Going for Broker?
Intermediation in Health Insurance Markets*

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

August 26, 2026

Preliminary and Incomplete – 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 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.

We begin with a stylized model to illustrate how steering and competitive-access distortions arise and

[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).

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 results in a 6% reduction in the share of brokers contracting with that insurer, and the increase in commissions has an opposite effect. For the pricing margin, we use this variation as an instrument for broker network breadth: relative to a control group of brokers that did not contract with the insurer and unexposed to the commission change, a reduction in the contracting network raises per-member annual premiums by about \$731 across competing insurers, consistent with narrower broker-mediated choice sets soften insurer competitionsoftening.

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. Hassle costs are identidied by comparing firms' propensities to work with brokers across similar geographic markets. 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 \$2.55 relative to their own premium expenses. In other words, they pay \$2.55 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 60% higher premiums than non-broker firms for the same level of employee surplus. Third, brokers provide real value: they reduce employers' hassle costs of offering insurance by \$30 PMPM, or 7.8% of the average employer premium share. For a typical 200-employee firm in New York City, this amounts to approximately \$71,000 per year, approximately the average salary of hiring an benefits consultant employee. Finally, brokers' selective contracting imposes sizable premium costs. Compared with full-menu contracts, exclusive constracts allows insurers' markup to raise by about 40% for the same client.

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 effects 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

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

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). 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.[3] We contribute two insights. First, we provide direct evidence that commission incentives distort brokers' recommendations in health insurance. Second and more novel, we show that intermediaries affect welfare not only through steering but also through their contracting decisions, which endogenously determine the set of upstream suppliers competing for consumer demand.[4]

[3] 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.

[4] 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

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).[5] 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 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 specialized benefits staff[6]. As a result, brokers are often hired and serve as a critical touchpoint in helping employers navigate plan selection, pricing, and administration.

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.

[5] 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.

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

Broker-Employer Relationships. Employer-broker matches are typically highly localized and relationship-based. Although some large brokerage chains operate nationally, broker markets are essentially local: employers generally prefer working 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.[7] 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, with employers rarely switching (see section 2.2) except in discrete shocks such as changes in ownership or retirement of the broker. 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 both administrative and advisory services. For administrative services, they handle enrollment, regulatory compliance, renewal paperwork, and ongoing benefits management, reducing the hassle costs employers would otherwise incur to offer coverage.[8] For advisory services, they assess firm risk and workforce preferences, collect quotes, present plan options, and make product recommendations to the human resources department or the chief financial officers. The administrative role gives brokers value independent of plan choice; the advisory role creates scope for commission-based steering.

In addition, brokers can also reduce insurance costs by leveraging the breadth of their insurer relationships as negotiating leverage. 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." This non-monotonicity that more carriers help, but only up to a point, previews the tradeoff we formalize later: broader broker networks can increase competitive pressure, but maintaining those networks is costly.

Broker-Insurer Relationships. Broker-insurer contracting further shapes the set of products available to employers. While there are many health insurance carriers in New York, a broker is not required to have contracts with or sell every one of them. 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.[9] Agreements with carriers are typically not exclusive. Brokers may collect quotes without such an agreement, but they cannot complete a sale without a formal contract.

Brokers often choose not to contract with certain carriers when expected client volume is low or when

[7] 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)

[8] 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.

[9] 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).

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 selective contracting decisions are prevalent in the industry and central to the competitive-access channel: they determine which insurers compete for a given employer's demand.

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.[10] Carriers generally pass these payments through to employers in premiums and have limited 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.

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) enrollment under 42 C.F.R. §422.2274 (Ritter Insurance Marketing, 2024). 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 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 are required to submit these forms, while self-insured plans could 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 our data construction process in detail.

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

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 NY over the past two decades. Extracting and hand-coding the key compensation variables across insurers and years required substantial manual effort; to our knowledge, this is the first dataset to compile insurer commission schedules at this level of coverage and detail. The schedules report each carrier's standard broker compensation and vary across insurer identity, 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. Combining these sources allows us to track both the set of insurers a broker can represent and the timing of these relationships.

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.

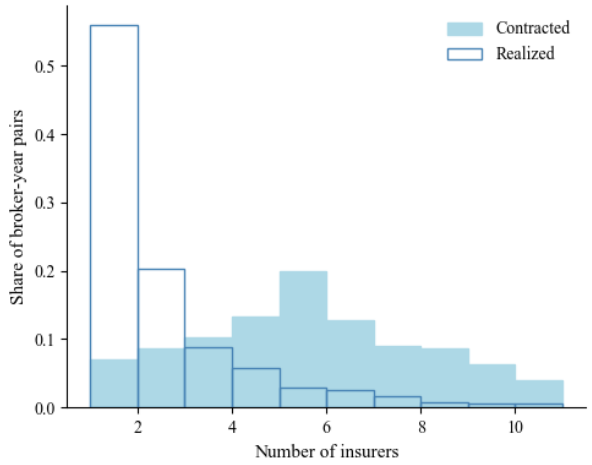
Broker-Insurer Contracts. Figure 1 illustrates the distribution of broker-insurer contracts and highlights three key patterns. First, selective contracting is widespread. Only 3% of brokers contract with all ten major insurers, and the median broker contracts with five insurers. Second, the set of insurers 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 2 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.

