information Frictions, Plan Choice, and Consumer Welfare," *American Economic Review*, 2015.

HealthPartners, "Should I Use a Health Insurance Broker?," 2021. HealthPartners Plan Blog. link.

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

Hristakeva, Sylvia, "Vertical contracts with endogenous product selection: An empirical analysis of vendor allowance contracts," *Journal of Political Economy*, 2022, *130* (12), 3202–3252.

Karaca-Mandic, Pinar, Roger Feldman, and Peter Graven, "The role of agents and brokers in the market for health insurance," *Journal of Risk and Insurance*, 2018, *85* (1), 7–34.

KFF, "Health Care Costs and Affordability," 2025. link.

__ , "Employer-Sponsored Health Insurance 101," 2026. link.

Kim, Gi Heung, "The equilibrium impacts of broker incentives in the real estate market," 2024.

Lee, David S. and Thomas Lemieux, "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, 2008, *48*, 281–355.

Li, Anran, Becky Staiger, Diane Alexander, and Molly Schnell, "Enrollment Brokers Did Not Increase Medicaid Enrollment, 2008–18: Study examines the effectiveness of Medicaid enrollment brokers.," *Health Affairs*, 2022, *41* (9), 1333–1341.

Li, Fei, Charles Murry, Can Tian, and Yiyi Zhou, "Intermediaries in Decentralized Markets: Evidence from Used-Car Transactions," Technical Report, National Bureau of Economic Research 2025.

Liu, Chenyuan and Justin Sydnor, "Dominated Options in Health Insurance Plans," *American Economic Journal: Economic Policy*, 2022, *14* (1), 277–300.

Liu, Tong, "Bargaining with private equity: Implications for hospital prices and patient welfare," *Unpublished Working Paper*, 2022.

Marketing, Ritter Insurance, "Contract & Get Appointed to Sell with Health Insurance Companies," 2025. Ritter Insurance Marketing. link.

Marone, Victoria and Adrienne Sabety, "When Should There Be Vertical Choice in Health Insurance Markets?," *American Economic Review*, 2022, *112* (1), 304–342.

Marsh McLennan Agency, "Changing Your Insurance and Employee Benefits Broker," 2026. Accessed October 11, 2026.

McCrary, Justin, "Manipulation of the running variable in the regression discontinuity design: A desntiy test," *Journal of Econometrics*, 2008, *142* (2), 698–714.

Myerson, Rebecca and Honglin Li, "Information gaps and health insurance enrollment: evidence from the Affordable Care Act navigator programs," *American journal of health economics*, 2022, *8* (4), 477–505.

Nevo, Aviv, "Measuring market power in the ready-to-eat cereal industry," *Econometrica*, 2001, *69* (2), 307–342.

New York State Insurance Department, "OGC Opinion No. 10-06-03: HMO Commissions to Brokers," 2010. Office of General Counsel, New York State Department of Financial Services. link.

New York Times, "Financial Brokers Must Now Act in Your Best Interest. What Does That Mean?," 2020. link.

Pakes, Ariel, Jack Porter, Kate Ho, and Joy Ishii, "Moment inequalities and their application," *Econometrica*, 2015, *83* (1), 315–334.

PeopleKeep, "10 Tips on Hiring a Health Insurance Professional for Your Small Business," 2013. PeopleKeep Blog. link.

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

Reuters, "Under the microscope: Spitzer-era ban on insurer fees," 2008. link.

Rey, Patrick and Thibaud Vergé, "Resale Price Maintenance and Interlocking Relationships," *The Journal of Industrial Economics*, 2010, *58* (4), 928–961.

__ **and** __ , "Secret Contracting in Multilateral Relations," 2016.

Ritter Insurance Marketing, "2025 Maximum Broker Commissions for Medicare Advantage & Medicare Part D," 2024. Ritter Insurance Marketing. link.

Robles-Garcia, Claudia, "Competition and incentives in mortgage markets: The role of brokers," *Unpublished working paper*, 2019.

Shaffer, Greg, "Slotting Allowances and Optimal Product Variety," *The BE Journal of Economic Analysis & Policy*, 2005, *5* (1), 000010151515380637l083.

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

Sidley Austin LLP, "CMS Finalizes Proposal to Overhaul Medicare Advantage Broker Payments," 2024. link.

Tilipman, Nicholas, "Employer incentives and distortions in health insurance design: Implications for welfare and costs," *American Economic Review*, 2022, *112* (3), 998–1037.

Tong, Joy Tianjiao, "Advising sophisticated customers: Evidence from health insurance brokers," *SSRN Electronic Journal, October. https://doi. org/10.2139/ssrn*, 2022, *4597241*.

Wooldridge, Jeff, "What's New in Econometrics? Lecture 6: Control functions and related methods," *NBER. http://www. nber. org/WNE/lect*, 2007, *6*, 420–445.

Online Appendix

Contents

A Additional Tables and Figures **57**

A1 Figures . . . 57

A2 Tables . . . 68

B Additional Details on Data **77**

B1 Panel Construction Using Form 5500 Filings . . . 77

B2 Data from the Large National Representative Insurer . . . 79

B3 Clarivate Data and Imputed Product Characteristics . . . 79

C Theoretical Derivations and Microfoundation **82**

C1 Strategic Commissions and Equilibrium Existence . . . 82

C2 Menu Formation Without Steering . . . 82

C3 Menu Formation With Steering . . . 82

C4 Welfare Decomposition . . . 83

C5 Extensions . . . 84

C6 Micro-foundation of the Steering Specification . . . 84

D Additional Model and Estimation Details **91**

D1 Employee Surplus From Having Access to Employer-Sponsored Insurance . . . 91

D2 Marginal Cost and the Enrolled Pool . . . 94

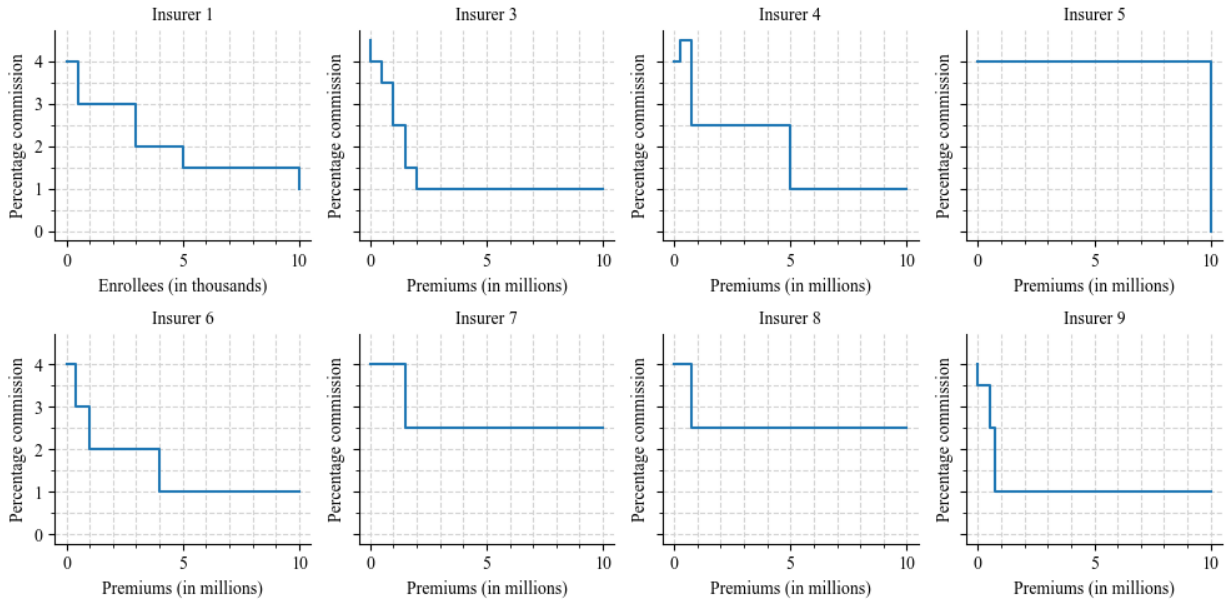
D3 Distance is the Primary Factor in Broker-Employer Matches . . . 96

D4 Multiplicity of Equilibria . . . 97

A. Additional Tables and Figures

A1. Figures

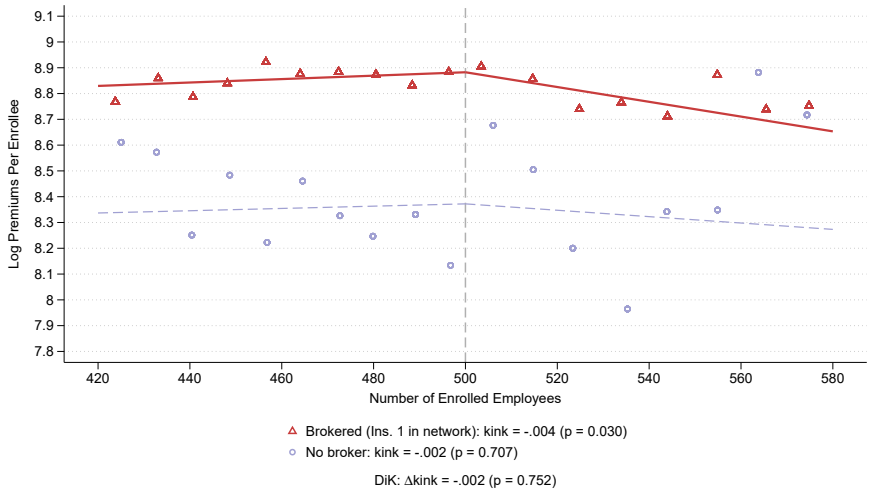
Figure A1. Commission schedules by insurer



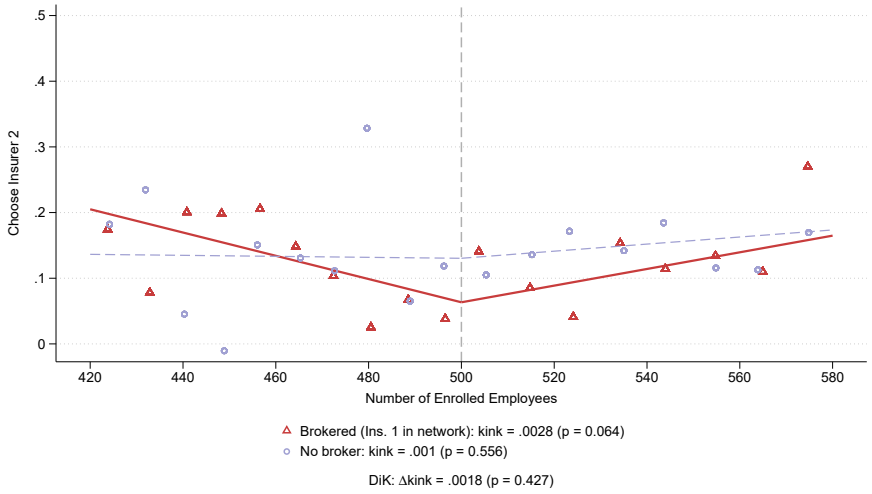
Notes: This figure plots the snapshot of comission schedules in 2017 for major insurers in NY ESI. Insurer 2 has per member per month commissions. Insurer 10's commission schedule is unknown.

Figure A2. Difference-in-Kinks: Brokered vs. Non-Brokered Employers

(a). Premiums Per Enrollee

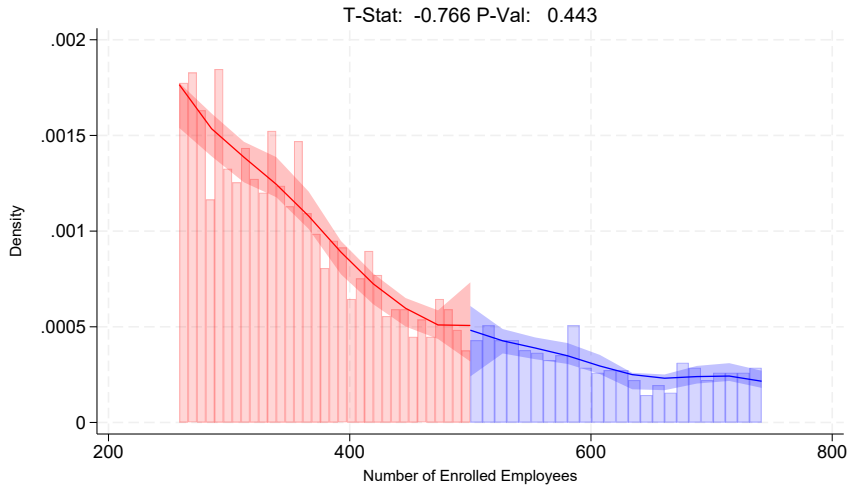


(b). Probability of Choosing Insurer 2



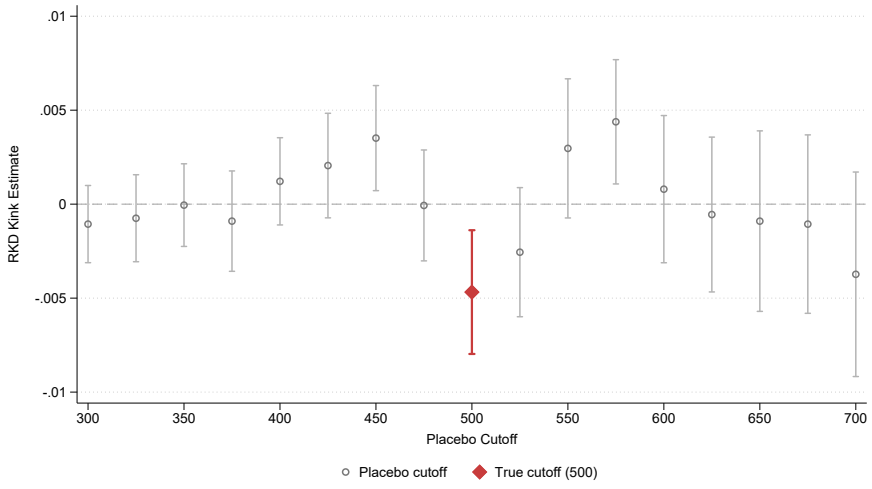
Notes: The figure reports difference-in-kinks estimates from equation (4). Panel (a) uses log premiums per enrollee as the outcome; panel (b) uses an indicator for choosing Insurer 2. The treatment group consists of brokered employers whose broker contracts with Insurer 1; the control group consists of non-brokered employers. The running variable is enrolled employees, centered at Insurer 1's 500-enrollee commission cutoff. Binned means and fitted lines are constructed as in Figure 7.

Figure A3. McCrary Density Test



Notes: This figure reports the density test of McCrary (2008) for the running variable (number of enrolled employees) at the commission cutoff of 500. A discontinuity in the density would suggest manipulation of the running variable. The analysis sample is restricted to fully insured, brokered employers who have Insurer 1 in their choice set and more than 100 enrolled employees.

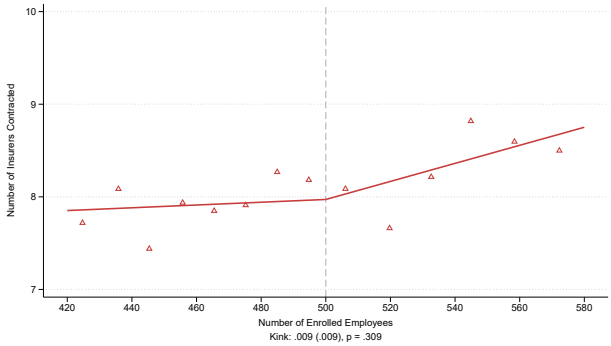
Figure A4. Placebo Cutoff Test



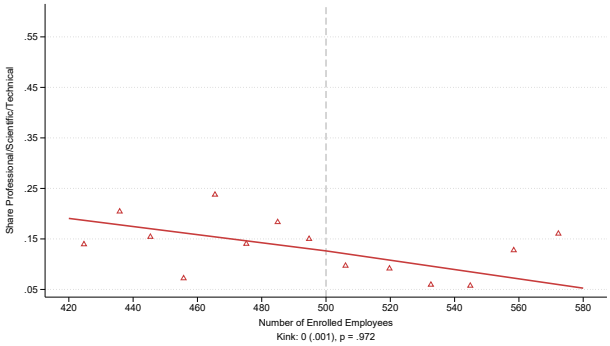
Notes: This figure reports RKD kink estimates from equation (3) at placebo cutoffs ranging from 300 to 700 enrolled employees, in increments of 25. The true cutoff at 500 (red diamond) is the commission kink for Insurer 1. Placebo cutoffs within one bandwidth (80 employees) of the true cutoff are sensitive to contamination by the treatment Vertical bars denote 95% confidence intervals. All estimates use a fixed bandwidth (CER optimal) and include region covariates.