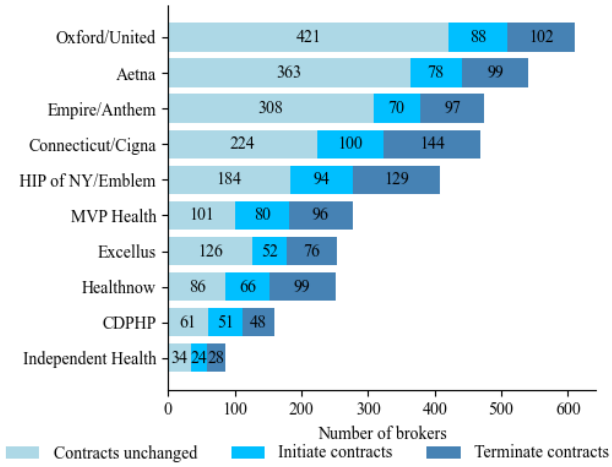
Limited Employer-Broker Switching. Table 3 reports statistics on employer-broker matching dynamics. Two patterns emerge. First, broker switching is infrequent. Employers work with an average of 1.68 brokers over the sample period, and only 13% of employer-year observations involve a switch from one broker 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 1. Broker-Insurer Contracts Among Fully-Insured Employers

(a). Insurers contracted versus recommended



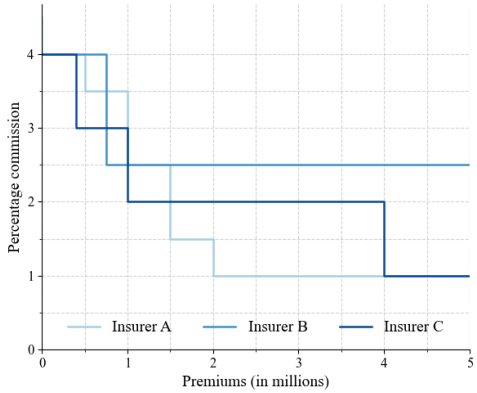
(b). Changes in contracting networks over time



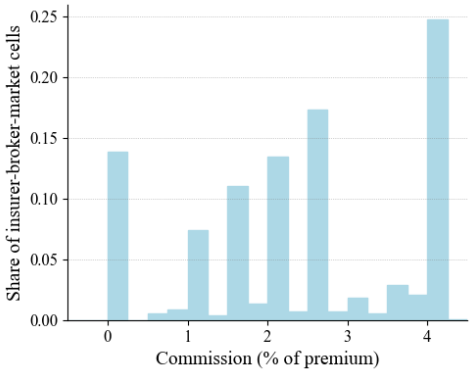
Notes: The figure depicts broker-insurer relationships among brokers that serve fully insured New York State ESI employers in 2011-2020. The sample is restricted to the ten major insurers reported in Table 2 in the empirical analysis. Panel (a) plots the broker-year-level distribution of the number of 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 2. 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.[11] The game proceeds in three stages. First, the broker observes commissions[12] 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

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

[12] 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:

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

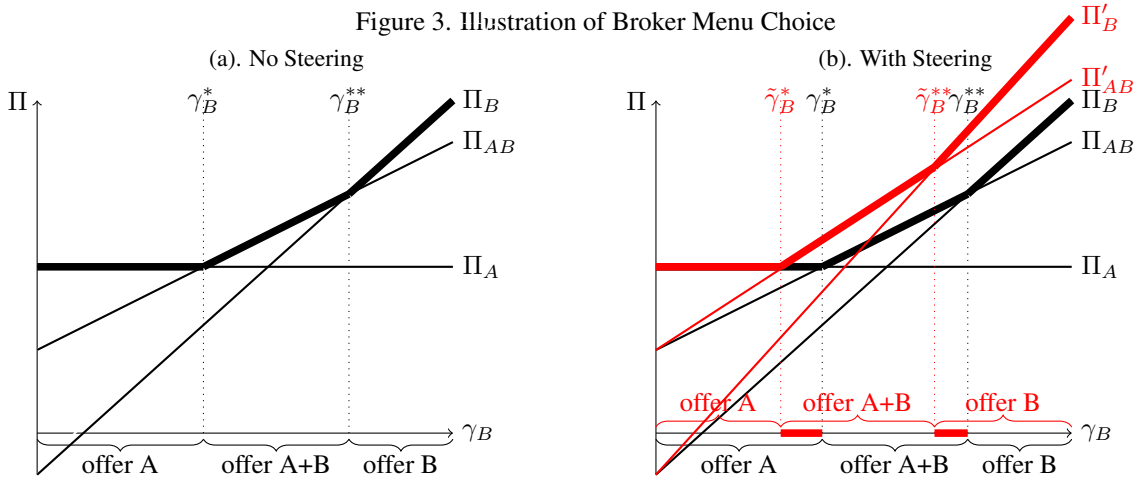
The first condition states that employers inherently prefer A to B. The second condition states that B is chosen less often than it would be if offered exclusively against the outside option. The third condition states that A is sufficiently attractive that under combined contracting, it remains profitable 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 3a. 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 3. 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 3b. 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

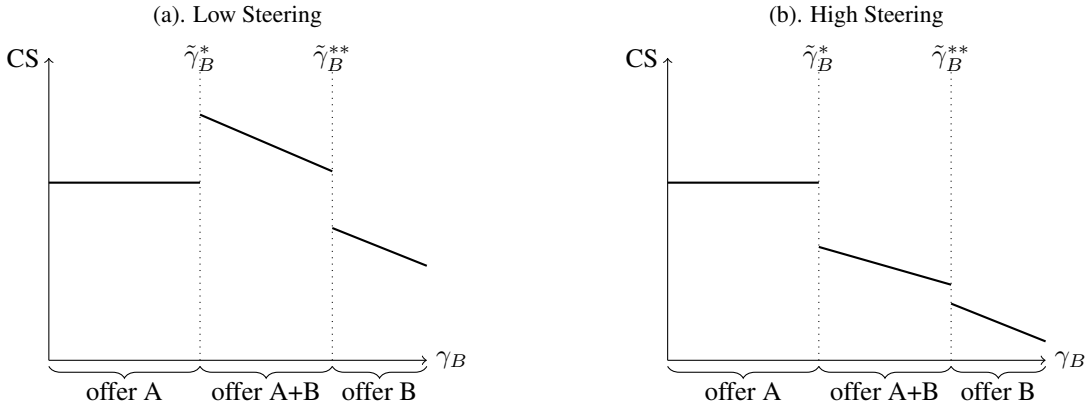
We now consider employer surplus, taking into account the broker's menu choice. We measure surplus in utility units and suppress the common $1/\alpha$ scaling. Expanding the menu from A to A+B creates two benefits: more product variety and stronger price competition, while steering toward the high-commission 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:

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

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 4a and 4b 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 4. 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*

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 5, 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).

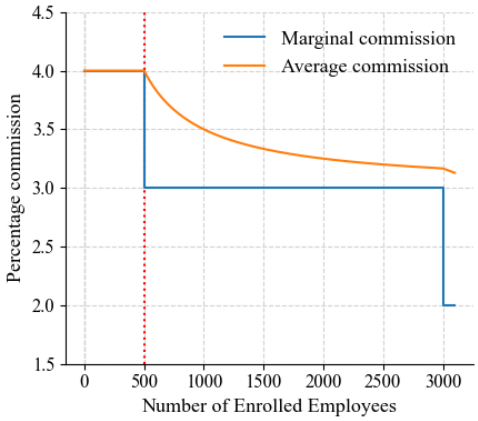
We exploit this kink to test whether employers on either side of the threshold differ in their propensity to choose Insurer 1. We focus on employers whose broker has Insurer 1 in their network and estimate:

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

where Y_f is an indicator for whether employer f chooses Insurer 1 and $K = 500$ is the commission kink cutoff[13]; $g()$ 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

[13]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 5. 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):

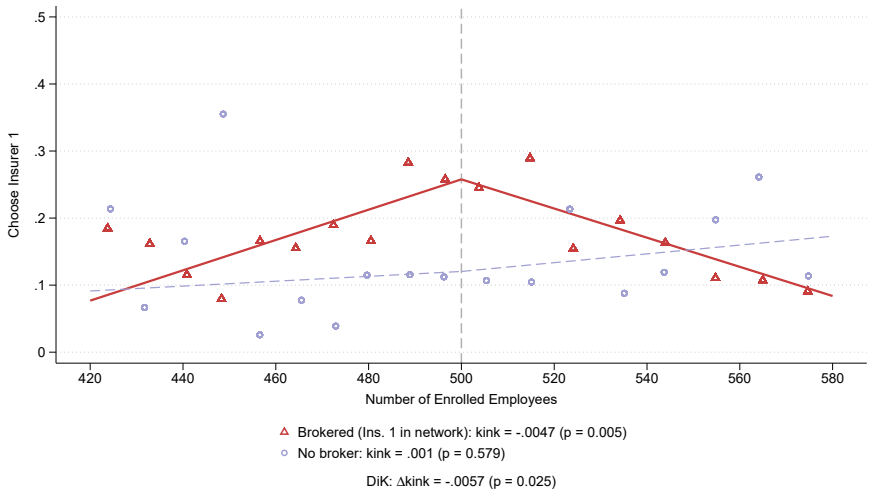
$$\begin{aligned} Y_f =& \beta_0 + \beta_1 \mathbb{1}_{fb} + g_0(X_f - K) + \mathbb{1}_{fb} \cdot g_1(X_f - K) + \tau_0 \cdot \mathbb{1}[X_f \geq K] \cdot (X_f - K) \\ &+ \tau^{DiK} \cdot \mathbb{1}_{fb} \cdot \mathbb{1}[X_f \geq K] \cdot (X_f - K) + \varepsilon_f, \end{aligned} \tag{4}$$

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. The remaining threat is therefore specific: 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, the control group is in fact flat, exhibiting no kink in the probability of choosing Insurer 1 and none in premiums; because these employers also purchase Insurer 1, the absence of a kink argues against a supply-side change by Insurer 1 at the cutoff.

Figure 6. 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 5). 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.

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.[14] We restrict to fully insured, single-carrier employers, who are best represented in Form 5500, and include covariate adjustments for the NYC and Mid-Hudson regions to absorb compositional differences between downstate and upstate NY.

Figure 6 reports the main estimates. The differential slope change is sharp and statistically significant: each additional ten enrollees beyond the threshold reduces the probability of choosing Insurer 1 by about 5.7 percentage points relative to the pre-kink trend. The components are consistent with the design: brokered employers kink downward by 4.7 percentage points per ten enrollees, 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

[14] 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).