Figure A5. RKD Covariate Smoothness

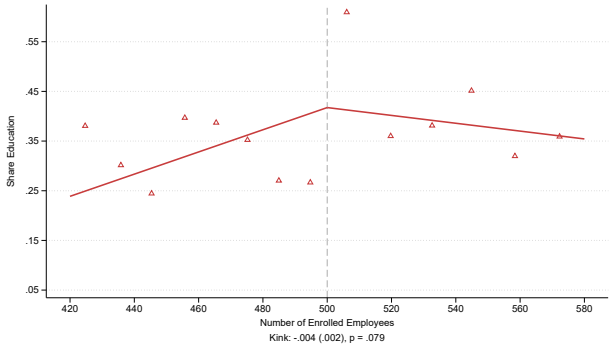
(a). Number of Insurers Contracted



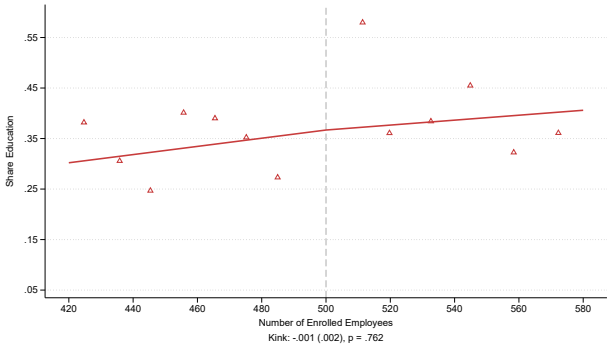
(b). Share Professional/Scientific/Technical



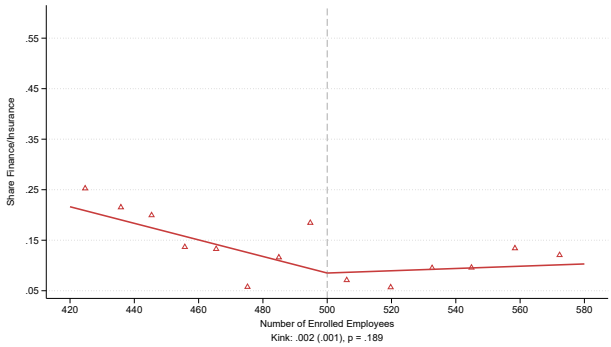
(c). Share Education



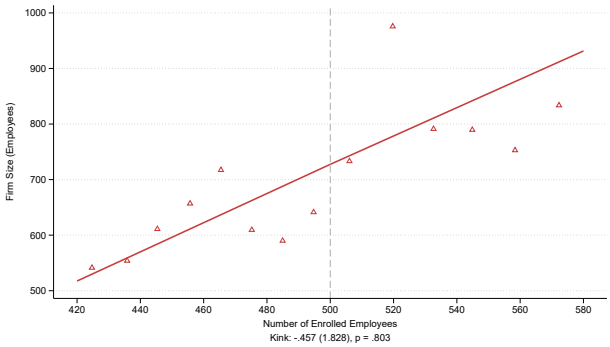
(d). Share Education: Donut (± 10)



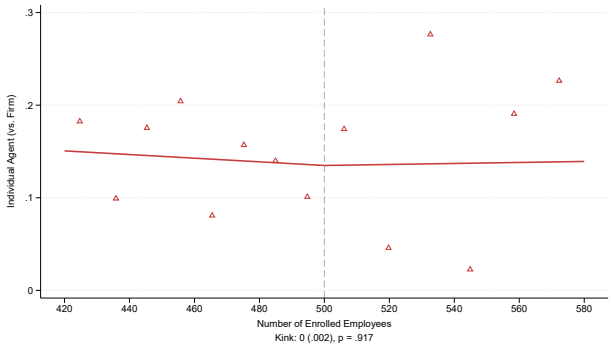
(e). Share Finance/Insurance



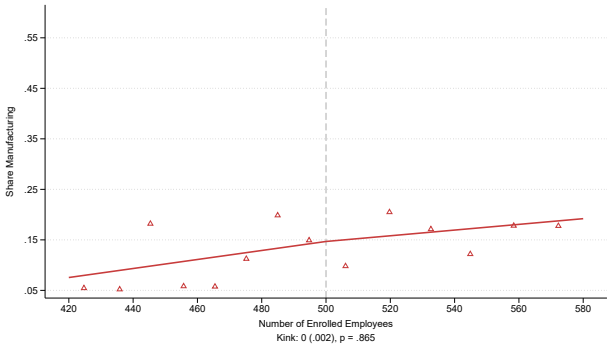
(f). Firm Size



(g). Individual Agent (vs. Firm)



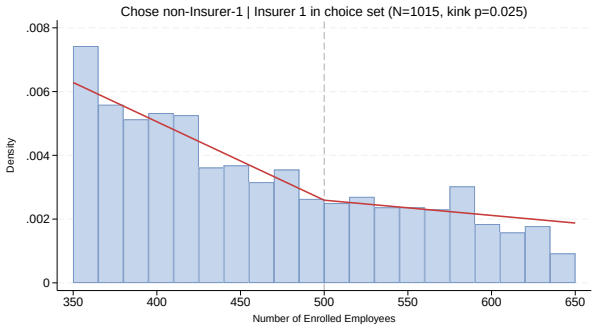
(h). Share Manufacturing



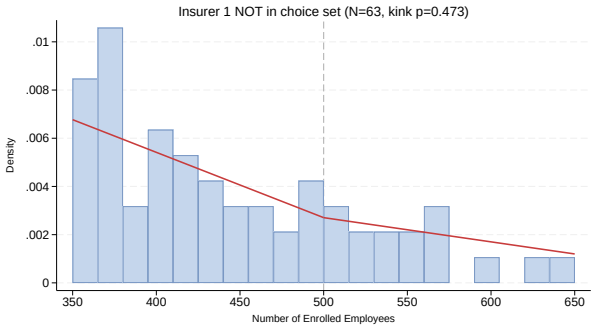
Notes: This figure tests whether predetermined covariates exhibit a kink at the commission cutoff of 500 enrolled employees. Each panel reports the RKD kink estimate from equation (3) with the covariate as the dependent variable. The specification matches the main analysis: local linear polynomial ($p = 1$), bandwidth of 80, triangular kernel, and region covariates (NYC, Mid Hudson). Panel (d) repeats the education test excluding employers within ± 10 enrolled employees of the cutoff, where a small cluster of education firms generates compositional noise; the kink estimate becomes insignificant. Insignificant kink estimates support the identifying assumption that the composition of employers evolves smoothly through the cutoff.

Figure A6. Conditional Density Test of Broker Steering

(a). Chose a non–Insurer 1 carrier (Insurer 1 in choice set)

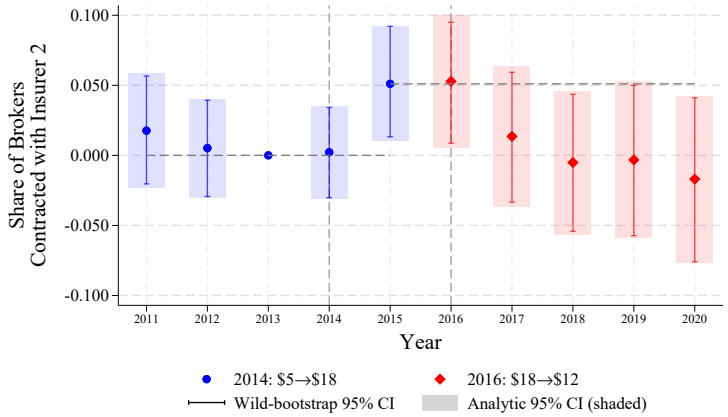


(b). Insurer 1 not in choice set (placebo)



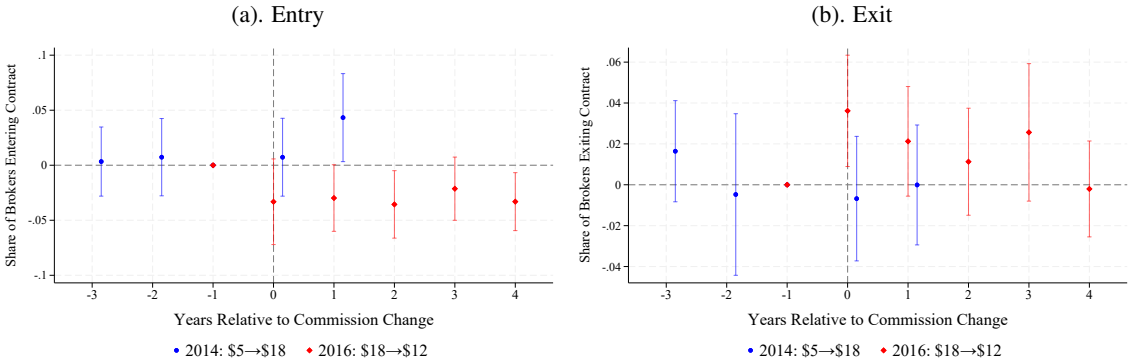
Notes: This figure gives a density-based view of the steering RKD. Panel (a) plots the density for brokered employers who have Insurer 1 in their choice set but chose a different carrier. Panel (b) plots the density for employers whose broker does not carry Insurer 1, and who therefore face no Insurer 1 commission incentive. Histograms use 15-employee bins; red lines are local linear-kink fits. Reported p-values are for the slope change at 500, from the equivalent membership regression kink design (a local-linear RKD, $p = 1$, deriv. 1, on a group indicator over the pooled sample). The sample matches the main RKD: fully insured, single-carrier employers, 2012–2019, with at least 100 enrolled employees.

Figure A7. Effect of commission changes on broker networks, alternative specification



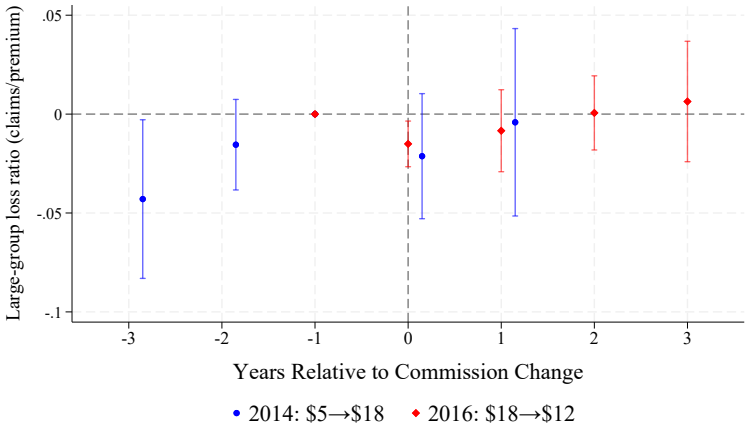
Notes: This figure plots the point estimates and 95% confidence interval from the estimation of equation (5). We use a single event-study specification with 2013 as the benchmark year. The outcome is the county-year level share of brokers contracting with each insurer, and estimates compare Insurer 2 with other insurers. Solid bars denote wild-bootstrap 95% confidence intervals and shaded bands denote analytic 95% confidence intervals, both clustered by county. Estimates include insurer and county-by-year fixed effects.

Figure A8. Broker Network Response: Entry and Exit Margins



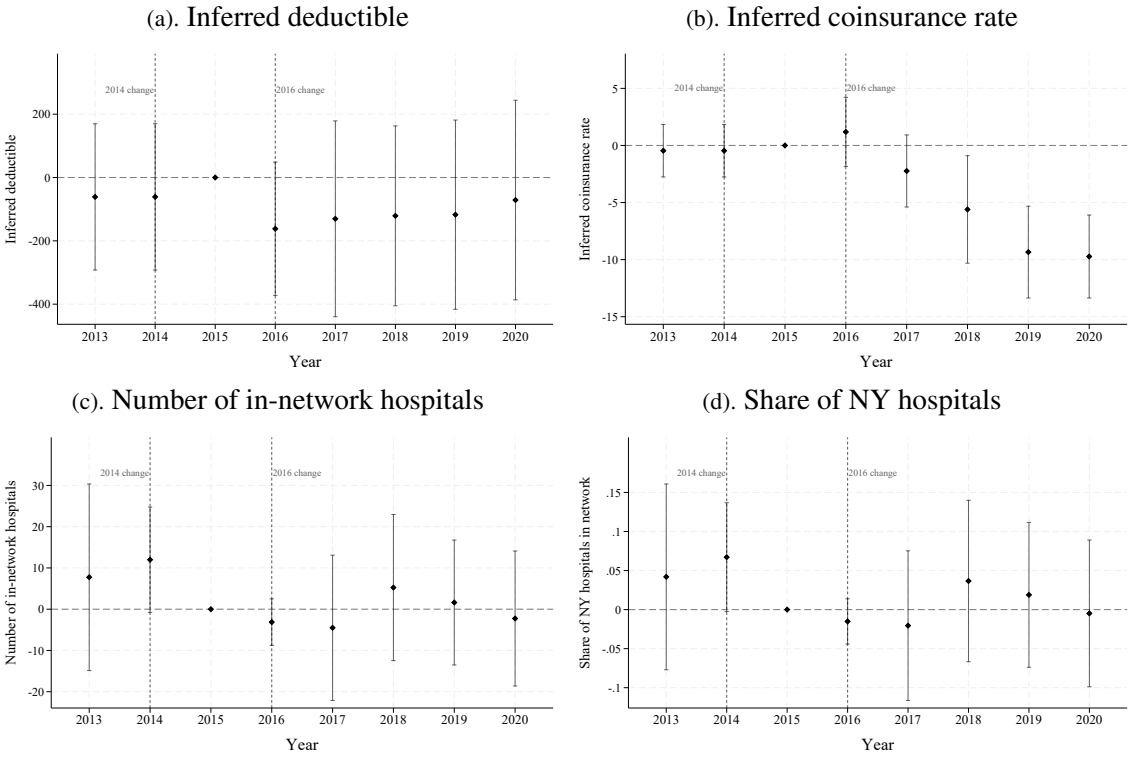
Notes: This figure decomposes the broker-network event study in Figure 8 into the share of brokers entering (Panel A) and exiting (Panel B) a contracting relationship with each insurer, measured at the insurer-county-year level. Entry and exit are defined using observed broker-insurer relationships in the panel: a broker-insurer pair is classified as an entry if it appears after being absent in the previous year, and as an exit if it is absent in the following year. By construction, net entry equals the year-over-year change in the stock of broker-insurer contracts shown in Figure 8. Specifications include insurer and county-by-year fixed effects. Standard errors are clustered at the county level, as in Figure 8, and bars denote 95% confidence intervals. Blue circles correspond to the 2014 commission increase, and red diamonds correspond to the 2016 commission decrease.

Figure A9. Placebo Event Study: Insurer 2's Large-Group Loss Ratio



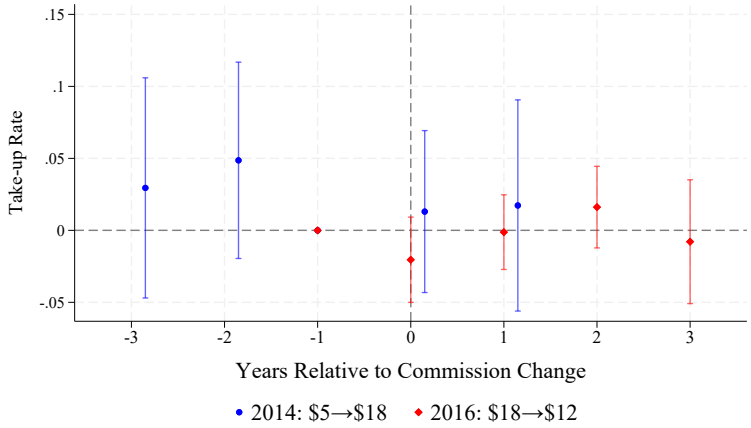
Notes: This figure presents stacked event-study estimates of the large-group medical loss ratio (claims divided by premiums), constructed from CMS Medical Loss Ratio (MLR) filings for New York and aggregated to the NAIC group-by-year level for 2011-2019. Blue circles correspond to the 2014 commission increase ($5 → $18), while red diamonds correspond to the 2016 commission decrease ($18 → $12). The omitted reference period is the year immediately preceding each event. All specifications include insurer and year fixed effects. Standard errors are clustered at the insurer-group level (10 groups), and bars denote 95% confidence intervals.

Figure A10. Placebo Event Studies: Insurer 2's Plan Design and Provider Network Breadth



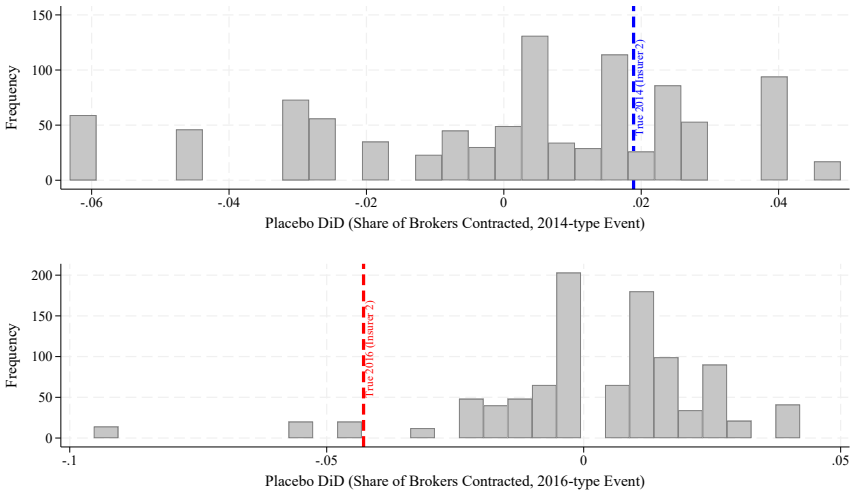
Notes: This figure presents stacked event-study estimates of plan characteristics, including the deductible (panel (a)) and coinsurance rate (panel (b)), inferred from claims data at the insurer-by-region-by-year level between 2013 and 2020, and the number of New York hospitals observed in-network for each insurer in the claims data (panel (c)) and the share of all claims-active New York hospitals covered by the insurer (panel (d)), measured at the insurer-year level from 2013 to 2020. We estimate the event study by pooling all years together, using 2015 as the omitted reference period. Vertical lines mark the 2014 and 2016 commission changes. Panels (a) and (b) include insurer and region-by-year fixed effects, are weighted by the number of patients, and cluster standard errors by region. Panels (c) and (d) include insurer and year fixed effects and cluster by insurer group. Bars denote 95% confidence intervals.