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, log premiums per enrollee fall by roughly 4% per ten additional enrollees 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 6. 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.

The validity of the DiK design rests on two requirements, and we verify each. 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. Were this occurring, enrollment would bunch at the threshold. It does not: 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).[15]

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

[15]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.

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- and MSE-optimal bandwidths, quadratic polynomials, a uniform kernel, alternative covariate adjustments, and donut-hole exclusions (Table A1). The estimate retains its sign, magnitude, and significance. 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

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.

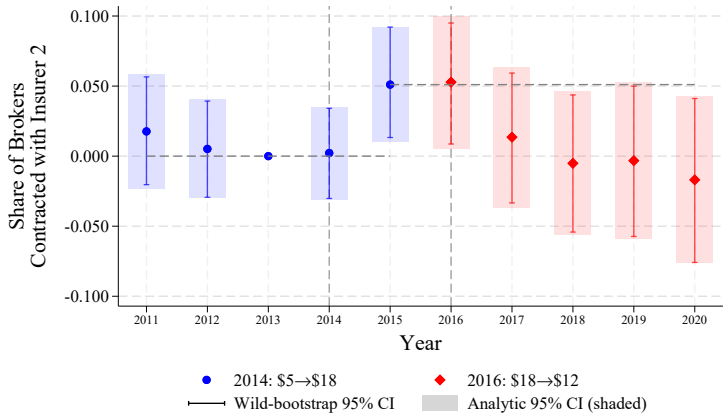
Panel A of Figure 7 traces Insurer 2's broker network share relative to a 2013 baseline across the full 2011–2020 panel, using a single-baseline event study that absorbs county-by-year fixed effects. The pre-period estimates for 2011 and 2012 are small and statistically indistinguishable from zero, confirming the absence of differential pre-trends. Following the 2014 commission increase from \$5 to \$18 per member per month, Insurer 2's broker share rises sharply and peaks in 2015, indicating that brokers responded to the higher compensation by expanding their contracting with Insurer 2. After the 2016 commission cut from \$18 to \$12, the share reverses and declines monotonically through 2020, unwinding much of the earlier gain. This non-monotone trajectory, a rise under the 2014 increase followed by a decline under the 2016 cut, motivates separate event studies for each shock. We implement these in Panel B using a stacked design. For the 2016 event, we restrict the estimation window to 2015–2020, starting the sample one year before the shock rather than in 2011. Because the 2014 commission increase had already shifted broker networks by 2015, including 2014 in the pre-period of the 2016 event study would confound the baseline with the lingering treatment effects of the earlier shock; 2015 instead represents the last year in which Insurer 2's elevated broker share was in a new steady state prior to the cut.

Panel B of Figure 7 displays the results for the two events separately, with each event's window truncated so that the two do not overlap. 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 a 33%

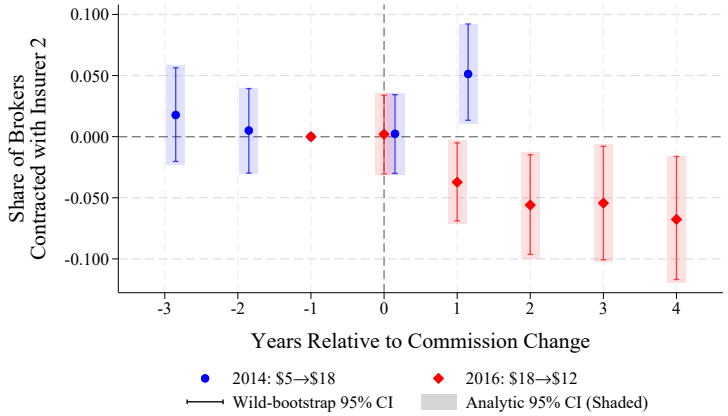
reduction in commissions in 2016 lowered Insurer 2's broker contracting share by about 5 to 7 percentage points by two or four years out, or roughly 6 to 8 percent relative to the baseline mean. Figure A7 further decomposes the contracting response into entry and exit. The 2014 increase raises entry 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.

Figure 7. Effect of commission changes on broker networks

(a). Event Study (2013 Benchmark)



(b). Separate Event Studies



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. Panel (a) reports estimates from a single event-study specification using 2013 as the benchmark year. Panel B reports 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. In panel (b), estimates are normalized 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.

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 A8), showing no evidence of contemporaneous

changes in pricing, underwriting, or risk selection. Second, plan generosity (deductibles and coinsurance) and provider access (the number and share of in-network NY hospitals) (Figure A9) stay flat relative to other carriers, except for coinsurance at the 2016 cut.[16] Third, employee take-up: the share of eligible employees who enroll in Insurer 2's plans when offered is unchanged at both events (Figure A10). 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 A11 compares Insurer 2's pooled difference-in-differences estimates with a null distribution constructed from 1,000 placebo draws. In each draw, one placebo insurer is selected uniformly at random from the nine non-Insurer-2 carriers, along with two ordered placebo treatment years chosen from 2012–2019 (excluding the actual treatment years 2014 and 2016), subject to a minimum two-year gap between the placebo events. The upper panel shows that Insurer 2's true 2014 estimate lies in the right tail of the placebo distribution, with a one-sided randomization p-value of 0.260. Although this estimate is not statistically distinguishable from the placebo distribution, its positive sign is consistent with the event-study evidence of a positive broker-share response following the 2014 commission increase. The lower pannel shows that Insurer 2's true 2016 estimate falls in the left tail, with a one-sided p-value of 0.027. This result indicates that placebo shocks do not generate systematic changes in broker market shares, providing additional supprt for interpreting the observed response as a causal effect of the commission change.

4.3. Premium Responses to Commission-Induced Network Changes

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.

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

[16]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 7), 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. A plan-generosity channel would have to move contracting at both.

other insurers, excluding Insurer 2:

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

where P_{jfbct} is the premium per enrollee 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_{ct} and α_{jb} are employer, county-year and insurer-broker fixed effects that control for time-invariant employer risks, local market shocks, and the negotiation skills of the insurer-broker pairs.

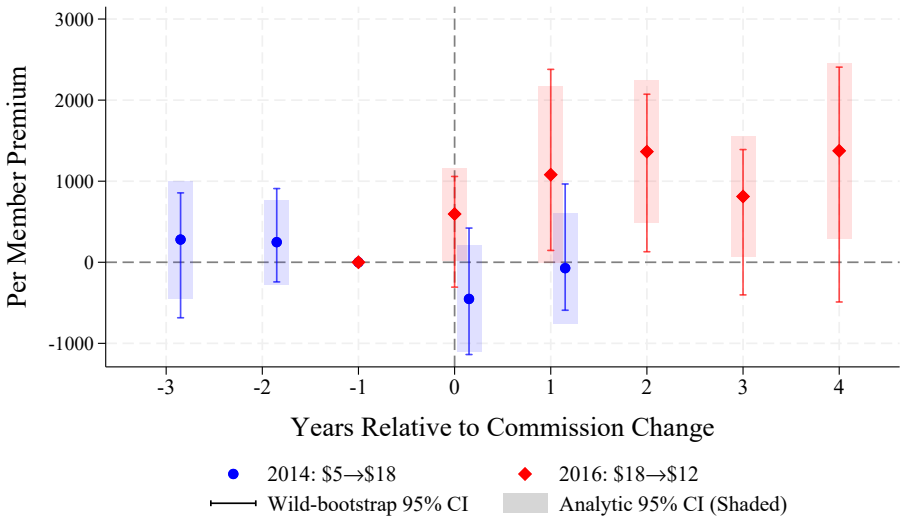
The identifying assumption is that Insurer 2's commission changes affect rivals' premiums only by changing whether brokers contract with Insurer 2, and not through any contemporaneous change in Insurer 2's own conduct that would move rivals' prices directly. Under this assumption, comparing newly exposed brokers to already-contracted brokers isolates how a commission-induced change in network breadth feeds through to competition, net of any common shock to all brokers in a county-year.

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, the variation is clean. 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 A12). So the premium response does not appear to reflect a broader re-optimization of broker portfolios. We cannot rule out small, 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.

Estimates in Figure 8 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 estimate implies a rival-premium increase of approximately $932 per enrollee per year for brokers exposed to a commission-induced network contraction, relative to comparison brokers. This finding indicates that broker networks set the scope of rivalry by expanding or contracting the set of insurers employers can reach.

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 A13) 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 A14). 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

Figure 8. 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, insurer-by-broker, 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.

low-overlap rivals alike. Table A2 presents analogous triple-difference estimates, and confirms that the response is larger among rivals whose pre-period broker-network overlapped most with Insurer 2's network.

Finally, Figure A15 reports randomization inference for equation (6). In each of 1,000 replications, two ordered placebo event years are drawn from 2012–2019, excluding the actual treatment years 2014 and 2016. For each placebo event, the exposed broker-county cohort is independently re-randomized while preserving the share of treated observations among event-eligible broker-county cell, and the pooled difference-in-differences specification is re-estimated. The true estimates lie in the predicted tails of their respective placebo distributions, with one-sided randomization p-values of 0.050 (2014 shock) and 0.002 (2016 shock). These results indicate that the estimated rival-premium responses are not artifacts of timing or cohort composition.

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 commis-

sion 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.[17] 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 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:

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

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 employee, weighted by its price sensitivity α_{fj}. The employer pays share $1 - \iota_f$ of the per-member premium p_{fjm}; and takeup$_{fjm}$ is the employee take-up rate.

[17] 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.

The second term, $\mathcal{W}_{fjm}(p_{fjm}, \text{takeup}_{fjm})$, measures the average ex-ante surplus of 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 PMPM commissions are converted to premium-equivalent rates. We assume brokers affect employers' choices by adjusting how much employers value their employees' welfare:

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

$\alpha_{1,b(f)}$ captures non-commission differences in the way brokered employers' insurer choices load on employee surplus 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.[18]

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

[18] 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.

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

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

where $\overline{c}_{fj}(p_{fjm})$ 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 $\overline{c}_{fj}$ on price captures adverse selection: a higher premium disproportionately deters lower-cost employees, raising the average cost of those who remain. Appendix D2 derives $\overline{c}_{fj}(\cdot)$ 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

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

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, which causes differences 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, if the employer forgoes broker services, it must pay the hassle costs of completing administrative paperwork or searching for a suitable insurer.