Figure A11. Placebo Event Study: Insurer 2's Take-up Rate



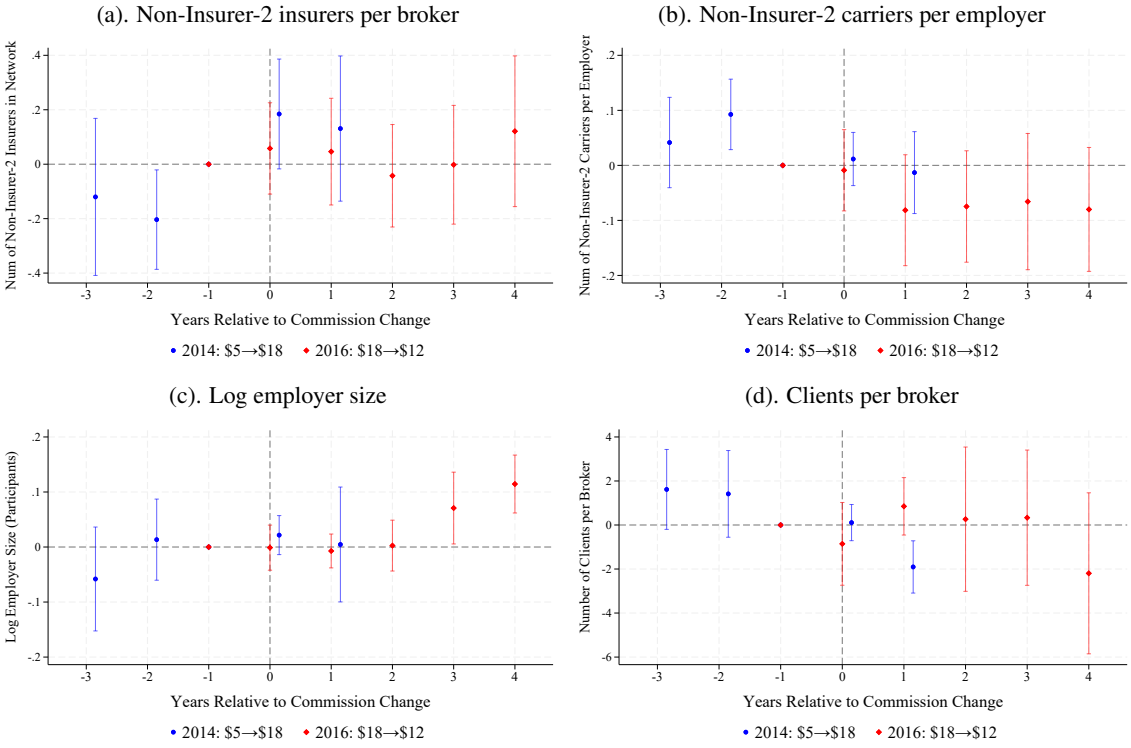
Notes: The figure plots event-study estimates of Insurer 2's take-up rate around its 2014 commission increase and 2016 commission decrease. The outcome is the share of eligible employees enrolling in Insurer 2 when it is offered. Estimates are normalized to the year before each event. Bars denote 95 percent confidence intervals. Standard errors are clustered at the county level.

Figure A12. Broker-contracting Placebo: Insurer 2 vs. Placebo Insurers



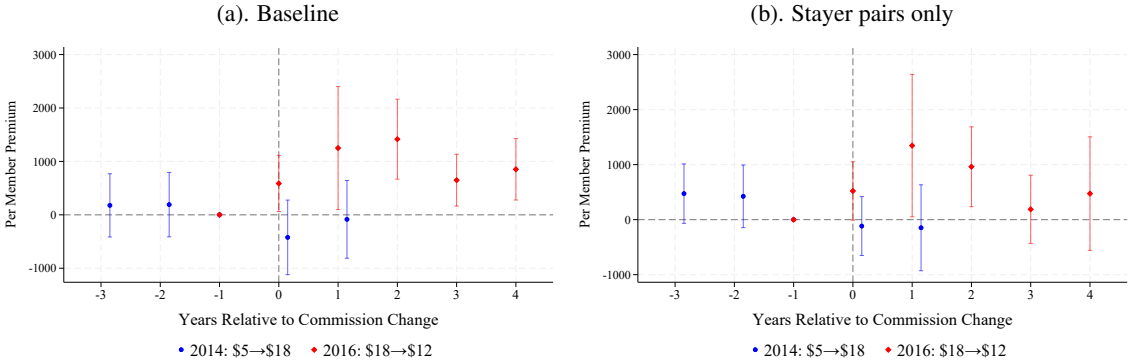
Notes: This figure reports the null distribution of 1,000 placebo difference-in-differences (DiD) estimates against Insurer 2's true estimate (vertical line). For each replication, we draw one placebo insurer uniformly from the nine non-Insurer-2 carriers and two ordered placebo event years (gap$\geq$2 years) from $\{2012, 2013, 2015, 2017, 2018, 2019\}$, which excludes the true 2014 and 2016 event years. The same placebo insurer is used for both placebo events within a draw, mirroring Insurer 2's paired shock history. For each draw we estimate two pooled pre/post DiD regressions: one using the 2014-type sample window and one using the 2016-type window with insurer and county-by-year fixed effects and standard errors clustered by county. The true estimates are Insurer 2's pooled DiD coefficients from the actual commission changes.

Figure A13. Placebo Event Studies: Treated Brokers' Choice Sets and Client Composition



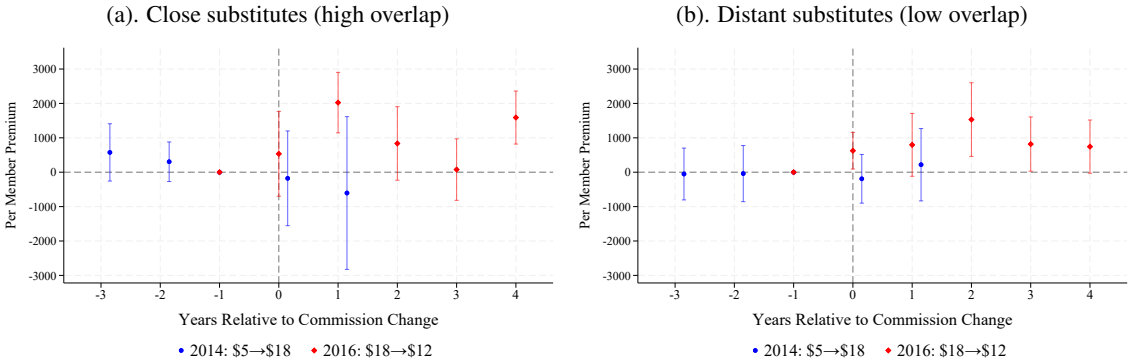
Notes: This figure presents broker-cohort event-study estimates for outcomes other than contracting with Insurer 2. The outcomes are the number of non-Insurer 2 insurers in a broker's network (Panel A), the number of non-Insurer 2 carriers per client employer-year (Panel B), the log number of participants per employer (Panel C), and the number of client employers per broker-county-year (Panel D). Treated cohorts consist of brokers without Insurer 2 in 2011-2013 for the 2014 event (blue circles) and brokers with Insurer 2 in 2014-2015 for the 2016 event (red diamonds). The omitted reference period is the year immediately preceding each event. Specifications include broker and county-by-year fixed effects, plus employer fixed effects in Panels B and C. Panels B to D use fully insured employers with at least 100 participants. Standard errors are clustered at the county level, and error bars denote 95% confidence intervals.

Figure A14. Rival-Premium Response: Baseline Versus Stayer Employer-Insurer Pairs



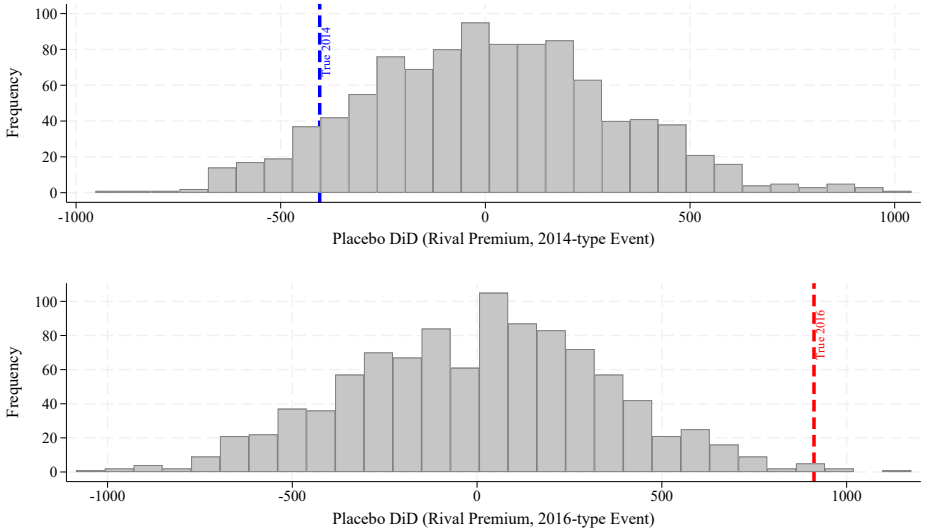
Notes: This figure replicates the rival-premium event study shown in Figure 9. Panel A reproduces the baseline specification using per-member premiums, winsorized at the 1st and 99th percentiles, for fully insured employers with at least 100 participants; Insurer 2 is excluded from the sample. Panel B further restricts the analysis to employer-insurer pairs observed both before and after the event within each event window, thereby shutting down employer-switching and other compositional margins. Specifications include employer, broker, insurer, and county-by-year fixed effects. Standard errors are clustered at the county level, as in Figure 9, and bars denote 95% confidence intervals.

Figure A15. Rival-Premium Response by Broker-Network Overlap with Insurer 2



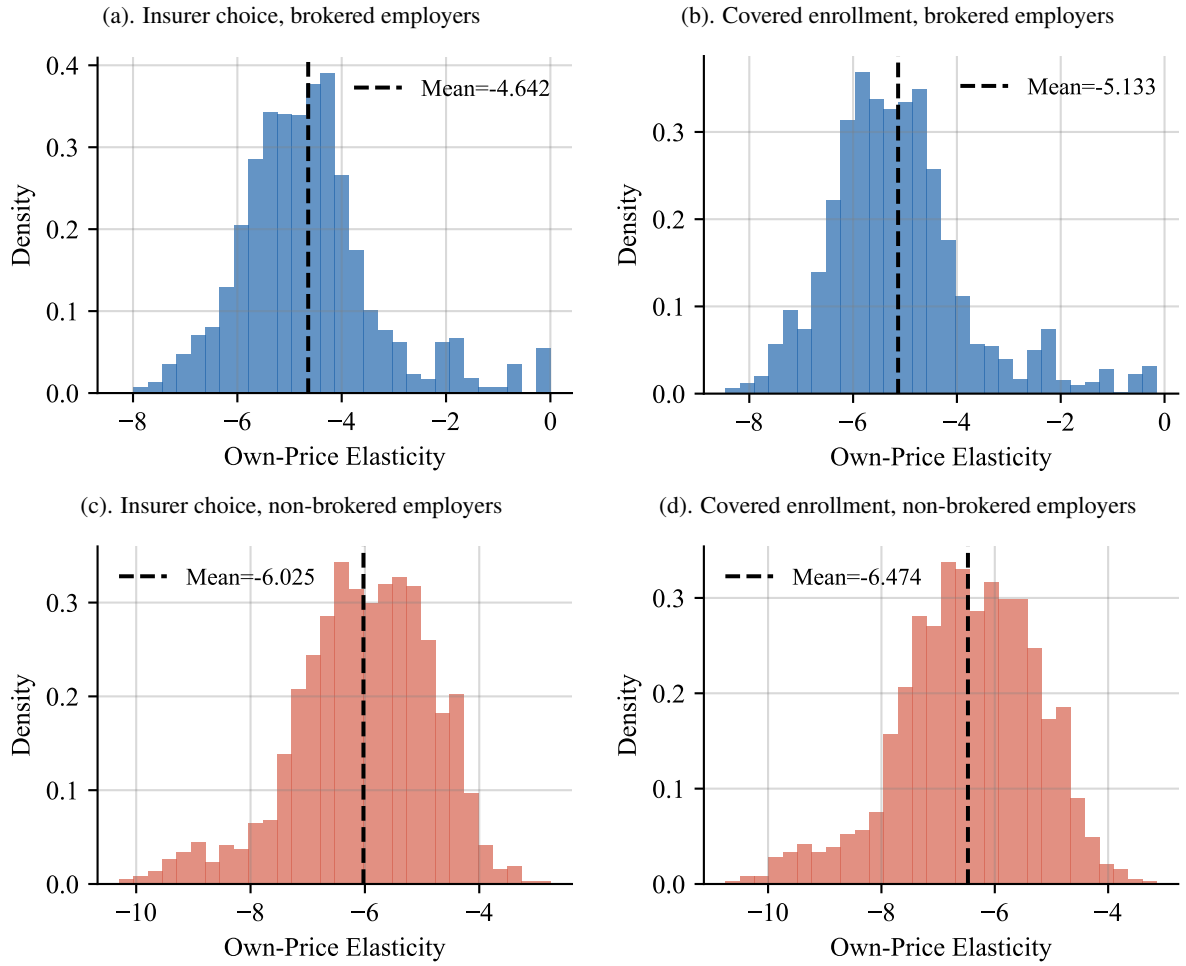
Notes: This figure decomposes the rival-premium event study in Figure 9 by each rival carrier's pre-period broker-network overlap with Insurer 2. Overlap is defined as the share of a carrier's contracted brokers that also contract with Insurer 2 during the pre-event period (2011–2013 for the 2014 event and 2014–2015 for the 2016 event). Rival carriers are split at the median overlap across the nine competing insurers. Panel A reports estimates for carriers with above-median overlap, which can be viewed as closer substitutes for Insurer 2, while Panel B reports estimates for carriers with below-median overlap. Specifications are identical to those in Figure A14.

Figure A16. Rival-premium Placebo: Event-specific Null Distributions



Notes: This figure plots the null distribution of 1,000 placebo difference-in-differences (DiD) estimates against the true estimate (vertical line). For each replication, we draw two ordered placebo event years (gap≥2 years) from {2012, 2013, 2015, 2017, 2018, 2019}, which excludes the true 2014 and 2016 event years; the earlier year corresponds to the 2014-type shock and the later year to the 2016-type shock. For each draw, broker-county cohorts are independently re-randomized into treatment and control groups at the observed treated shares from the actual events. We then estimate the pooled pre/post DiD on rival insurer premiums for each placebo event using its corresponding sample window, with employer, broker, insurer, and county-by-year fixed effects, clustering standard errors by county. The true estimates use Insurer 2's actual broker-cohort assignments and the actual 2014 and 2016 commission changes.

Figure A17. Own-Price Elasticities: Employer Choice and Covered Enrollment



Notes: This figure plots the distribution of own-price elasticities at the status quo equilibrium, separately for employers with and without brokers. Each observation is an employer–insurer cell. The choice elasticity is the response of the employer's probability of selecting an insurer to that insurer's premium, holding other premiums fixed; the enrollment elasticity additionally folds in workers' take-up and gives the response of the insurer's covered enrollment. Dashed lines mark means. For non-brokered employers, both objects include movement to the no-offer option. These are responses to a single insurer's price, not to a uniform change in all premiums.

A2. Tables

Table A1. RKD Robustness

	Linear ($p = 1$)			Quadratic ($p = 2$)		
Bandwidth	Kink Est.	SE	N	Kink Est.	SE	N
Panel A: Bandwidth and Polynomial Order						
Half bandwidth (40)	-0.0036	(0.0047)	286	0.0203	(0.0171)	286
Baseline (80)	-0.0047***	(0.0017)	573	-0.0043	(0.0063)	573
1.5x bandwidth (120)	-0.0025***	(0.0009)	927	-0.0076**	(0.0035)	927
CER-optimal	-0.0033***	(0.0011)	756	-0.0076*	(0.0045)	756
MSE-optimal	-0.0012*	(0.0006)	1307	-0.0068***	(0.0023)	1307
Panel B: Covariates ($p = 1$, $h = 80$)						
No covariates	-0.0040**	(0.0018)	573			
Baseline covariates	-0.0047***	(0.0017)	573			
Employer fixed effects	-0.0024***	(0.0009)	573			
Panel C: Other Robustness ($p = 1$, $h = 80$)						
Uniform kernel	-0.0040***	(0.0012)	583			
Donut hole (+/-10)	-0.0061***	(0.0020)	511			
Donut hole (+/-20)	-0.0027	(0.0026)	431			

Notes: This table reports RKD kink estimates for the probability that an employer chooses Insurer 1 under alternative specifications. The sample comprises fully insured, single-carrier employers whose broker has Insurer 1 in network; the dependent variable and estimating equation are those of Figure 7. Each row varies a single dimension of the local-polynomial estimator relative to the baseline (local linear, bandwidth 80, triangular kernel, with covariate adjustments for the NYC and Mid-Hudson regions). Columns report the kink estimate, its standard error, the bandwidth, and the effective number of observations. "CER" and "MSE" denote the coverage-error- and mean-squared-error-optimal bandwidths of Calonico et al. (2014), and "donut $\pm k$" excludes observations within k enrolled employees of the cutoff. At wider bandwidths, the quadratic estimates are larger and more precise, absorbing curvature that would otherwise bias the linear fit; the employer fixed-effects estimate is smaller, consistent with within-firm variation being less noisy. Standard errors are heteroskedasticity-robust and bias-corrected following Calonico et al. (2014). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A2. First Stage of Employer Preference Estimation

Premium per Employee	Coefficient	Standard Error
Leave-out (Hausman) premium instrument	2.8858	0.0442
WTP per employee	1.3841	0.0162
has broker	4.6808	2.5813
has broker $\times$ effective commission	212.3700	74.6452
Constant	0.0000	1.8596
F-stats	4,269.8	

Notes: This table reports the first stage of employer preferences estimates, where the endogenous regressor is premium per employee and the excluded instrument is the leave-out Hausman instrument. We include insurer fixed effects.

Table A3. Margins by Menu Breadth

Number of insurers in menu	Take-up-weighted price-cost ratio	Price / standardized marginal cost	Menu configurations	Regions
1	1.549	1.477	27	7
2	1.808	1.721	47	6
3	1.383	1.327	46	6
4	1.293	1.243	26	4
5	1.264	1.216	8	3
6	1.255	1.207	1	1

Notes: This table reports equilibrium price-cost ratios across all solved broker menus containing the indicated number of insurers, including both observed and counterfactual menus. The take-up-weighted price-cost ratio uses the average marginal cost of workers who endogenously enroll and corresponds to the markup measure used in the model. Price relative to standardized marginal cost instead uses a menu-invariant, worker-type-weighted cost benchmark, thereby holding enrollee composition fixed. Both measures are weighted equally across menus within region and employer type and then by the observed mass of brokered employers across employer types and regions. The final two columns report the number of unique region-menu configurations and contributing regions. Singleton menus prices are subject to the model's singleton pricing cap: a 1.5 times price-to-cost ratio. The full-network category is excluded because the number of available insurers differs across regions.

Table A4. Robustness of the Estimated Hassle Costs

	(1) Baseline	(2) Lagged use	(3) CF-RE	(4) Switchers
(a). Estimates				
Mean broker-use wedge, $\widehat{\mu}$ (\$ PMPM)	68.5	27.9	99.7	24.4
	(33.5)	(22.1)	(51.2)	(52.0)
Persistence contribution (\$ PMPM)	—	67.7	69.7	—
Total index-equivalent wedge (\$ PMPM)	68.5	95.7	169.4	24.4
Coefficient on ΔV_{ft}	0.0141	0.0136	0.0057	0.0082
(b). Specification				
IV probit	Yes	Yes	No	Yes
Lagged broker use	No	Yes	Yes	No
Control-function residual	No	No	Yes	No
Firm random effect	No	No	Yes	No
Mundlak control	No	No	Yes	No
Switcher sample	No	No	No	Yes
(c). Sample and fit				
Employer-year observations	5,965	5,965	5,965	610
Employers	2,170	2,170	2,170	148
Observed broker-use rate	0.831	0.831	0.831	0.593
Predicted broker-use rate	0.835	0.833	0.816	0.605
First-stage F statistic	131.5	130.7	60.2	27.3

Notes: The table reports estimates of the employer broker-use wedge, expressed in dollars per member per month. Column (1) reports the baseline IV probit without lagged broker use. Column (2) adds lagged broker use. Column (3) reports a two-stage control-function random-effects probit with a Mundlak control for the firm-level mean of ΔV_{ft}. Column (4) estimates the lagged-use IV probit among firms observed both with and without a broker. The first row excludes the contribution of lagged broker use in Columns (2)-(4). Thus, excluding lagged use is a cost-accounting convention; these three models include lagged use during estimation. The second dollar-valued row adds the lagged-use contribution back to the reported index. In Column (3), the reported hassle-cost component also excludes the control-function residual, Mundlak term, and firm random effect. For brokered employers, V_f^{wb} is constructed using the observed broker. For non-brokered employers, V_f^{wb} is the equally weighted average across brokers operating in the employer's county and year. In all columns, V_f^{nb} is calculated using the full direct-market menu, including the no-offer option. The instrument is the firm leave-out mean ΔV generated by the relevant broker for its other actual clients. For non-brokered employers, the instrument is averaged across the same feasible broker set used to construct V_f^{wb}. All specifications include firm-size categories and industry, region, and year indicators. Standard errors derived from the delta method are reported.

Table A5. Prices per Enrollee under Partial Equilibrium Scenarios

	All employers			Brokered			Not brokered		
	Data	Model: observed menus	Model: simulated status quo	Data	Model: observed menus	Model: simulated status quo	Data	Model: observed menus	Model: simulated status quo
Panel A: Equilibrium outcomes									
Broker use (%)	85.4	81.8	70.3						
Not offering (%)	7.0	8.6	14.4	0.0	0.0	0.0	48.0	47.5	48.6
Take-up, offering firms (%)	75.9	75.1	74.2	75.9	74.3	72.4	75.9	82.7	82.5
Premium per enrollee	507.1	532.0	540.0	506.8	538.6	555.7	510.8	475.5	467.4
Commission (%)				3.32	3.22	3.33			
Insurers per menu				3.96	4.09	3.13			
Panel B: Market share by insurer (% of offering employers)									
Insurer A	26.4	26.4	27.1	25.8	25.6	26.4	32.5	34.0	30.5
Insurer B	19.7	14.0	19.0	21.3	14.5	20.2	1.3	9.1	13.4
Insurer C	13.9	13.9	13.4	13.5	13.8	13.4	18.2	14.7	13.2
Insurer D	9.9	10.6	9.5	9.7	10.4	9.3	11.7	11.7	10.4
Insurer E	11.6	10.2	11.2	11.9	10.3	11.9	7.8	9.2	8.3
Insurer F	6.7	11.0	7.0	6.4	11.5	6.5	10.4	6.7	9.3
Insurer G	4.7	6.1	4.2	4.7	6.5	4.1	3.9	3.2	4.5
Insurer H	2.1	2.6	2.8	2.1	2.4	2.7	2.6	3.5	3.2
Insurer I	2.4	2.5	2.8	2.5	2.4	2.9	1.3	3.1	2.5
Insurer J	2.7	2.8	3.0	2.0	2.6	2.6	10.4	4.6	4.7
Employers offering (N)	941	925	866	864	828	711	77	97	155
Panel C: Premium per enrollee by insurer									
Insurer A	576.2	575.7	580.8	576.2	586.5	599.2	576.5	506.0	507.6
Insurer B	413.0	541.1	493.2	413.4	552.8	507.8	389.5†	382.0	392.4
Insurer C	571.3	586.8	593.0	571.1	594.1	606.4	573.1	528.6	530.4
Insurer D	496.9	515.7	526.6	496.3	519.0	534.7	502.5	490.7	493.2
Insurer E	622.4	629.2	646.5	623.9	637.3	660.5	596.3	552.3	554.1
Insurer F	413.9	434.8	448.0	413.9	438.2	466.8	413.4	384.6	387.4
Insurer G	349.2	356.9	379.7	349.1	357.6	390.8	352.4†	345.0	332.4
Insurer H	440.3	465.4	466.7	434.4	471.4	475.6	441.1†	430.3	431.6
Insurer I	400.5	427.1	469.2	404.3	429.5	480.0	387.5†	410.9	412.7
Insurer J	382.6	431.0	447.5	398.1	443.0	476.3	349.7	373.3	373.6
Panel D: Insurer on the broker's menu (% of brokered employers)									
Insurer A				61.6	62.4	58.6			
Insurer B				30.3	30.8	23.4			
Insurer C				59.8	61.6	50.0			
Insurer D				61.6	64.2	44.2			
Insurer E				57.1	58.9	53.1			
Insurer F				33.3	34.1	15.9			
Insurer G				32.8	33.5	16.6			
Insurer H				44.1	46.3	35.8			
Insurer I				9.8	10.4	9.8			
Insurer J				6.0	6.8	5.7			

Table A6. Equilibrium Objects under Various Commission Caps

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Status	Fixed networks			Endogenous networks		
	quo	3%	2%	1%	3%	2%	1%
Employers							
Employer surplus	-12.15	-8.45	-3.79	2.15	-11.26	-15.40	-30.05
Choice value	17.41	19.58	22.41	26.12	18.52	18.36	18.87
Hassle costs	29.55	28.03	26.20	23.97	29.78	33.76	48.92
Broker use (%)	70.3	72.1	74.0	76.5	69.7	65.5	48.9
Not offering (%)	14.4	13.6	12.7	11.6	14.7	16.7	24.4
Employees							
Employee WTP	165.86	168.76	172.23	176.44	166.37	163.47	148.37
Take-up (%)	63.5	64.6	65.8	67.3	63.7	62.5	56.7
Insurers							
Premium per enrollee							
All employers	540.0	535.5	529.9	523.7	537.2	535.6	539.6
Brokered	555.7	549.5	541.8	533.5	552.9	554.4	576.4
Not brokered	467.4	465.3	463.4	461.1	467.3	466.5	472.2
Markup (p/mc)							
All employers	1.35	1.34	1.33	1.31	1.35	1.34	1.35
Brokered	1.39	1.38	1.36	1.33	1.39	1.39	1.44
Not brokered	1.18	1.18	1.18	1.18	1.18	1.18	1.18
Insurer profit	150.78	151.97	153.31	154.86	150.52	147.39	132.60
Brokers							
Insurers per menu	3.13	3.14	3.14	3.15	3.06	2.89	2.43
Commission (%)	3.33	2.72	1.95	1.00	2.68	1.93	1.00
Broker commissions	19.44	16.19	11.89	6.28	15.37	10.32	3.92
Broker contracting cost	8.03	8.03	8.03	8.03	6.98	5.45	2.98
Broker profit	11.40	8.16	3.86	-1.76	8.40	4.88	0.94

Table A7. The Effects of Commission Caps by Regions

	All	Upstate NY	Capital Region	New York City	Mid Hudson	Long Island	Western NY
A. Change in employer surplus, endogenous networks (N3–S0)							
3% cap	0.89	7.35	0.92	-0.31	-0.51	-0.80	-5.30
2% cap	-3.25	5.09	1.08	-3.82	-6.17	-5.74	-18.26
1% cap	-17.90	-13.44	-13.87	-17.62	-21.24	-18.00	-30.34
Best cap	3%	3%	2%	None	None	None	None
B. Change in employer surplus under the 3% cap, by step							
Fixed networks, F3–S0	3.70	1.70	1.48	5.20	4.61	3.10	2.21
Insurer choice, F1	0.23	0.00	0.03	0.38	0.26	0.24	0.06
Insurer pricing, F2	2.62	1.34	1.11	3.61	3.25	2.00	1.84
Broker use, F3	0.85	0.36	0.33	1.21	1.10	0.85	0.32
Network adjustment, N3–F3	-2.81	5.65	-0.56	-5.51	-5.12	-3.90	-7.51
Broker menus, N1	-2.01	2.98	-0.21	-3.26	-3.29	-3.30	-6.17
Insurer pricing, N2	-1.69	0.14	-0.36	-2.17	-1.55	-2.24	-4.06
Broker use, N3	0.89	2.53	0.00	-0.07	-0.28	1.64	2.72
Total changes, N3–S0	0.89	7.35	0.92	-0.31	-0.51	-0.80	-5.30
C. Market structure in the status quo							
Employers	1,012	213	64	460	78	111	86
Insurers in market		3	4	5	6	5	3
Commission dispersion (pp)		0.62	0.90	1.03	0.93	1.03	0.75
Average commission (% of premium)	3.33	3.49	2.81	3.32	3.42	3.26	3.48
Insurers per brokered menu	3.13	1.20	2.99	3.81	4.13	3.79	2.48
% of insurers in market		40	75	76	69	76	83
Contracting costs (% of commissions)	38	25	33	40	41	48	53
D. Broker networks and premiums							
Broker commissions, 3% cap, F3–S0 (%)	-16.7	-15.9	-8.8	-17.6	-17.3	-16.5	-17.1
Broker profits, 3% cap, F3–S0 (%)	-28.5	-21.6	-13.6	-30.8	-31.4	-34.4	-39.3
Insurers per menu, 3% cap	3.06	1.54	2.99	3.60	3.97	3.64	2.34
% of insurers in market		51	75	72	66	73	78
Insurers per menu, 2% cap	2.89	1.71	2.97	3.30	3.68	3.37	1.97
Insurers per menu, 1% cap	2.43	1.39	2.72	2.90	2.99	2.84	1.29
Premium change, 3% cap, N1 ($)	-3.0	-11.2	0.5	0.0	-1.7	-3.8	-1.5
Premium change, 3% cap, N2 ($)	7.2	-0.2	0.3	9.4	9.2	8.1	17.2
E. Other outcomes, 3% cap, N3–S0							
Broker use (pp)	-0.6	1.5	0.3	-0.7	-0.7	-1.5	-4.4
Not offering coverage (pp)	0.2	-0.9	-0.1	0.3	0.3	0.7	2.1
Employee WTP ($)	0.51	5.05	1.14	-0.64	-0.60	0.59	-4.13
Insurer profit ($)	-0.27	-1.47	0.21	0.85	-0.13	-1.46	-2.19
Broker profit ($)	-3.01	-2.52	-1.16	-3.72	-3.33	-2.63	-1.95

Table A8. Equilibrium Objects under Commission Rules, by Level

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	Status	Commission cap			Uniform commission rate								Fiduciary
	quo	3%	2%	1%	0.5%	1%	1.5%	2%	2.5%	3%	3.5%	4%	duty
Employers													
Employer surplus	-12.15	-11.26	-15.40	-30.05	-43.24	-30.05	-21.28	-14.78	-9.21	-4.81	-1.68	0.33	-10.77
Choice value	17.41	18.52	18.36	18.87	26.55	18.87	17.39	18.62	20.43	22.10	23.28	23.86	17.93
Hassle costs	29.55	29.78	33.76	48.92	69.79	48.92	38.67	33.40	29.64	26.91	24.97	23.53	28.70
Broker use (%)	70.3	69.7	65.5	48.9	24.9	48.9	60.4	65.9	69.5	71.9	73.4	74.5	71.2
Not offering (%)	14.4	14.7	16.7	24.4	35.3	24.4	19.1	16.5	14.7	13.5	12.7	12.2	14.0
Employees													
Employee WTP	165.86	166.37	163.47	148.37	130.57	148.37	158.69	163.94	167.02	169.37	170.67	170.78	167.14
Take-up (%)	63.5	63.7	62.5	56.7	49.8	56.7	60.7	62.7	63.9	64.9	65.4	65.5	64.0
Insurers													
Premium per enrollee													
All employers	540.0	537.2	535.6	539.6	520.1	539.6	537.7	535.1	534.9	532.8	532.4	535.7	537.7
Brokered	555.7	552.9	554.4	576.4	582.6	576.4	561.6	553.5	550.2	545.7	543.8	546.8	552.2
Not brokered	467.4	467.3	466.5	472.2	481.0	472.2	467.5	466.2	467.5	469.8	472.0	473.9	467.6
Markup (p/mc)													
All employers	1.35	1.35	1.34	1.35	1.31	1.35	1.35	1.34	1.35	1.34	1.35	1.35	1.35
Brokered	1.39	1.39	1.39	1.44	1.54	1.44	1.40	1.39	1.38	1.38	1.38	1.39	1.39
Not brokered	1.18	1.18	1.18	1.18	1.17	1.18	1.18	1.18	1.18	1.18	1.18	1.18	1.18
Insurer profit	150.78	150.52	147.39	132.60	103.24	132.60	143.53	147.13	148.70	149.00	148.98	149.29	149.22
Brokers													
Insurers per menu	3.13	3.06	2.89	2.43	1.73	2.43	2.71	2.93	3.13	3.31	3.46	3.56	3.06
Commission (%)	3.33	2.68	1.93	1.00	0.50	1.00	1.50	2.00	2.50	3.00	3.50	4.00	3.32
Broker commissions	19.44	15.37	10.32	3.92	0.94	3.92	7.38	10.78	14.19	17.62	21.01	24.34	19.52
Broker contracting cost	8.03	6.98	5.45	2.98	1.29	2.98	4.44	5.53	6.47	7.38	8.22	8.96	7.96
Broker profit	11.40	8.40	4.88	0.94	-0.35	0.94	2.94	5.26	7.72	10.24	12.79	15.38	11.56

Notes: Employers in the multi-insurer regions of New York, 2018. Column (1) reports the status quo; the other columns report long-run levels under a commission cap, a uniform commission rate, or a fiduciary duty, with insurer choice, prices, broker use and broker menus re-optimized, before broker exit. Dollar rows are per eligible employee per month. Premium per enrollee and markup are averages over all employers, over firms that use a broker, and over firms that do not; insurers per menu and the commission rate are averages among firms that use a broker; the remaining rows are averages over all employers.

Table A9. Equilibrium Objects under Commission Rules, by Level, with Broker Exit

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	Status	Commission cap			Uniform commission rate								Fiduciary
	quo	3%	2%	1%	0.5%	1%	1.5%	2%	2.5%	3%	3.5%	4%	duty
Employers													
Employer surplus	-12.15	-11.32	-15.79	-32.62	-43.10	-32.62	-22.46	-15.06	-9.26	-4.82	-1.68	0.34	-10.78
Choice value	17.41	18.95	20.22	29.42	44.61	29.42	22.20	20.09	20.92	22.28	23.37	23.90	18.13
Hassle costs	29.55	30.28	36.01	62.04	87.71	62.04	44.65	35.15	30.18	27.11	25.05	23.56	28.91
Broker use (%)	70.3	69.1	63.1	35.4	9.6	35.4	54.4	64.1	68.9	71.6	73.3	74.5	71.0
Not offering (%)	14.4	14.9	17.8	30.7	42.9	30.7	21.9	17.3	15.0	13.6	12.8	12.2	14.1
Employees													
Employee WTP	165.86	166.14	162.21	140.93	122.01	140.93	155.30	162.99	166.77	169.28	170.63	170.77	167.04
Take-up (%)	63.5	63.6	62.0	53.6	46.0	53.6	59.2	62.3	63.8	64.8	65.4	65.5	64.0
Insurers													
Premium per enrollee													
All employers	540.0	536.2	532.1	516.5	487.8	516.5	529.4	532.3	533.8	532.4	532.3	535.7	537.3
Brokered	555.7	551.9	551.5	555.5	533.5	555.5	555.4	551.1	549.3	545.3	543.7	546.7	551.9
Not brokered	467.4	467.7	467.9	475.8	478.7	475.8	469.6	467.3	467.8	469.9	471.9	473.8	467.7
Markup (p/mc)													
All employers	1.35	1.35	1.34	1.30	1.23	1.30	1.33	1.34	1.34	1.34	1.34	1.35	1.35
Brokered	1.39	1.39	1.38	1.42	1.49	1.42	1.39	1.38	1.38	1.38	1.38	1.38	1.39
Not brokered	1.18	1.18	1.18	1.18	1.18	1.18	1.18	1.18	1.18	1.18	1.18	1.18	1.18
Insurer profit	150.78	149.67	143.82	111.59	75.81	111.59	134.44	144.36	147.80	148.66	148.84	149.23	148.88
Brokers													
Insurers per menu	3.13	3.06	2.90	2.44	1.71	2.44	2.73	2.94	3.14	3.31	3.46	3.56	3.06
Commission (%)	3.33	2.69	1.93	1.00	0.50	1.00	1.50	2.00	2.50	3.00	3.50	4.00	3.32
Broker commissions	19.44	15.27	9.97	2.85	0.36	2.85	6.67	10.49	14.08	17.57	20.99	24.33	19.47
Broker contracting cost	8.03	6.86	5.10	1.83	0.26	1.83	3.77	5.24	6.36	7.32	8.19	8.95	7.91
Broker profit	11.40	8.40	4.87	1.01	0.09	1.01	2.90	5.25	7.72	10.25	12.80	15.38	11.56
Share active (%)	100.0	97.2	88.1	37.1	13.5	37.1	70.7	90.6	97.1	98.9	99.6	99.8	99.0

Notes: As Table A8, after broker exit: brokers whose commission revenue falls below the exit cutoff leave the market, and their clients face the no-broker options. Exit is an expected share under the contracting-cost draws. "Share active" is the share of brokers that remain; it is 100 at the status quo, where no exit is applied.