Let V_f^{wb} and V_f^{nb} denote utilities with and without a broker, and let u_{fjm}^{wb} and u_{fjm}^{nb} denote the flow utility in equation (7) evaluated at $\mathbb{1}_{fb} = 1$ and $\mathbb{1}_{fb} = 0$, respectively. Both the choice-quality term $\rho_{1,b(f)}$ and the steering term $\rho_{2,b(f)}\gamma_{fj}$ enter the employer's *choice rule* but not its *experienced utility*: at the point of deciding whether to hire a broker, the employer does not yet know how broker intermediation will shift its preferences, so it evaluates both options at its pre-broker valuation ρ_0. The wedge between decision utility and experienced utility is therefore $-(\rho_{1,b(f)} + \rho_{2,b(f)}\gamma_{fj})$ for an employer using a broker, and zero otherwise. Following Allcott (2013), we explicitly account for this wedge as a surplus correction:

$$V_f^{wb} = \frac{1}{\alpha_f}\mathbb{E}\Big[\underbrace{\max_{j\in m}}_{\text{broker menu } m} u_{fjm}^{wb}\Big] + \frac{1}{\alpha_f}\sum_j \Big(s_{fjm} \cdot \underbrace{(-\rho_{1,b(f)} - \rho_{2,b(f)} \cdot \gamma_{fj})}_{\text{wedge: broker effects on preferences}} \cdot \text{WTP}_{fjm}\Big), \tag{12}$$

$$V_f^{nb} = \frac{1}{\alpha_f}\mathbb{E}\Big[\underbrace{\max_{j \in J \cup U}}_{\text{all options on market}} u_{fjm}^{nb}\Big], \tag{13}$$

where WTP_{fjm} is the willingness-to-pay derived from $\mathcal{W}_{fjm}$.[19]

Employer f uses a broker if its expected utility gain exceeds a random hassle cost H_f:

$$\mathbb{1}_{fb} = 1 \quad \text{if} \quad V_f^{wb} > V_f^{nb} - \underbrace{H_f}_{\text{hassle cost}}, \quad H_f \sim N(\mu_f, \sigma^2). \tag{14}$$

We further parameterize μ_f as a function of employers' types, i.e., by industry, firm sizes, geographic markets, with an idiosyncratic normally-distributed error term of standard deviation σ.

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, that switching rates are similar in thick and thin broker markets, and that 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 quote 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. To assess the robustness of our findings, we consider alternative specifications that strengthen the outside option. For instance, 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.

[19] The subscript m in u_{fjm}, p_{fjm}, s_{fjm}, takeup_{fjm}, $\mathcal{W}_{fjm}$, and π^I_{fjm} denotes the broker menu under which equilibrium prices are quoted. In the no-broker branch, m degenerates to the full market $J \cup U$, since employers without brokers face prices set under the no-broker market structure.

The expected commission income from menu m is

$$\Pi_b(m) = \mathbb{E}_{\xi,\omega,H}\Big[\sum_{f \in \mathcal{F}_b} \underbrace{\mathbb{1}_{fb}}_{\substack{\text{whether employer} \\ \text{uses broker}}} \cdot \sum_{j \in m} \underbrace{s_{fjm} \cdot \text{takeup}_{fjm}}_{\substack{\text{demand from} \\ \text{employer } f\text{, plan } j}} \cdot \underbrace{p_{fjm} \cdot \gamma_{fj}}_{\substack{\text{commission} \\ \text{income}}} \Big], \tag{15}$$

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.

$$m = \text{argmax} \Big(\underbrace{\Pi_b(m)}_{\text{commission income}} - \underbrace{K(m)}_{\text{contracting costs}} \Big), \quad K(m) = \sum_{j \in m} K_j, \tag{16}$$

where total contracting costs are additive across per-insurer fixed cost parameters K_j.

Equation (15) and (16) 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 $\sum_j K_j$. Third, brokers take into account that their menu decision affects insurer pricing p_{fjm} and consumer choices $s_{fjm}, \text{takeup}_{fjm}$. Equation (16) 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 (16); each insurer's premium satisfies the first-order condition in equation (9); each employers' broker-use decision satisfies the threshold rule in equation (14); 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.[20] 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

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

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.[21] 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 and Identification

6.1. Employers' Choice of Insurer

Key primitives include employers' weights for their own expenses, α_f; the weights for employee surplus ρ_0 (non-brokered baseline), $\rho_{1,b(f)}$ (broker-status wedge), and $\rho_{2,b(f)}$ (broker steering response to commissions) in equation (8); and brand perception, δ_j^f.

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

We draw on two additional data inputs to facilitate the estimation. First, since the Form 5500 data only reports premiums for the realized insurance choice, we assume that employers in the same market, type, 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, where the hat denotes 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. The identification of the steering coefficient, $\rho_{2,b(f)}$, exploits discontinuities in commission schedules *within* insurers, similar to the intuition in Section 4. Conditional on the same insurer, the differential choice patterns of employers who work with brokers and are just below and above the commission change cutoff, along with their correlations with the magnitude of commission changes, identify how commission payments affect choice rules.

The baseline weight ρ_0 is identified by non-brokered employers' choices: conditional on premiums, insurer effects and other controls, it measures how strongly these employers choose insurers with higher employee surplus. We interpret ρ_0 as the revealed weight on non-brokered employers, not as the counterfactual preference of all employers absent brokers. The broker-status wedge $\rho_{1,b(f)}$ is identified by the differences in the choice patterns among brokered and non-brokered employers. Because employers may select into broker use, this parameter is descriptive. We interpret $\rho_{1,b(f)}$ as a reduced-form wedge that absorbs non-commission differences between brokered and non-brokered choice behavior, but not the causal

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

effect of brokers on choice quality.