Table A10. Equilibrium Objects under Access Rules, by Level

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Status	Minimum menu size (number of insurers)					Fiduciary duty + minimum menu size				
	quo	2	3	4	5	6	2	3	4	5	6
Employers											
Employer surplus	-12.15	-0.98	7.83	12.63	28.48	30.17	0.82	10.36	16.29	33.83	35.51
Choice value	17.41	22.55	25.80	28.64	39.91	41.21	23.54	27.45	31.16	44.06	45.37
Hassle costs	29.55	23.53	17.97	16.01	11.43	11.03	22.72	17.09	14.87	10.23	9.86
Broker use (%)	70.3	74.0	77.8	80.0	86.5	87.0	74.9	78.7	81.3	87.9	88.4
Not offering (%)	14.4	12.3	10.3	9.2	6.4	6.1	11.9	9.8	8.6	5.7	5.5
Employees											
Employee WTP	165.86	170.35	176.07	178.44	185.68	186.20	171.90	177.91	180.75	188.30	188.82
Take-up (%)	63.5	65.3	67.7	68.6	71.3	71.5	65.9	68.4	69.5	72.3	72.5
Insurers											
Premium per enrollee											
All employers	540.0	538.3	525.7	522.5	519.6	519.4	535.2	521.8	517.8	514.5	514.2
Brokered	555.7	549.6	530.9	526.4	522.7	522.4	545.4	526.1	520.7	516.9	516.5
Not brokered	467.4	476.8	491.3	493.6	482.7	482.3	477.1	492.5	494.2	482.1	481.7
Markup (p/mc)											
All employers	1.35	1.37	1.34	1.33	1.32	1.32	1.36	1.33	1.32	1.31	1.30
Brokered	1.39	1.40	1.36	1.35	1.33	1.33	1.39	1.35	1.34	1.31	1.31
Not brokered	1.18	1.18	1.17	1.17	1.17	1.17	1.18	1.17	1.17	1.17	1.17
Insurer profit	150.78	151.86	152.44	152.05	151.10	151.04	149.16	149.17	148.25	146.17	146.04
Brokers											
Insurers per menu	3.13	3.42	3.50	3.68	4.31	4.39	3.35	3.43	3.67	4.31	4.39
Commission (%)	3.33	3.24	3.25	3.24	3.21	3.21	3.22	3.24	3.22	3.16	3.16
Broker commissions	19.44	19.52	20.54	20.99	22.58	22.72	19.56	20.58	21.07	22.36	22.48
Broker contracting cost	8.03	8.53	10.44	11.31	15.48	16.17	8.40	10.35	11.33	15.49	16.17
Broker profit	11.40	10.99	10.10	9.69	7.10	6.55	11.16	10.23	9.74	6.88	6.31

Notes: Employers in the multi-insurer regions of New York, 2018. Column (1) reports the status quo; the other columns report long-run levels under a minimum menu size, alone or combined with a fiduciary duty, with insurer choice, prices, broker use and broker menus re-optimized, before broker exit. The minimum is capped at the number of insurers active in a region. Dollar rows are per eligible employee per month. Premium per enrollee and markup are averages over all employers, over firms that use a broker, and over firms that do not; insurers per menu and the commission rate are averages among firms that use a broker; the remaining rows are averages over all employers.

Table A11. Equilibrium Objects under Access Rules, by Level, with Broker Exit

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Status	Minimum menu size (number of insurers)					Fiduciary duty + minimum menu size					No
	quo	2	3	4	5	6	2	3	4	5	6	brokers
Employers												
Employer surplus	-12.15	-1.06	3.61	7.49	19.10	19.22	0.71	5.88	10.83	23.44	23.37	-44.74
Choice value	17.41	23.26	28.19	31.38	42.39	43.48	24.18	29.65	33.69	45.99	47.01	56.15
Hassle costs	29.55	24.32	24.59	23.89	23.29	24.26	23.47	23.77	22.86	22.55	23.64	100.89
Broker use (%)	70.3	73.3	73.1	74.1	75.0	73.8	74.2	74.0	75.3	75.9	74.5	0.0
Not offering (%)	14.4	12.7	12.8	12.3	11.9	12.5	12.2	12.4	11.8	11.5	12.1	48.0
Employees												
Employee WTP	165.86	170.09	172.31	173.55	176.36	175.55	171.64	174.00	175.65	178.32	177.39	114.85
Take-up (%)	63.5	65.2	66.0	66.5	67.5	67.2	65.8	66.7	67.4	68.3	67.9	43.0
Insurers												
Premium per enrollee												
All employers	540.0	536.9	526.1	523.2	516.8	515.9	533.9	522.5	518.7	511.9	510.9	474.7
Brokered	555.7	548.7	535.9	532.3	524.6	523.8	544.6	531.2	526.6	518.6	517.7	—
Not brokered	467.4	475.2	475.2	473.4	471.7	473.0	475.6	475.4	472.6	471.4	473.0	474.7
Markup (p/mc)												
All employers	1.35	1.36	1.33	1.32	1.31	1.30	1.36	1.32	1.31	1.29	1.29	1.18
Brokered	1.39	1.40	1.36	1.35	1.33	1.33	1.39	1.35	1.33	1.31	1.31	—
Not brokered	1.18	1.18	1.18	1.18	1.18	1.18	1.18	1.18	1.18	1.18	1.18	1.18
Insurer profit	150.78	150.64	146.65	145.17	138.90	137.38	148.05	143.67	141.65	134.26	132.71	59.29
Brokers												
Insurers per menu	3.13	3.44	3.53	3.72	4.33	4.38	3.36	3.46	3.70	4.33	4.38	—
Commission (%)	3.33	3.24	3.25	3.24	3.21	3.21	3.23	3.24	3.22	3.15	3.15	—
Broker commissions	19.44	19.37	19.51	19.72	19.68	19.32	19.42	19.56	19.80	19.39	18.98	—
Broker contracting cost	8.03	8.34	8.90	9.45	11.36	11.35	8.22	8.82	9.47	11.26	11.22	—
Broker profit	11.40	11.04	10.61	10.27	8.32	7.96	11.20	10.74	10.33	8.13	7.76	—
Share active (%)	100.0	98.2	94.4	92.8	84.1	81.2	98.4	94.5	92.7	83.6	80.5	—

Notes: As Table A10, after broker exit, as in Table A9. The last column removes brokers entirely: every employer faces the no-broker options, no broker is active, and the rows defined among brokered firms and the broker rows are left blank.

B. Additional Details on Data

B1. Panel Construction Using Form 5500 Filings

Panel Construction. Below, we describe the data cleaning steps used to construct the final analytic dataset from Form 5500 filings and related sources.

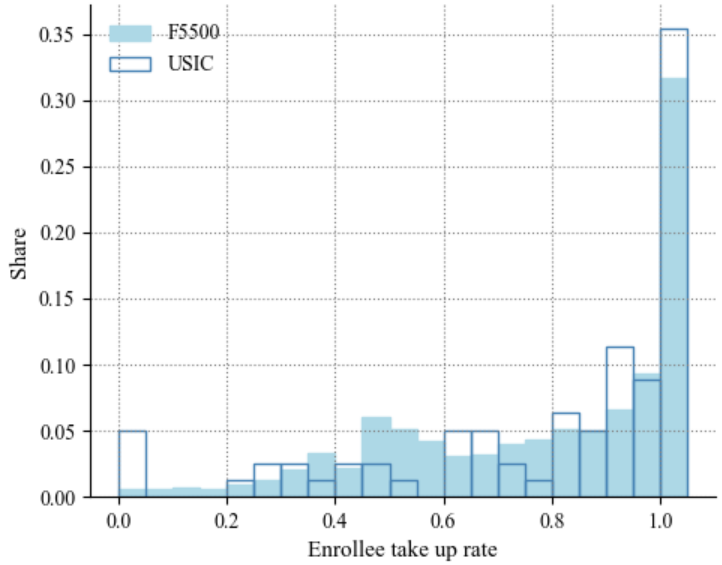
First, we begin with Schedule A filings of Form 5500, which records employers' benefit contracts. Each filing reports the carrier, premiums paid, and the number of covered enrollees (employees and dependents) for each plan. Second, we restrict attention to health insurance filings. Other benefit types, such as life, disability, and unemployment insurance, are excluded. Third, we clean and standardize the carrier identifiers. Schedule A provides insurers' NAIC codes, which we use to identify health insurers. We harmonize these codes and consolidate carriers under common parent companies. Fourth, we limit the sample to New York employers. Fifth, we resolve multiple forms within a reporting-carrier-employer-year. If the form is an amendment, only the final submission is retained. If several forms correspond to different products under the same carrier, we collapse them into a single record. Sixth, we incorporate information on brokers. Schedule A Part I reports broker names, locations, and commission payments. We standardize broker names and merge them to the main Schedule A record at the reporting-carrier-employer-year level. Seventh, we merge the sample with the main Form 5500 database, which adds the employer-year level information, including the total number of enrolled employees that aggregates across carriers, industry, and location. We aggregate to one observation per employer–year–carrier, applying the same collapsing rules as above. If multiple brokers are associated with a filing, we retain the broker accounting for the largest share of premiums. Finally, we augment the dataset with Form 5500 group health plan information from the Bureau of Labor Statistics, which provides a self-insured indicator. We also bring in the BLS-recovered total number of enrolled employees at the employer-year level, aggregating across carriers when multiple carriers are offered.

The resulting dataset is structured at the employer–year–realized insurer level. Each record contains the insurance carrier contracted by the employer, the broker through which the contract was obtained, the number of enrollees (employees plus dependents), premiums paid, the number of enrolled employees (unlike enrollees or premiums, this is an aggregate across carriers), and employer characteristics.

Takeup Rate. We next describe how we measure firm size and insurance take-up. Form 5500 filings contain records on both welfare benefits (including medical, drug, dental, vision, disability, and life insurance, and others) and pension benefits (including defined benefits plans, defined contribution plans, and pension plans). We use these filings to proxy for firm size by taking the maximum number of enrolled employees across all reported welfare and pension plans for each employer, assuming that every employee takes up at least one benefit the employer offers. Because not all employees necessarily take up the benefits offered, this measure may underestimate firm sizes.

We then compute the take-up rate of employer-sponsored insurance as the ratio of enrolled employees in health insurance benefits to our imputed firm size. We benchmark our imputed take-up rates against two data sources. First, the average health insurance take-up rate in our sample is 79%, which is close to 71% for firms of similar sizes in the Medical Expenditure Panel Survey. Second, we compare our imputed take-up rates for the large national insurer against those reported in its records, which provide exact firm size and

Figure A18. Benchmark Takeup-Rate Across Data Sources



Notes: This figure plots imputed takeup-rate for tha large national insurer in Form 5500 and its accurate enrollment records.

enrollment information. Figure A18 shows our imputed takeup rate distribution for that large representative insurer closely resembles the true takeup measure. This similarity provides support for the validity of our firm size and take-up rate measures.

Household Characteristics A limitation of the Form 5500 data is that, while we observe the number of employees who enroll in coverage through the firm, we do not observe the demographics of those employees. We thus constructed a simulated sample of employee demographics relying on distributions from an external data source: the Medical Expenditures Panel Survey (MEPS). The MEPS Household Component (HC) is a nationally representative sample of households and contains information on demographic characteristics and health insurance coverage information. Using MEPS, we identify individuals with employer-sponsored insurance through their own employer and construct demographic distributions of age, gender, and household size by firm size and industry, which we then sample into the Form 5500 dataset.

Premiums The Form 5500 contains the total premiums earned for each employer, which we use as our primarily variable for constructing the prices faced by employers (per employee per month) for each insurance plan. As the Form 5500 data only contains *realized* premiums (i.e. premiums conditional on firms having chosen the insurer), we do not directly observe premiums of the full choice set facing the employer. Thus, we rely in imputation methods to simulate the likely premiums that a firm would face for an unobserved plan.

To do so, we implement a hedonic premium regression in the spirit of Dafny et al. (2013). Specifically, for our New York sample between 2011-2020, we regress the log of premiums per enrollee on a rich set of observables, including insurer fixed effects, geographic region, firm size, year, and industry. We also leverage broker network information in the regression by including indicators for the number of insurers in a broker's network for the specific employer. Given that a key factor we are exploring is the competitive effects of broker-insurer networks, we expect premiums for the same insurer and same employer to vary

depending on the broker.

B2. Data from the Large National Representative Insurer

We supplement our analysis with data from a large health insurance provider. We focus on the data for New York State. The data provides information on health insurance coverage for employers in the small group market. The data cover the years 2013 through 2017. We observe data at the employer-year and enrollee-year levels as well as linked enrollee claims data. The employer-year level data include the number of employees eligible for health insurance coverage, the number of employees enrolled, the number of plans offered by the employer, characteristics of each plan, including plan type (EPO, PPO, POS), coinsurance, physician copay, in-network deductible, in-network out-of-pocket maximum, out-of-network deductible, and out-of-network out-of-pocket maximum, and the total premium paid to the insurer for each plan. The enrollee-year data include age, gender, and health plan choice from the employer's offerings. We calculate a household's premium by multiplying the enrollee's premium amount by the number of enrollees in each household.

One limitation of the data is that while we know the *number* of eligible employees who do not take up offered coverage, we do not know any of their characteristics. Thus, we supplement the data again with demographics from the MEPS HC, as with the Form 5500 sample above. Here, we identify individuals who worked for a small group employer and were offered health insurance from that employer.

Table A12 reports the sample statistics at the enrollee level. The first two columns represent all firms, while the last two columns represent firms with multiple plans. We include the distribution of employees who take up coverage at the firm (i.e., those covered in the large national insurers' records) and those who opt for the outside option (i.e., calibrated from the Medical Expenditure Panel Survey). Employees who do not participate in the employer-sponsored plan tend to be older on average, are less likely to be men, and are more likely to have a household size of one. There is no clear pattern of firm size differences, except that they are less likely to be 100-499 firms, and more likely to be smaller. Comparing the single-plan and multi-plan firms, the age distribution and family size distribution appear similar, but multi-plan firms are clearly larger. Enrollment-weighted plan characteristics, such as premiums and deductibles, are also similar.

Table A13 reports sample statistics at the firm level. Multi-plan firms are bigger and have slightly higher take-up rates. In general, approximately 60% of all firms have 1-49 employees, 30% have 50-99 employees, and 10% have more than 100 employees. However, for multi-plan firms, the rates are 10%, 60%, and 30%, respectively.

B3. Clarivate Data and Imputed Product Characteristics

We derive insurance product characteristics from the Clarivate Real-World Data (RWD) Product, previously known as Decision Resources Group Real-World Data. This dataset covers over 300 million longitudinal patient records in the United States, constructed from medical claims submitted through billing software used by a wide array of healthcare providers, including hospitals, physician offices, long-term care facilities, etc. The raw dataset contains billions of patient claims with detailed information on providers, patients, and payers, and a granular breakdown of patient payments such as deductibles, coinsurance, and copayments. In addition, the data report full diagnosis information in terms of ICD codes, medical procedures via HCPCS

Table A12. Summary Statistics by Coverage Decision and Firm Type

	All Firms				Multi-Plan Firms			
	(1)		(2)		(3)		(4)	
	Inside		Outside		Inside		Outside	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
A. Demographics:								
Age 18-34	0.623	0.485	0.487	0.500	0.619	0.486	0.479	0.500
Age 35-54	0.283	0.451	0.339	0.473	0.293	0.455	0.342	0.474
Age 55-64	0.074	0.262	0.111	0.314	0.070	0.255	0.114	0.317
B. Firm Size:								
Firm Size: 1-49 employees	0.063	0.243	0.082	0.274	0.019	0.136	0.032	0.176
Firm Size: 50-99 employees	0.497	0.500	0.439	0.496	0.507	0.500	0.523	0.499
Firm Size: 100-499 employees	0.437	0.496	0.303	0.460	0.474	0.499	0.445	0.497
Firm Size: 500+ employees	0.003	0.057	0.176	0.381	0.000	0.000	0.000	0.000
Percent Male	0.526	0.409	0.436	0.496	0.535	0.415	0.440	0.496
Household Size	1.874	1.316	1.000	0.000	1.818	1.293	1.000	0.000
C. Plan Characteristics:								
Monthly Premium	103.377	77.757			107.285	75.152		
Deductible	566.032	920.844			634.305	974.874		
Out-of-Pocket Maximum	2407.116	2019.124			2637.779	2024.493		
Coinsurance Rate	0.682	0.432			0.719	0.410		
EPO	0.329	0.470			0.375	0.484		
POS	0.661	0.473			0.618	0.486		
PPO	0.009	0.096			0.007	0.083		

Notes: Summary statistics at the household level. Inside option refers to households enrolled in employer coverage, outside option to those declining coverage. Take-up rate is 59.9% for all firms and 76.2% for firms offering multiple plans.

Table A13. Firm Characteristics Summary Statistics

	All Firms		Multi-Plan Firms	
	Mean	SD	Mean	SD
A. Firm Size:				
Number of Enrollees	34.5	44.4	69.7	48.8
Number Eligible	45.5	65.6	81.9	38.8
B. Size Distribution:				
1-49 employees	0.6	0.5	0.1	0.3
50-99 employees	0.3	0.5	0.6	0.5
100-499 employees	0.1	0.3	0.3	0.4
500+ employees	0.0	0.1	0.0	0.0

Notes: Summary statistics at the firm-year level. N = 590 total firms, of which 194 offer multiple plans.

codes, patient demographics (e.g., 3-digit ZIP code, gender, age), and insurer type (commercial, Medicare, or Medicaid). A key feature of the dataset is its open-network structure, which enables us to track patients over time and across providers, even when patients switch insurers within the observation period.

We limit our analysis sample to New York and include all commercial claims, which encompass both professional and institutional claims, and cover patient visits to hospitals (inpatient and outpatient), ambulatory surgical centers (ASCs), clinics, and physician offices. Starting from the raw claims data, we restrict the sample to patients aged between 18 and 64 and exclude individuals enrolled in privately administered Medicaid or Medicare plans. We further require all claims to have non-missing information on patient gender, year of visit, total allowed amounts, and detailed out-of-pocket components, including deductibles and coinsurance. Because most health plans do not count preventative care toward deductibles or maximum

out-of-pocket limits, we exclude claims associated with preventative care procedures to avoid bias. Our final analysis sample focuses on the four largest commercial payers in New York State—Aetna, Connecticut/Cigna, Empire/Anthem, and Oxford/United—while pooling all other relatively smaller insurers into a single category.[1] labeled as "Other Insurers." In total, the sample comprises more than 12.5 million commercial claims spanning from 2013 to 2022.

To infer plan characteristics across carriers and regions, we aggregate claims for each patient within a calendar year, summing allowed amounts, deductibles, coinsurance, and other out-of-pocket costs. Although the dataset does not provide the start and end dates of annual health plan cycles, we use the calendar year as a proxy. Deductibles accrue until a patient reaches the plan's annual maximum deductible, at which point the insurer begins covering subsequent medical expenses (with patients still responsible for coinsurance or copayments). To ensure that we are observing patients who likely reached their deductible limits, we focus on those with annual medical expenditures exceeding $6,000. For carrier-region-year cells with fewer than 20 such patients, we relax this threshold to $2,000 to preserve sample size. We define a cell as the interaction of a carrier, region, and year. Regions are based on the ten economic development regions designated by the New York State Department of Economic Development: Western NY, Finger Lakes, Southern Tier, Central NY, Mohawk Valley, Capital Region, Mid-Hudson, New York City, Long Island, and North Country (https://esd.ny.gov/regions). Each patient is mapped to a region according to their ZIP code.

Within each cell, the inferred deductible is calculated as the average deductible paid by patients whose annual medical spending exceeds $6,000. The inferred coinsurance rate is defined as one minus the mean ratio of patients' total coinsurance payments to their total medical expenses net of deductibles, thereby capturing the average share of costs borne by insurers. Summary statistics of the inferred deductibles and coinsurance rates are presented in Table A14.

Table A14. Plan Characteristics

		Sample Mean					
	N	Patient Age	Female	Patient Paid	Insurer Paid	Deductible	Coinsurance (%)
Oxford/United	2,784,846	45.0	0.60	266.1	574.4	702.2	84.8
Aetna	730,811	42.6	0.58	285.9	416.9	996.8	88.4
Connecticut/Cigna	527,818	41.2	0.60	218.0	322.0	1088.2	86.0
Empire/Anthem	472,644	45.3	0.58	399.7	546.8	929.8	70.3
Other Insurers	1,696,603	44.7	0.61	123.3	432.0	817.9	84.4

[1]The cleaned sample provides relatively limited coverage of Excellus. As a result, we combine Excellus with the group of other insurers and instead single out the next largest insurer, Empire/Anthem, as a separate category.

C. Theoretical Derivations and Microfoundation

This appendix collects the derivations underlying the conceptual framework of Section 3. We present, in turn, the broker profit slopes and menu thresholds without and with steering, the welfare decomposition, and the qualifications omitted from the main text.

C1. Strategic Commissions and Equilibrium Existence

In the main text, we treat commissions as exogenous. If insurers were instead allowed to set commissions strategically before brokers chose menus, they would compete by raising commissions until one insurer's net profit is driven to zero. This logic is in a similar spirit to manufacturers paying lump-sum transfers to retailers to carry their products (Shaffer, 2005; Hristakeva, 2022). Such settings, however, raise concerns about equilibrium existence when upstream firms are not homogeneous or single-product, because best responses may be discontinuous (Rey and Vergé, 2010, 2016). We therefore hold commissions fixed and study the effect of exogenous commission caps.

C2. Menu Formation Without Steering

We start by considering the baseline case without steering ($\rho = 0$). The profit from offering only A is unaffected by B's commission, while the profits from offering A+B or only B increase in γ_B at rates determined by the share of employers choosing B under combined or exclusive contracting, respectively:

$$\frac{\partial \Pi_A}{\partial \gamma_B} = 0, \quad \frac{\partial \Pi_{AB}}{\partial \gamma_B} = s_B^c, \quad \frac{\partial \Pi_B}{\partial \gamma_B} = s_B^m.$$

Our assumptions on employer preferences (equation (1)) imply that the profit line for offering only B is steeper than that for offering both products, while the intercept of offering both exceeds that of offering only B:

$$\frac{\partial \Pi_B}{\partial \gamma_B} > \frac{\partial \Pi_{AB}}{\partial \gamma_B}, \qquad \Pi_{AB}\big|_{\gamma_B=0} > \Pi_B\big|_{\gamma_B=0}.$$

The broker chooses the menu that maximizes its profit, corresponding to the upper envelope of the three profit lines. When γ_B is low, the broker offers only A. As γ_B rises past γ_B^*, it expands the menu to A+B; and at a higher threshold γ_B^{**}, expected revenue from B dominates and the broker contracts only with B, forgoing the contracting cost of A and further raising B's monopoly share. Hence, a commission reduction can either expand the menu, from B only to A+B when crossing γ_B^{**} from above, or shrink it, from A+B to A only when crossing γ_B^* from above, depending on the baseline commission.

C3. Menu Formation With Steering

We next analyze the case with steering toward B, $\rho > 0$. Because steering reduces the perceived utility of non-steered options, the employer's perceived flow utility when B is offered is

$$V_A^c = \delta_A - \alpha p_A^c - \rho \gamma_B + \epsilon_A; \quad V_U^l = \delta_0 - \rho \gamma_B + \epsilon_0, \; l \in \{m, c\}.$$

When B is absent from the menu, the perceived utilities of A and U revert to their baseline. Since we hold γ_A fixed, γ_B can equivalently be interpreted as the commission differential $\gamma_B - \gamma_A \geq 0$. Denoting variables under steering with tildes, the slopes of broker profits with respect to B's commission become

$$\frac{\partial \tilde{\Pi}_{AB}}{\partial \gamma_B} = \tilde{s}^c_B + \rho \tilde{s}^c_B (1 - \tilde{s}^c_B)\gamma_B - \rho \tilde{s}^c_A \tilde{s}^c_B \gamma_A > \tilde{s}^c_B > s^c_B = \frac{\partial \Pi_{AB}}{\partial \gamma_B},$$

$$\frac{\partial \tilde{\Pi}_{B}}{\partial \gamma_B} = \tilde{s}^m_B + \rho \tilde{s}^m_B (1 - \tilde{s}^m_B)\gamma_B > \tilde{s}^m_B > s^m_B = \frac{\partial \Pi_{B}}{\partial \gamma_B}.$$

The first inequality in the $\partial \tilde{\Pi}_{AB}/\partial \gamma_B$ expression holds because $\rho \tilde{s}^c_B (1 - \tilde{s}^c_B)\gamma_B - \rho \tilde{s}^c_A \tilde{s}^c_B \gamma_A = \rho \tilde{s}^c_B (1 - \tilde{s}^c_B)(\gamma_B - \gamma_A) + \rho \tilde{s}^c_U \tilde{s}^c_B \gamma_A > 0$. Thus, steering increases the responsiveness of broker profits to commissions relative to the no-steering case, effectively steepening both slopes. This widens the range of commissions over which B is offered, and a higher steering ability ρ further enlarges this range.

The ordering in Figure 4b should be interpreted locally. With steering, the shares $\tilde{s}^c_B$ and $\tilde{s}^m_B$ vary with γ_B, so $\tilde{\Pi}_{AB}$ and $\tilde{\Pi}_B$ need not be globally linear in γ_B. The figure illustrates the range in which $\tilde{s}^m_B < (1 + \rho\gamma_B)/(2\rho\gamma_B)$ and γ_A is sufficiently small so that $\frac{\partial \tilde{\Pi}_B}{\partial \gamma_B} > \frac{\partial \tilde{\Pi}_{AB}}{\partial \gamma_B}$. Under this ordering, the broker moves from offering A, to offering A+B, and then to offering B as γ_B rises. At sufficiently high values of γ_B, the ordering can reverse, so that $\frac{\partial \tilde{\Pi}_B}{\partial \gamma_B} < \frac{\partial \tilde{\Pi}_{AB}}{\partial \gamma_B}$, and the broker may eventually move from B only back to A+B. This far-right region does not affect the local comparative statics emphasized in the main text, which concern commission changes around the thresholds at which B enters the menu and A is dropped. Under this local ordering, steering widens the range of commissions over which B is offered, and a higher steering ability ρ further enlarges this range.

C4. Welfare Decomposition

We now consider employer surplus under each menu, which follows a standard log-sum format (ignoring the constant term):

$$\begin{aligned}
\text{CS}_{A+B} &= \ln\left(\exp(\delta_A - \alpha p^c_A - \rho\gamma_B) + \exp(\delta_B - \alpha p^c_B) + \exp(\delta_0 - \rho\gamma_B)\right), \\
\text{CS}_{A} &= \ln\left(\exp(\delta_A - \alpha p^m_A) + \exp(\delta_0)\right), \\
\text{CS}_{B} &= \ln\left(\exp(\delta_B - \alpha p^m_B) + \exp(\delta_0 - \rho\gamma_B)\right).
\end{aligned}$$

Comparing the A-only and B-only menus is straightforward: surplus is higher under A only, both because employers prefer A to B ($\delta_A > \delta_B$) and because the B-only menu imposes the steering penalty $-\rho\gamma_B$ on non-steered options. The A+B versus A comparison in equation (2) is more nuanced, splitting into a positive competitive effect and a negative steering effect.

The competitive effect can be further decomposed into two components:

$$\underbrace{\ln\left[\frac{\exp(\delta_A - \alpha p^c_A) + \exp(\delta_0)}{\exp(\delta_A - \alpha p^m_A) + \exp(\delta_0)}\right]}_{\text{Price Competition } (+)} + \underbrace{\ln\left[\frac{\exp(\delta_A - \alpha p^c_A) + \exp(\delta_B - \alpha p^c_B) + \exp(\delta_0)}{\exp(\delta_A - \alpha p^c_A) + \exp(\delta_0)}\right]}_{\text{Product Variety } (+)} > 0.$$

The first term represents the price competition effect, namely the welfare gain from A's price falling from monopoly to competitive levels; the second represents the product variety effect, namely the welfare gain from adding B to the choice set. Both components are always positive. The steering effect can be rewritten as

$$\ln\left(1-(s_A^c+s_U^c)\left(1-\exp(-\rho\tilde{\gamma}_B^*)\right)\right)<0,$$

where s_A^c, s_U^c are the shares of employers choosing A or U without steering. This expression shows that the welfare loss from steering is increasing in the steering incentive $\tilde{\gamma}_B^*$, the steering ability ρ, and the baseline share of employers preferring a non-steered option, $s_A^c + s_U^c$. The steering effect is always negative, representing a pure deadweight loss from distorted choices.

C5. Extensions

Finally, we note robustness to several model extensions. The qualitative predictions are unchanged if the broker receives percentage-of-premium rather than flat commissions. They also hold if steering operates in both directions rather than only toward B, in which case one would rescale the brand fixed effects to obtain similar derivations.

C6. Micro-foundation of the Steering Specification

Equations (7) and (8) in the main text specify employer plan choice as a single *joint* decision rule: the weight α_{fj} that the employer attaches to its own premium already embeds the broker's commission-driven steering incentive through the term $\mathbb{1}_{fb}\,\alpha_{2,b(f)}\,\gamma_j$. The main text adopts that formulation for tractability.

This appendix provides a microfoundation for the reduced-form formulation. We show that a model with two separately optimizing agents reproduces the joint rule, and we use the model to assign structural content to the steering parameter $\alpha_{2,b(f)}$. Our microfoundation rationalizes the sign and interpretation of the steering term, while making explicit where the empirical approximation enters.

Setup, timing, and information We focus on the *ex-post recommendation stage* of the game. By this stage: (i) the broker's menu m has been chosen in the ex-ante network-formation stage (Section 5.4) and is taken as given, together with the associated contracting costs, which are sunk; (ii) employer f has already decided to engage the broker, so $\mathbb{1}_{fb}=1$; and (iii) premiums $\{p_{fjm}\}_{j\in m}$ and the preference and cost shocks (ξ_{fj},ϵ_{fj}) are realized. We use the qualifier *ex-post* deliberately, to distinguish this recommendation problem from the ex-ante menu problem of Section 5.4, in which the broker maximizes *expected* payoffs over the distribution of shocks before committing to m.

Both the employer and the broker observe the employer's match value u^E_{fjm} defined below for every plan $j\in m$; the econometrician observes neither ξ_{fj} nor ϵ_{fj}. The broker issues a single recommendation $r_f\in m$, and the employer acts on it.

Assumption 1 (Compliance) *An employer that engages a broker chooses the plan the broker recommends.*

Assumption 1 is the standard delegation assumption: having chosen to rely on the broker's expertise, the employer adopts the broker's advice. It is the natural limiting case of a model in which the broker's informational advantage (i.e., a more precise signal of plan-employer match quality or a lower cost of processing

plan and network information) makes its recommendation pivotal, and it isolates the broker's incentives as the operative force behind observed choices. The disciplining force on the broker that excessive steering lowers the value of broker services and deters engagement operates at the prior stage (Section 5.3) and is preserved.

The case $\mathbb{1}_{fb} = 0$ is degenerate: absent a broker there is no recommendation problem and the employer maximizes its own objective directly. The microfoundation therefore focuses on $\mathbb{1}_{fb} = 1$; we nonetheless carry the indicator $\mathbb{1}_{fb}$ through the algebra so that the resulting expressions nest both cases.

The employer's objective Independent of the broker's commission motive, the employer f ranks plans by the objective

$$u^E_{fjm} = -\alpha^E_{fj}(1-\iota_f)\, p_{fjm}\, \text{takeup}_{fjm} \;+\; \rho_f\, W_{fjm} \;+\; \delta^f_j + \xi_{fj} + \epsilon_{fj}, \tag{A1}$$

with objective weight $\alpha^E_{fj} = \alpha_0 + \mathbb{1}_{fb}\,\alpha_{1,b(f)}$. This is equation (7) of the main text with the welfare weight restricted to $\alpha_0 + \mathbb{1}_{fb}\alpha_{1,b(f)}$, that is, equation (8) *without* the steering term $\mathbb{1}_{fb}\alpha_{2,b(f)}\gamma_j$. All notation is as in the main text: $(1-\iota_f)\, p_{fjm}\, \text{takeup}_{fjm}$ is the employer's premium contribution per employee, ρ_f the (broker-invariant) weight the employer places on employee surplus, W_{fjm} the average ex-ante employee surplus (evaluated at the equilibrium premium p_{fjm} and take-up rate takeup_{fjm}), δ^f_j brand perception, and ξ_{fj}, ϵ_{fj} the structural and idiosyncratic shocks.

The two components of α^E_{fj} have distinct interpretations. First, α_0 is the employer's pre-broker valuation of a dollar of its own premium outlay—the weight that governs choices when no broker is involved. Second, $\mathbb{1}_{fb}\alpha_{1,b(f)}$ is the *non-conflicted* refinement that a broker's expertise brings to the trade-off between own expenses and employee surplus, for example by reducing the informational frictions the employer faces in assessing how well plan j matches its workforce. The defining property of u^E_{fjm} is that it carries *no commission term*: it is the objective that a broker with no commission motive—an interest-aligned agent—would pursue on the employer's behalf.

The broker's ex-post recommendation problem The broker is not a perfect agent of the employer. From recommending plan j to employer f, the broker receives an ex-post payoff equal to its commission income plus a weight on the employer's objective:

$$\Pi^B_{fjm} \;=\; \Pi^{\text{com}}_{fjm} \;+\; \theta_{b(f)}\, u^E_{fjm}, \qquad \theta_{b(f)} \geq 0. \tag{A2}$$

Commission income. Under the percentage commissions that govern small- and mid-sized groups, the broker earns a fraction γ_j of the premium. Per employee, commission income is

$$\Pi^{\text{com}}_{fjm} \;=\; \gamma_j\, b_{fjm}, \qquad b_{fjm} = p_{fjm}\, \text{takeup}_{fjm}, \tag{A3}$$

where b_{fjm} is the *commission base*, namely the premium per employee.[2] Aggregating the ex-post realiza-

[2]Commissions are levied on the full premium, of which the employer pays the share $1-\iota_f$ and employees the remainder. The commission base b_{fjm} is therefore the employer's own premium-contribution variable in equation (7) rescaled by the constant $(1-\iota_f)^{-1}$, a fact used in the proof of Proposition 1.

tions Π^{com}_{fjm} across employers and integrating over shocks yields the expected commission income $\Pi_b(m)$ that the broker maximizes at the ex-ante menu stage (equation (13)). In this sense, the ex-post recommendation problem and the ex-ante menu problem are consistent in their commission component.[3]

Weight on the employer's objective. The parameter $\theta_{b(f)}$ is the weight the broker places on the employer's objective. We index it by the broker type $b(f)$, consistent with the paper's assumption that steering intensity is homogeneous within broker type; it varies across employers only through the type of the broker they meet. A strictly positive $\theta_{b(f)}$ captures the broker's relationship and reputational motive: broker-of-record agreements in this market are written on a month-to-month basis (Section 2.1), so an employer can dismiss a broker that serves it poorly, and the broker therefore internalizes the employer's objective to a degree governed by $\theta_{b(f)}$. The polar cases are instructive: $\theta_{b(f)} \to \infty$ describes a broker that behaves as a perfect agent, while $\theta_{b(f)} \to 0$ describes a broker that maximizes commissions alone.

The broker recommends the plan that maximizes its ex-post payoff,

$$r_f \;\in\; \operatorname{argmax}_{j \in m}\; \Pi^{B}_{fjm}. \tag{A4}$$

The microfounded decision rule

Proposition 1 *Suppose Assumption 1 holds and $\theta_{b(f)} > 0$. The plan chosen by an employer that engages a broker is the maximizer over $j \in m$ of*

$$V_{fjm} = -\,\alpha_{fj}(1-\iota_f)\,p_{fjm}\,\textit{takeup}_{fjm} \;+\; \rho_f\,W_{fjm} \;+\; \delta^f_j + \xi_{fj} + \epsilon_{fj}, \tag{A5}$$

with

$$\alpha_{fj} = \alpha_0 + \mathbb{1}_{fb}\alpha_{1,b(f)} + \mathbb{1}_{fb}\alpha_{2,b(f)}\gamma_j, \qquad \alpha_{2,b(f)} = -\,\frac{1}{\theta_{b(f)}}\cdot\frac{1}{1-\iota_f}. \tag{A6}$$

By equation (A6), the steering coefficient is invariant to the insurer j and the menu m; it varies across broker types through $\theta_{b(f)}$, and across employers through the premium-contribution share ι_f.