The correlation of choice patterns and insurance premiums identifies α_f. Insurers' knowledge of employers' unobserved preferences when making pricing decisions creates a correlation between residuals and prices. To address this endogeneity, we use a control function based on Hausman instruments (Nevo, 2001). 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 variations in prices in other markets can signal an insurer's cost changes in all markets, which also shifts the equilibrium pricing choices in the focal market. It is plausible that residual demand shocks in other markets are mean independent of residual demand shocks 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.

Results. Table 4 reports estimates of employers' preferences.

Table 4. Estimates of Employers' Preferences

Realized Choice	Coefficient	Standard Error
WTP per employee	0.0495	0.0133
has broker × WTP	0.0037	0.0044
has broker × WTP × effective commission	0.1662	0.0266
Premium per employee	−0.0194	0.0049
Control function for premium	0.0156	0.0049

Looking first at the premium coefficients, the results indicate that employers incur disutility from their portion of the total premium paid, as expected. The coefficient indicates that, for firms not using a broker (and hence given the option not to offer insurance), a 10% increase in premiums results in a 4.8 percentage point (pp) decrease in the probability of offering insurance. This is well within established estimates of offer elasticities and, in particular, is quite similar to estimates from Gruber and Lettau (2004). On the intensive margin, if the price of a major insurer increases by 10%, holding all other prices constant, the probability of a no-broker firm offering that insurer decreases by approximately 16%.

The coefficient on the control-function residual is positive and statistically significant. This coefficient is not a preference parameter. Rather, it indicates 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 W_{fjm}, δ_j^f, and observed controls.

We also compare employers' experienced-utility weight on employee surplus to their weight on premium expenditure, α_f. Per the framework in Section 5, this weight is ρ_0 for employers without a broker and $\rho_0 + \rho_{1,b(f)}$ for employers with a broker (the steering term $\rho_{2,b(f)}\gamma_{fj}$ enters the choice rule but not experienced utility). Under a benchmark with full wage offsets and no tax wedge, equal marginal weights would be efficient: i.e., $\rho_0 = \alpha_f$ without a broker and $\rho_0 + \rho_{1,b(f)} = \alpha_f$ with a broker. The rationale is that 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.

Three key patterns are outlined. First, employers who do not work with brokers value \$1 of employees' welfare relative to their own premium expenses at \$2.55. In other words, we find that, for each dollar in value of employee surplus, employers pay, on average, \$2.55 in insurance premiums. Non-brokered employers

place a revealed weight on employee surplus that exceeds the premium coefficient, i.e., "overweighting" health insurance relative to average employee valuations, which could be rationalized by the tax treatment of ESI (Gruber, 2010) or other employer objectives.

Second, however, brokers tend to increase employers' valuation of employee WTP by \$0.20 to \$0.58, implying that brokers worsen the matches between employers and insurers, and encourage employers to over-insure their workforce. Our estimates imply that, at the most extreme, firms with brokers pay about 20% more in premiums to generate the same dollar of employee value as non-broker firms. This "overweighting" is consistent with the estimates from Tilipman (2022). We note that two possible reasons could explain this difference. Employers with brokers may differ in unobserved ways that render them more likely to offer generous insurance than employers without brokers. Alternatively, it is possible that even in the absence of commissions, brokers may recommend different types of insurers than employers would have chosen on their own.

This leads to our final point: commissions lead to a higher employer weight on employee surplus. Indeed, we estimate that about 20% of the difference in generosity of offers between no-broker and broker firms reported above can be attributed to broker commissions. For a prominent insurer in our data, decreasing commissions by 1pp reduces employers' likelihood of choosing the plan by about 3pp. This is of a similar magnitude to our reduced form estimates of brokers' steering effects in Section 4. It also implies that, given our intensive margin elasticities of employees reported in D1, employers with brokers appear to be more sensitive to the same percentage change in commissions than in raw premium rates.

6.2. Insurer Pricing

The key primitive in this stage is marginal costs, mc_{fj}, by insurer, employer types, and geographic markets.

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, region, industry, and firm-size fixed effects by ordinary least squares to obtain smoothed cost estimates.

Results. Table 5 reports our estimates. The magnitude of the monthly per-member marginal cost estimates is sensible. First, the marginal cost-to-price ratio is approximately 85%, consistent with MLR requirements. Second, big insurers have higher markups than small insurers. Third, New York City and urban areas incur higher claims expenses than rural areas, such as the Finger Lakes. Fourth, larger employers have lower claims costs than smaller employers.

6.3. Employers' Choice of Whether to Use a Broker

Primitives to be estimated in this stage are the parameters governing employers' hassle costs associated with finding and evaluating insurers without a broker's assistance.