By Assumption 1, the employer's chosen plan equals $r_f = \operatorname{argmax}_{j\in m}\Pi^{B}_{fjm}$. Because $\theta_{b(f)} > 0$ and argmax is invariant to multiplication by a positive scalar,

$$\operatorname{argmax}_{j\in m}\Pi^{B}_{fjm} = \operatorname{argmax}_{j\in m}\;\tfrac{1}{\theta_{b(f)}}\Pi^{B}_{fjm} = \operatorname{argmax}_{j\in m}\left\{\,u^{E}_{fjm} + \tfrac{1}{\theta_{b(f)}}\Pi^{\text{com}}_{fjm}\right\}.$$

Define $V_{fjm} \equiv u^{E}_{fjm} + \frac{1}{\theta_{b(f)}}\Pi^{\text{com}}_{fjm}$. Substituting the employer's objective (A1) and the commission income (A3), and recalling that commissions are paid only under a brokered contract and hence the commission

[3] The brokers' ex-post utility has an extra $\theta_{b(f)}u^{E}_{fjm}$ term that we do not include in the ex-ante menu problem. One could microfound $\theta_{b(f)}$ in a repeated relationship model in which poor recommendations reduce the broker's continuation value, for example by increasing the probability that the employer dismisses the broker. We do not model this dynamic problem explicitly. Instead, $\theta_{b(f)}$ summarizes the reduced-form discipline that such relational concerns impose on ex-post recommendations. The ex-ante menu problem retains the tractable commission-income-minus-cost structure used for estimation and abstracts from a direct utility weight on the employer's objective, except through the effect of menu quality on the employer's broker-use decision in Section 5.3.

term carrying the indicator $\mathbb{1}_{fb}$ yields

$$V_{fjm} = -\left(\alpha_{0,f} + \mathbb{1}_{fb}\alpha_{1,b(f)}\right)(1-\iota_f)\,p_{fjm}\,\text{takeup}_{fjm} + \rho_f\,W_{fjm} + \mathbb{1}_{fb}\,\tfrac{1}{\theta_{b(f)}}\,\gamma_j\,b_{fjm} + \delta_j^f + \xi_{fj} + \epsilon_{fj}.$$

Express the steering term in units of the employer's own expenses. Because the commission base $b_{fjm} = p_{fjm}\text{takeup}_{fjm}$ is the employer's premium-contribution variable $(1-\iota_f)p_{fjm}\text{takeup}_{fjm}$ rescaled by $(1-\iota_f)^{-1}$,

$$\tfrac{1}{\theta_{b(f)}}\,\gamma_j b_{fjm} \;\equiv\; -\,\alpha_{2,b(f)}\,\gamma_j\,(1-\iota_f)\,p_{fjm}\,\text{takeup}_{fjm}, \qquad \alpha_{2,b(f)} = -\tfrac{1}{\theta_{b(f)}}\,(1-\iota_f)^{-1},$$

and collecting the coefficients on $(1-\iota_f)p_{fjm}\text{takeup}_{fjm}$ delivers equations (A5)-(A6), the exact counterpart of equations (7)-(8) of the main text. Displaying both of its dimensions, the exact coefficient $\alpha^{exact}_{2,b(f)f} = -\theta^{-1}_{b(f)}(1-\iota_f)^{-1}$ varies with the broker type $b(f)$ through $\theta_{b(f)}$ and with the employer f through ι_f. The baseline empirical specification retains the first and restricts the second: it estimates a separate $\alpha_{2,b(f)}$ for each broker type while replacing the employer-specific factor $(1-\iota_f)^{-1}$ with a common one. The restriction is immaterial in our implementation, because the premium-contribution share is the sole channel through which the exact coefficient varies across employers, and Form 5500 reports only aggregate premium payments rather than the employer-employee split, so we calibrate ι_f to a common value (Appendix D1). Under that calibration, $\alpha_{2,b(f)}$ is exact rather than approximate; were employer-specific contribution shares observed, the ι_f-varying coefficient would be recovered by rescaling.

Two remarks follow. First, the employer's preference terms $(\delta_j^f, \xi_{fj}, \epsilon_{fj}, W_{fjm})$ survive in V_{fjm} only because $\theta_{b(f)} > 0$: if $\theta_{b(f)} = 0$ the broker would recommend the plan yielding the highest commission income mechanically, and V_{fjm} would collapse to pure steering. A positive weight on the employer's objective is precisely what makes the decision rule a genuine perturbation of that objective rather than its replacement. Second, with ϵ_{fj} drawn i.i.d. type-I extreme value, the maximizer of V_{fjm} delivers the conditional-logit choice probabilities estimated in Section 6.1.

Normalizations and maintained assumptions. Two features of the construction above deserve comment. The first clarifies the information structure needed for welfare interpretation. The second explains the empirical approximation that maps the exact commission-income wedge to the constant-within-broker-type premium-weight specification in equations (7) and (8).

(i) Information structure. Whether ϵ_{fj} is interpreted as the employer's idiosyncratic match value observed by the broker who knows its client, or as an idiosyncratic shock to the broker's recommendation, the broker's argmax yields the same logit kernel. This equivalence is an equivalence for choice probabilities, not necessarily for welfare. In the baseline welfare accounting, we interpret the relevant unobserved match component as part of the employer's experienced utility. Under an interpretation in which ϵ_{fj} is pure recommendation noise, the choice probabilities are unchanged, but the welfare correction would exclude that noise from experienced utility.

(ii) Functional form of the steering wedge. The exact broker problem delivers the wedge $V_{fjm} - u^E_{fjm} = \mathbb{1}_{fb}\theta^{-1}_{b(f)}\gamma_j p_{fjm}\text{takeup}_{fjm}$. This wedge is (a) zero whenever $\mathbb{1}_{fb} = 0$, because commissions arise only under

a brokered contract; (b) zero whenever $\gamma_j = 0$, because a broker paid no commission has no reason to depart from the employer's objective; and (c) linear in γ_j, because the commission rate enters the broker's payoff linearly. Equation (8)'s term $\mathbb{1}_{fb}\alpha_{2,b(f)}\gamma_j$, entering the weight the decision rule places on the employer's premium contribution, is precisely the object with these three properties, and is the first-order representation of any smooth commission technology. It is also appropriate over the narrow range of commission rates observed in our data (2-4%); we do not extrapolate the steering response beyond this range. Expressing the wedge as a component of the premium weight keeps both channels of broker influence, i.e., the choice-quality term $\alpha_{1,b(f)}$ and the steering term $\alpha_{2,b(f)}\gamma_j$, inside the single composite weight α_{fj}, so that $\alpha_{0,f}$, $\alpha_{1,b(f)}$, and $\alpha_{2,b(f)}$ are reported in common, directly interpretable units: the value the decision rule attaches to a dollar of the employer's own premium.

Two representations of the steering wedge. By equation (A6), $\alpha_{2,b(f)} = -\theta_{b(f)}^{-1}(1-\iota_f)^{-1}$, where $1-\iota_f$ is the employer's share of the premium. The change of variable that carries the exact wedge $\theta_{b(f)}^{-1}\gamma_j b_{fjm}$, with dollar commission base $b_{fjm} = p_{fjm}\text{takeup}_{fjm}$, into the premium weight therefore involves a single scale factor, $(1-\iota_f)^{-1}$, which carries no insurer index j, no menu index m, and no equilibrium object. Equation (8), which treats $\alpha_{2,b(f)}$ as a parameter (constant within broker type), is exact whenever employers share a common contribution share, as under the calibration of ι_f described in Appendix D1, and is a deliberate simplification only to the extent that ι_f varies across employer types; we argue that it is an innocuous one. The contrast with the alternative representation is instructive. Had the wedge instead been expressed in units of employee surplus, interacting γ_j with W_{fjm} rather than with the employer's premium contribution, the implied coefficient would be $\theta_{b(f)}^{-1} b_{fjm}/W_{fjm}$, and the scale variable would itself be an equilibrium object. Equations (A7)-(A8) of Appendix D1 imply, for a representative employee of employer f (the relation holds employee by employee), the identity $W_{fjm} = (\alpha_f^w)^{-1}\ln[1+\exp(\bar{u}_{fjm})] = -(\alpha_f^w)^{-1}\ln(1-\text{takeup}_{fjm})$, where α_f^w is the representative *employee*'s premium sensitivity, distinct from the employer's weight $\alpha_{0,f}$, $\bar{u}_{fjm}$ is the mean employee valuation of plan j, and the second equality uses $1-\text{takeup}_{fjm} = [1+\exp(\bar{u}_{fjm})]^{-1}$. Hence $b_{fjm}/W_{fjm} = \chi_{fjm} \equiv \alpha_f^w\, p_{fjm}\, g(\text{takeup}_{fjm})$ with $g(s) \equiv s/[-\ln(1-s)]$.

Crucially, the factor χ_{fjm} genuinely varies: α_f^w differs across employer types; g depends on take-up rates (though it is smooth, bounded in $(0,1)$, and mild over the observed range, e.g. $g(0.70) \approx 0.58$, $g(0.85) \approx 0.45$); and p_{fjm} is an equilibrium price set by insurers in Bertrand-Nash competition (Section 5.2), not a primitive. Placing the steering wedge on the premium term removes that source of variation entirely, since the commission base and the employer's premium contribution differ only by the factor $(1-\iota_f)$. Three considerations make the constant-$\alpha_{2,b(f)}$ specification innocuous.

First, the steering wedge the microfoundation actually delivers is $\theta_{b(f)}^{-1}\gamma_j p_{fjm}\text{takeup}_{fjm}$: it is the term $\mathbb{1}_{fb}\,\theta_{b(f)}^{-1}\gamma_j b_{fjm}$ in the choice rule displayed in the proof of Proposition 1, obtained simply by dividing the broker's payoff (A2) by the scalar $\theta_{b(f)}$. Because $\theta_{b(f)}$ is a structural primitive carrying no j or m index, and depending on the employer only through the type of its broker, the coefficient $\theta_{b(f)}^{-1}$ on the variable $\gamma_j p_{fjm}\text{takeup}_{fjm}$ is constant within broker type, and this representation of the wedge involves no approximation, only the variable it multiplies depending on (f,j,m). Equation (8) re-expresses the same wedge through the premium weight, interacting γ_j with $(1-\iota_f)p_{fjm}\text{takeup}_{fjm}$ in place of $p_{fjm}\text{takeup}_{fjm}$ so that $\alpha_{0,f}$, $\alpha_{1,b(f)}$, and $\alpha_{2,b(f)}$ share common units; the only employer-level dependence

of $\alpha_{2,b(f)} = -\theta_{b(f)}^{-1}(1-\iota_f)^{-1}$ beyond the broker type arises from that change of variable, and it involves no equilibrium object. The exact, ι_f-varying specification is thus nested and available; the constant-$\alpha_{2,b(f)}$ form is a parsimony choice.

Second, the simplification leaves intact every qualitative property the microfoundation establishes: the steering wedge is gated by $\mathbb{1}_{fb}$, vanishes at $\gamma_j = 0$, is increasing in γ_j, and is absent from experienced utility.

Third, it coincides with a modeling stance the paper adopts that steering intensity is taken to be homogeneous within broker type, "a tractable approximation of average steering effects." Under that reading, $\alpha_{2,b(f)}$ is the average steering intensity, which is precisely the object the demand estimates and the aggregate-welfare counterfactuals require.

Partial compliance. Assumption 1 is a full-delegation benchmark. The interpretation of the steering coefficient does not require a literal one-for-one compliance with broker advice. We relax it by letting the employer act on a weighted combination of its own objective and the broker's recommendation with influence weight $q_f \in [0,1]$, nesting full compliance ($q_f = 1$) and full independence ($q_f = 0$). The realized choice then maximizes

$$P_{fjm}^{pc} = q_f \frac{\exp(V_{fjm})}{\sum_{k\in m}\exp(V_{fkm})} + (1-q_f)\frac{\exp(u_{fjm}^E)}{\sum_{k\in m}\exp(u_{fkm}^E)},$$

where V_{fjm} is the broker-influenced decision rule and u_{fjm}^E is the employer's objective. For the local variation in commissions used to identify steering, this mixture has the same first-order implication as a single-index decision rule in which the commission wedge is attenuated by q_f:

$$V_{fjm}^{pc} = u_{fjm}^E + \mathbb{1}_{fb} q_f \theta_{b(f)}^{-1}\gamma_j b_{fjm}.$$

Equivalently, when the wedge is expressed in units of the employer's own expenses,

$$\alpha_{2,b(f)f}^{pc} = -\,q_f \theta_{b(f)}^{-1}\frac{1}{1-\iota_f}.$$

Thus the estimated constant-within-broker-type coefficient $\alpha_{2,b(f)}$ should be interpreted as an average of broker influence (q_f) times broker misalignment $(\theta_{b(f)}^{-1})$, scaled by the reciprocal of the employer's premium share. Full compliance corresponds to $q_f = 1$; when employers ignore broker advice, $q_f = 0$, and the steering wedge vanishes.

Interpretation: steering ability as misaligned broker preference Beyond rationalizing the reduced form, the microfoundation gives the steering parameter structural content. By equation (A6), $\alpha_{2,b(f)} = -\theta_{b(f)}^{-1}(1-\iota_f)^{-1} < 0$: a higher commission lowers the weight the decision rule attaches to the employer's own premium, so brokered choices tilt toward high-commission, high-premium plans. Steering ability, measured by the magnitude $|\alpha_{2,b(f)}|$, is therefore inversely proportional to $\theta_{b(f)}$, the weight the broker places on the employer's objective. Under the specification of equation (8), in which the steering wedge is measured in units of the premium weight, this is the reduced-form coefficient $\alpha_{2,b(f)} = -\theta_{b(f)}^{-1}(1-\iota_f)^{-1}$ reported in

the main text. Steering ability and regard for the employer are thus two sides of the same primitive. A broker that behaves as a perfect agent ($\theta_{b(f)} \to \infty$) has $\alpha_{2,b(f)} \to 0$, and the decision rule V_{fjm} coincides with the employer's own objective u^E_{fjm}; a broker that maximizes commissions alone ($\theta_{b(f)} \to 0$) has $\alpha_{2,b(f)} \to -\infty$, and recommendations track commissions. Because the scale factor $(1-\iota_f)^{-1}$ carries no broker index, estimated heterogeneity in $\alpha_{2,b(f)}$ across broker types, i.e., individual agents versus brokerage firms, reflects heterogeneity in $\theta_{b(f)}$ alone.

The microfoundation also locates the steering wedge within the paper's welfare accounting. Consistent with Section 5.3, the employer's (objective) experienced utility is evaluated at the pre-broker weight $\alpha_{0,f}$, so the entire broker-induced shift $\mathbb{1}_{fb}(\alpha_{1,b(f)} + \alpha_{2,b(f)}\gamma_j)$ in the premium weight enters the decision rule V_{fjm} but not experienced welfare. The microfoundation decomposes this decision-experienced wedge into its two sources: $\alpha_{1,b(f)}$, the broker's information refining the employer's objective u^E_{fjm}, and the steering distortion $V_{fjm} - u^E_{fjm} = -\mathbb{1}_{fb}\alpha_{2,b(f)}\gamma_j(1-\iota_f)p_{fjm}\text{takeup}_{fjm}$ that the broker's conflict of interest adds on top. Only the latter is commission-driven. This is exactly the wedge between decision utility and experienced utility that Section 5.3 carries, following Allcott (2013), here derived explicitly from the broker's recommendation problem.

D. Additional Model and Estimation Details

D1. Employee Surplus From Having Access to Employer-Sponsored Insurance

Below, we derive employees' decisions on whether to participate in the insurance plan that their employer sponsors and the surplus they derive from it.

Employees' Decision. The employee i decides whether to participate in the plan offered by insurer j through their employer f. For notational simplicity, we omit the subscript for geographic markets. Employee i's flow utility from choosing product j is:

$$u_{ij} = -(\alpha_0 + \underbrace{\alpha_1 \cdot z_i}_{\text{employee demo.}}) \cdot \underbrace{\iota_f \cdot p_j}_{\text{employee paid premium}} + (\beta_0 + \beta_1 \cdot z_i) \cdot \underbrace{x_j}_{\text{financial char.}} + \delta_j^w + \varepsilon_{ij}; \quad u_{i0} = \varepsilon_{i0}. \tag{A7}$$

where p_j is the insurance premium of plan j; ι_f is the employee's share of premium contribution at employer f; x_j is the plan's financial attributes, including deductibles, coinsurance rates; δ_j^w is insurer-specific brand effects perceived by workers. ε_{ij} is independent and identically distributed from the extreme-value type-I distribution. We allow each employee's preferences for insurance attributes to vary by their characteristics z_i, including age, gender, and household size. The outside option $j = 0$ is opting out of employer-sponsored plans, where the employee remains uninsured or enrolls in spousal coverage or other insurance programs.

Let takeup$_{fjm}$ denote the share of employees who take up plan j sponsored by employer f under broker menu m. The average welfare of employee i of employer f from having access to the plan j is

$$\mathcal{W}_{fjm} = \frac{1}{n_f} \sum_{i \in f} \ln(\exp(u_{ij}) + \exp(u_{i0})) \frac{1}{\alpha_i}, \tag{A8}$$

where n_f is the employer size. We set $\mathcal{W}_{f,j=U,m} = 0$ if the employer does not offer insurance coverage.

Estimation. Employees' primitives include their premium sensitivity α_i, preferences for financial attributes β_i, and brand perception δ_j^w.

Since Form 5500 lacks information on the exact attributes of insurance plans, we supplement it with enrollment records from a large, nationally representative insurer and conduct the estimation in two steps. In the first step, we exploit enrollment records of fully insured plans of the national insurer to estimate the premium sensitivity α_i and preferences for financial attributes β_i by worker demographics. We use maximum simulated likelihood estimators that derive choice patterns from equation (A7), substituting brand effects with an inside-option fixed effects.[45] We assume preference estimates α_i, β_i from this single large insurer accurately depict valuations for plan attributes of the NY employee population and are consistent across insurers (except insurer brand effects).

[4] Since both Form 5500 and the large insurer only include aggregate premium payments, we calibrate the premium contribution of employees to be 25% from the Kaiser/HRET annual employer survey. https://www.kff.org/series/employer-health-benefits-survey/

[5] We observe enrollment records, plan attributes, and demographics for employees who opt into employer-sponsored coverage in USIC data, and simulate the demographics for employees who are eligible for employer coverage but do not take it up using the empirical distributions of demographics from the Medical Expenditure Panel Survey (MEPS) for employees in a similar enrollment situation, based on age, gender, and household sizes.

In the second step, we utilize Form 5500 to estimate brand effects for all insurers, δ_j^w. To do so, we use three inputs from outside the Form 5500 data: insurer-specific distributions of product attributes from RWD Clarivate, market-by-firm-type-specific distributions of employee demographics from MEPS, and employees' preferences for each product attribute estimated from the large national insurer. We then use maximum likelihood to recover δ_j^w that rationalize within-firm take-up rates for employers offering that particular insurer.[6] Appendix B1 describes how we calculate take-up rates in Form 5500.

Identification. In the first step, variations in choice patterns and their correlation with product characteristics identify employees' preferences for prices and financial attributes, α_i and β_i. Since employees at the same firm face the same set of insurance plan choices, this within-employer method controls for unobserved employer-level factors that might affect selection into plans. In particular, preferences for product attribute parameters are primarily identified through enrollment decisions made by employees in firms that offer multiple plan options with different attributes from the same insurer. The overall valuation of the inside option (as opposed to taking up spousal coverage or remaining uninsured) is primarily identified by variation in take-up rates across firms that choose the large national insurer, conditional on attributes and demographics.

Insurers' knowledge of employees' unobserved preferences across plans when making pricing decisions creates a correlation between residuals and prices. We construct a Nevo (2001)-style instrument to address this endogeneity concern. We instrument for prices facing employees using the average prices of plans with the same characteristics (e.g., high versus low deductible) offered by other employers in the same firm size category (e.g., 100-499 or over 500 employees) in different markets. The instrument's relevance is satisfied because plans of similar characteristics offered by employers with similar bargaining leverage with insurers have similar cost structures. We assume that the unobserved preferences are idiosyncratic across employers after controlling for plan types and attributes. Namely, average prices of other employers do not reflect unobserved differences in preferences that affect choice probabilities for one particular employer.[7] We implement the instrument using a control-function approach following Wooldridge (2007).

In the second step, the brand effects, δ_j^w, are identified by variations in take-up rates across employers that offer different insurers in Form 5500. Holding fixed plan attributes and employees' preferences for these attributes constant, variations in the take-up rates reflect the perceived quality of each insurer relative to the same outside option of being uninsured or having non-employer-sponsored insurance. In other words, variations in take-up rates of the employer-sponsored options identify insurer-specific brand effects.

Results. Table A15 reports estimates of employee preferences from the conditional logit model. The first two columns present results with no instruments, and the second two columns present instrumented results. Across all specifications, employees prefer plans with lower premiums or a higher degree of financial generosity, such as low deductibles or coinsurance rates. However, the parameters on other attributes tend to diverge from there. Having a full set of interactions of demographics and attributes (columns 2 and 4) tends to yield sensible results on premium preferences by age. Namely, consumers become less price sensitive as

[6] As Form 5500 does not break the number of enrolled employees by insurer, we conduct second-step estimation only for employers that offer a single insurer, which accounts for 87% of employer-year pairs in our sample.

[7] This exclusion restriction may be violated if unobserved demand shocks correlate across firms in different markets, for example, when all employers with 100-499 employees have similar unobserved preferences for certain plan attributes not controlled for in our model. Given that preferences for plan attributes are quite local, this seems likely in this setting.

the average age of the household increases. Moreover, older consumers receive less utility from a higher out-of-pocket maximum compared to their younger peers, who may be less sick and more liquidity-constrained. Larger households are significantly less sensitive to price, as expected.

Table A15. Estimates of Employees' Preferences

	No Instruments		**With Instruments**	
	(1)	(2)	(3)	(4)
Monthly Premium	-13.296***	-14.386***	-67.093***	-65.231***
	(4.376)	(4.694)	(2.392)	(2.234)
Deductible	-0.230	-0.280	-0.797***	-1.018***
	(0.141)	(0.190)	(0.033)	(0.043)
Out-of-Pocket Maximum	0.339***	0.380***	0.230***	0.485***
	(0.103)	(0.134)	(0.014)	(0.034)
Coinsurance Rate	1.022	0.995	2.934***	2.727***
	(1.246)	(1.260)	(0.170)	(0.173)
Inside Option	-2.136	-2.090	4.748***	3.788***
	(1.401)	(1.421)	(0.377)	(0.396)
Premium Interactions:				
35-54 × Monthly Premium	-0.763	0.990	1.221***	3.156***
	(0.568)	(0.894)	(0.281)	(0.655)
55-64 × Monthly Premium	-2.091**	2.459*	-0.133	2.665***
	(0.999)	(1.441)	(0.474)	(0.959)
x Male	0.850	0.807	-0.407	-0.171
	(0.557)	(0.557)	(0.281)	(0.294)
x HH Size	11.772***	11.859***	13.112***	16.482***
	(1.767)	(2.104)	(0.351)	(0.587)
Deductible Interactions:				
35-54 × Deductible		0.036		0.053
		(0.085)		(0.044)
55-64 × Deductible		0.151		0.088
		(0.149)		(0.061)
x HH Size		0.018		0.199***
		(0.074)		(0.025)
Out-of-Pocket Maximum Interactions:				
35-54 × Out-of-Pocket Maximum		-0.097*		-0.104***
		(0.051)		(0.032)
55-64 × Out-of-Pocket Maximum		-0.259***		-0.152***
		(0.089)		(0.045)
x HH Size		0.010		-0.146***
		(0.047)		(0.020)
Observations	81,412	81,412	76,522	76,522

Notes: Bootstrapped standard errors are reported in parentheses. Monthly premiums, deductibles, and out-of-pocket maximums are measured in thousands of dollars. The estimation sample is fully-insured firms with more than 100 employees. The base age group is 18-34 years old. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The biggest differences, however, arise from the inclusion of instruments (columns 3 and 4). In our preferred specification (column 4), the magnitude of the premium coefficient quadruples in size, suggesting a significant downward bias in the absence of instruments. Moreover, other coefficients flip signs towards a more expected direction. Most notably, the value of the inside option (being insured through the firm) was negative and insignificant without instruments, suggesting consumers do not value insurance. This is clearly the result of the model being unable to rationalize why consumers forego insurance at low levels of price

sensitivity. However, with instruments, consumers can be price sensitive and value insurance, which aligns more intuitively.

We convert the point estimates to derived statistics. First, Table A16 derives own-premium elasticities by demographic groups. On average, a 1% increase in monthly enrollee-share premiums results in a decrease of about 3.5% in market share. This varies significantly by household size, with single households being the most price sensitive, while households with 3 members are significantly less price sensitive (with the latter having an elasticity of -1.04 on average). Within those groups, there is also significant variation: three member households, for instance, vary between elasticities of -0.69 and -1.47.

Table A16. Own-Price Elasticities by Consumer Type

	Mean	Std. Dev.	25th pctl	75th pctl	N
Total	3.59	2.45	1.70	5.26	15198
A. Household Size					
1	4.78	1.98	3.68	5.85	10008
2	2.40	1.20	1.60	3.23	2221
3	1.04	0.63	0.69	1.47	1347
B. Firm Size					
1-49 employees	3.08	2.37	0.93	4.57	1010
50-99 employees	3.65	2.45	1.71	5.21	8051
100-499 employees	3.58	2.44	1.70	5.37	6137

Notes: Elasticity estimates from the conditional logit model with premium interactions and control function.

While higher than existing elasticity estimates, there are a few notes that distinguish our data from prior literature. First, most of the employers in our data are small firms, the enrollees of which are thought to have higher price elasticities. Second, the single households make up a majority of the data, driving the estimates higher. Third, most firms in our data only offer one plan, meaning in a majority of cases, this can be thought of as an extensive margin elasticity, which tends to be higher as individuals could simply take up spousal coverage.[8] Finally, our estimates are broadly consistent with recent literature estimating demand for large firms, including Ho and Lee (2017), which finds elasticities ranging from -1.23 to -2.95 for *total* (enrollee + employer) shares.

We next report the estimated value of the plan, along with its characteristics other than price. A 10pp increase in the coinsurance rate is valued at approximately $112 annually for a household of two, while a $1,000 decrease in deductible is valued at approximately $233 by two-member households and approximately $800-$1000 for larger families. The average surplus from having access to insurers on NY ESI is between $103 - $142 per employee per month, though with considerable variation across firm type, region, and industry, as is reported in Table A17.

D2. Marginal Cost and the Enrolled Pool

This appendix derives the average marginal cost $\overline{c}_{fj}(p_{fjm})$ that enters equation (10) and explains how it is calibrated.

[8] While the average elasticities in works such as Cutler and Reber (1998) are between -0.3 and -0.6, the *extensive margin* elasticities are around -2. The remaining discrepancy could be explained by our relatively smaller firm size and higher monthly premiums.

Table A17. Willingness to Pay by Insurer

Insurer	Mean	Min	Max
Aetna	124.22	102.10	147.05
MVP Health	142.72	120.74	164.87
Empire (Anthem)	115.90	67.83	147.63
Oxford (United)	115.05	95.55	142.36
Healthnow	115.27	99.20	145.41
Connecticut (Cigna)	142.86	124.59	182.43
Emblem/HIP of NY	103.63	88.82	118.62
Excellus	128.41	107.01	164.75
CDPHP	113.92	96.86	139.48
Independent Health	114.73	102.05	129.22

Notes: Willingness to pay estimates are in dollars per employee per month.

Worker-type cost primitive. Partition the employees of employer type f into demographic worker types w (by age group and household size), with population shares θ_{fw} summing to one within each f. The primitive cost object is defined at the worker-type level: the per-member marginal cost of insurer j for a worker of type w is

$$mc_{jw} = c_j \cdot \lambda_w, \tag{A9}$$

where c_j is an insurer-specific cost level reflecting insurer j's negotiated provider prices, and λ_w is a demographic risk multiplier, common across insurers, that scales costs with the expected medical spending of worker type w. Employer-type risk is therefore not a separate cost parameter: it operates entirely through the composition $\{\theta_{fw}\}_w$, so that an employer type drawing a higher-cost demographic mix faces higher costs from every insurer.

Average cost of the enrolled pool. The insurer does not bear mc_{jw} for the full workforce, but only for employees who take up coverage. Let $\text{takeup}_{fjw}(p)$ denote the take-up rate of worker type w at employer f for insurer j's plan at premium p, obtained from the employee choice model of Appendix D1. The insurer's average per-member marginal cost for employer type f is the cost of mc_{jw} averaged over the employees who enroll:

$$\overline{c}_{fj}(p_{fjm}) = \frac{\sum_w \theta_{fw}\, \text{takeup}_{fjw}(p_{fjm})\, mc_{jw}}{\sum_w \theta_{fw}\, \text{takeup}_{fjw}(p_{fjm})}. \tag{A10}$$

Equation (A10) is the source of the adverse selection channel. Take-up takeup_{fjw} declines in the premium, and it does so at different rates across worker types; raising p_{fjm} thus shifts the weights in equation (A10) toward the worker types whose take-up is least price-sensitive, changing the average cost of the pool. The cell-level cost in equation (10) is then $mc_{fj} = \overline{c}_{fj}(p_{fjm}) + \omega_{fj}$, with ω_{fj} an employer–insurer cost shock that does not operate through enrollee composition.

Calibration. We calibrate the risk multipliers λ_w from the Medical Expenditure Panel Survey (MEPS). For each worker type w we compute mean allowed medical spending and set λ_w equal to that quantity divided by the industry-wide mean, so that λ_w measures the relative riskiness of worker type w. Given the λ_w, the insurer cost levels c_j are recovered so that the take-up-weighted average in equation (A10) reproduces the cell-level marginal costs obtained by inverting the pricing first-order conditions. We do not estimate each insurer–worker-type cost freely: with only the pricing first-order conditions as moments, the fully flexible

cost surface is weakly identified and yields implausible estimates for some cells.[9]

D3. Distance is the Primary Factor in Broker-Employer Matches

In this section, we provide empirical evidence that geographic proximity is the most important determinant of employers' broker choices: employers overwhelmingly prefer to work with the closest broker in their community.

We estimate a series of logit models of employer–broker matches. For each employer in the Form 5500 data, we define the potential broker choice set as all active brokers having at least one branch in New York State in a year. In the baseline model (Model 1), the only covariate is the log of the distance between the employer's location and the broker's nearest office. Model 2 replaces distance with the number of insurers in the broker's contracting network. In Model 3, we include both the log distance and the number of insurers in the contracting network, as well as an indicator for whether the broker is an individual agent. Finally, Model 4, labeled as the full model, adds interaction terms between all three broker-level variables and the employer's industry code, along with additional employer-level controls: the number of brokers hired, the number of carriers ultimately selected, the number of plan participants, and an indicator for whether the plan is fully insured. We then compare model performance using the Pseudo R-square across specifications, as reported in Column (1) of Table A18.

Table A18. Employer-Broker Matching

	Evaluation Scores	
	(1)	(2)
Models	Pseudo R2	AUC Distance
Logit with Dist (Model 1)	0.1833	0.8340
Logit with #Carriers in Network (Model 2)	0.0056	0.5836
Logit with Dist+#Carriers in Network+Individual Agent (Model 3)	0.1921	0.8582
Logit with Full Model (Model 4)	0.1970	0.8595

Notes: This table reports the predictive performance of alternative models in explaining employer–broker matching outcomes.

As shown, the Pseudo R-square is substantially larger in Model 1 (0.1833) than in Model 2 (0.0056), indicating that geographic distance between employers and brokers is far more predictive of matching outcomes than the number of carriers in a broker's network. Adding further covariates provides only modest improvements in explanatory power: Model 3 increases the Pseudo R-square by 0.0088 relative to the baseline, while the full specification in Model 4 raises it by just 0.0137.

Instead of relying on Pseudo R-square, we evaluate model performance using an alternative measure—the Area Under the Receiver Operating Characteristic Curve (AUC) distance. Specifically, within each choice set, we compare the predictive probability assigned by the logit models to the realized (chosen) option versus each alternative. If an alternative receives a higher predicted probability than the chosen

[9]For example, an unconstrained specification implies higher costs for younger than for older enrollees for some insurers. The constant-ratio restriction $mc_{jw}/mc_{jw'} = \lambda_w/\lambda_{w'}$ implied by equation (A9) disciplines the cost surface while still allowing costs to vary across both insurers and worker demographics.

option, it is assigned a score of 1; if lower, a score of 0; and if equal, a score of 0.5. The AUC distance is then computed as the average score across all comparisons in the whole choice panel. By construction, a random choice model, where each option has equal probability $1/N$, yields an AUC distance of 0.5. A higher AUC distance therefore indicates stronger predictive performance.

Column 2 of Table A18 reports the AUC values for our four models. Model 1 achieves an AUC distance of 0.8340, substantially above 0.5, confirming that the baseline model significantly outperforms a random choice model. Consistent with the Pseudo R-square results, Model 1 also exhibits much higher predictive power than Model 2, underscoring that geographic distance is the dominant predictor of employer-broker matches. Adding additional covariates in Models 3 and 4 modestly improves predictive performance, but the gains are small relative to the explanatory power of distance alone.

Taken together, these results highlight distance as the single most important determinant of employer-broker matching. This finding supports our simplifying assumption that, conditional on using a broker, employers are most likely to select the broker located closest to them.

D4. Multiplicity of Equilibria

Sources. The Nash-Bertrand pricing stage in Section 5.2 can admit multiple equilibria. The source is the feedback from premiums to both demand and costs. First, a higher premium lowers employer utility directly through the premium term in equation (7). It also lowers employee surplus and employee take-up, which feed back into employer demand. This makes the resulting demand system not log-concave in price. Second, because take-up changes the composition of the enrolled pool, a change in the premium also changes the average marginal cost through $\overline{c}_{fj}(p_{fjm})$. These forces make insurer best responses nonlinear, so the pricing first-order conditions in equation (9) need not have a unique solution. This appendix describes how we compute the pricing equilibrium set and select the equilibrium used for estimation and counterfactuals.

Computation. For each broker employer type (f)-broker menu (m)-geographic market cell, we solve the pricing fixed point from multiple starting values and collect the distinct equilibria reached. We use eleven starting points: six multiplicative perturbations of the data-initialized price vector, with factors 0.5, 0.75, 1.0, 1.25, 1.5, and 2.0, and five random perturbations, where each price component is scaled by an independent uniform draw on $[0.5, 2.0]$. From each starting point, we iterate all insurers' best responses to convergence. We then cluster converged price vectors and treat distinct clusters as candidate pricing equilibria.

For each candidate equilibrium, we compute insurer profits and first-order-condition residuals. We classify a candidate as valid if all insurers earn weakly nonnegative profits and the maximum absolute first-order-condition residual is below 10^{-3}. The selected equilibrium is chosen from this valid set.

Taxonomy. The pricing equilibria fall into three broad families. Table A19 reports an example from a three-insurer market.

The first family is the competitive equilibrium. Premiums are relatively low and tightly clustered across insurers. In this equilibrium, all insurers actively compete for enrollment, and enrollment remains economically meaningful. This equilibrium also has the broadest basin of attraction in our multistart algorithm.

The second family is the asymmetric equilibrium. One or more insurers set high premiums and obtain very small enrollment, while the remaining insurers price lower. This equilibrium can be sustained because

Table A19. Example: Three Equilibrium Types in One Market

	Insurer X	Insurer Y	Insurer Z
Competitive	$498	$547	$472
Asymmetric	$1,336	$546	$446
High-price	$1,626	$2,049	$1,670

Notes: Equilibrium premiums (PMPM) at the three fixed points reached in a single broker menu-geographic market-employer type cell. The table reports the three fixed points reached in this single three-insurer market.

the high-premium insurer has little enrollment at risk: the demand it would gain from a price cut is small relative to the margin it would lose on inframarginal enrollees. The lower-premium insurers, in turn, face fewer effective competitors.

The third family is the high-price equilibrium. All insurers set high premiums, and many employers move to the outside option of not offering coverage. Although premiums and markups are high, total insurer profits can be low because enrollment collapses. These equilibria typically have narrow basins of attraction.

Selection rule. We select one pricing equilibrium per cell. Among valid equilibria, the baseline selection rule chooses the equilibrium with the lowest average premium. In the simulations, this rule selects the competitive equilibrium in nearly all cells. We use the same rule for baseline estimation and all counterfactuals. The purpose is to keep comparisons on a common pricing-equilibrium branch and avoid attributing policy effects to arbitrary switches between competitive and high-price market-unraveling equilibria. This is also conservative for our anti-competitive channel: selecting the most competitive baseline equilibrium minimizes the price-softening we attribute to broker contracting, so any anti-competitive effect is not an artifact of selection.