Table 5. Per-Member Monthly Marginal Cost Estimates

	Marginal Costs	Premiums	Cost Ratio (%)
By Insurer			
Aetna	677	772	87.74
MVP Health	463	515	89.87
Empire/Anthem	674	766	87.90
Oxford/United	639	773	82.71
Healthnow	618	688	89.90
Connecticut/Cigna	650	730	89.01
HIP of NY/Emblem	597	701	85.35
Excellus	539	643	83.84
CDPHP	586	668	87.71
Independent Health	541	617	87.75
By Region			
Capital Region	551	626	88.09
Central NY	546	630	86.76
Finger Lakes	531	602	88.32
Mid Hudson	624	717	87.16
Long Island	643	744	86.45
Mohawk Valley	522	594	88.05
New York City	658	758	86.90
North County	488	564	86.80
Southern Tier	580	672	86.54
Western NY	549	637	86.22
By Industry			
Natural resources	607	666	91.09
Mining	470	532	88.30
Construction	665	759	87.76
Manufacturing	601	685	87.84
Wholesale and retail trade	610	695	87.77
Transportation and utilities	653	758	86.35
Information	632	732	86.38
Financial activities	652	754	86.61
Professional and business services	648	748	86.76
Education, health, and social services	589	679	86.94
Leisure and hospitality	592	679	87.17
Other services	649	751	86.54
By Employer Size			
100-499 employees	636	730	87.12
500+ employees	574	662	86.82

Estimation. We estimate a random-effects probit model of employers' decisions to use a broker, employing a two-stage control function approach to address the endogeneity of the broker's value-added.

Let V_f^b and V_f^{nb} denote the employer's expected plan-choice value under broker and no-broker regimes, respectively, computed using the estimated preferences from the previous stage under no-broker (experienced-utility) weights. The key object is the value difference $\Delta V_f = V_f^b - V_f^{nb}$, which measures the employer's gain from the broker's insurer menu relative to searching on its own. A positive coefficient on ΔV_f implies that employers are more likely to use brokers when brokers provide more valuable menus, and the ratio of the remaining index to this coefficient delivers the implied hassle cost in dollar (PMPM) terms.

In the first stage, we regress ΔV_f on the instrument z_{ft} (described below) and controls (firm size, region, industry, year fixed effects, broker characteristics, lagged broker use, and the firm-level mean $\overline{\Delta V}_f$), and

recover the residual $\hat{v}_{ft}$.

In the second stage, we estimate a random-effects probit:

$$\Pr(\text{use broker}_{ft} = 1) = \Phi\Big(\beta_V \Delta V_{ft} + \beta_{\hat{v}} \hat{v}_{ft} + \beta_{\bar{V}} \overline{\Delta V}_f + \mathbf{X}'_{ft}\boldsymbol{\gamma} + \beta_\ell \text{broker_lag}_{ft} + u_f\Big), \tag{17}$$

where $u_f \sim N(0, \sigma_u^2)$ is a firm-level random effect capturing persistent unobserved heterogeneity in broker preference; $\hat{v}_{ft}$ is the control function residual from the first stage, which corrects for endogeneity of ΔV_{ft}; $\overline{\Delta V}_f$ is the Mundlak device (the firm-level time average of ΔV_{ft}), which allows β_V to be identified from within-firm variation even when $\text{Corr}(\Delta V_{ft}, u_f) \neq 0$; broker_lag_{ft} indicates whether the firm used *any* broker in the prior year, capturing inertia in broker use; and $\mathbf{X}_{ft}$ includes firm size, region and industry indicators, year fixed effects, and broker characteristics (log number of clients, number of counties served, individual agent indicator).

The implied hassle cost for firm f is therefore:

$$\widehat{F}_f = \frac{\mathbf{X}'_{ft}\hat{\boldsymbol{\gamma}} + \hat{\beta}_\ell \cdot \text{broker_lag}_{ft} + \hat{\beta}_0}{\hat{\beta}_V}. \tag{18}$$

Identification. The central endogeneity concern is with ΔV_{ft}: employers who use brokers may do so in part because the broker they match with offers a valuable menu, but that menu is itself an equilibrium object shaped by the broker's contracting decisions, which may respond to employer characteristics. For example, brokers serving employers in industries with complex benefit needs may contract with different insurer menus than those serving relatively simpler preferences, generating a correlation between ΔV_{ft} and unobserved employer heterogeneity. More broadly, any unobserved firm characteristic (e.g., HR sophistication, workforce health risk) that affects both the employer's outside option V_f^{nb} and its likelihood of using a broker will bias β_V.

We address this endogeneity using instruments that exploiting variation in the broker's value-added across its clients in different geographic markets. Specifically, we construct a leave-one-out instrument in the spirit of Berry et al. (1995): for each employer–broker–year observation, the instrument is the average ΔV among the broker's other clients in counties outside the employer's own county. This captures broker-level shifts in value-added that are plausibly exogenous to any single employer's unobserved demand shocks, conditional on county and year fixed effects. As noted above, we implement this using a control function approach (Petrin and Train, 2010).

The Mundlak device $\overline{\Delta V}_f$ addresses a distinct concern: even with the control function, the random effect u_f may be correlated with the time-average level of ΔV_{ft} across years. Including $\overline{\Delta V}_f$ as a regressor projects out this between-firm correlation, so that β_V is identified from *within-firm* variation in ΔV over time—the same variation exploited by fixed-effects estimators, but within the random-effects probit framework that allows estimation of σ_u (Mundlak, 1978; Chamberlain, 1984).

The random effect u_f captures persistent unobserved heterogeneity in firms' propensity to use brokers—for instance, differences in HR capacity, organizational complexity, or management preferences that are not captured by our observables or by broker networks and premiums. Crucially, u_f and the hassle cost parameters in $\boldsymbol{\gamma}$ are separately identified because the hassle cost covariates (industry, region, firm size) vary

across firms, while u_f is identified from the *within-firm serial correlation* in broker use over time. σ_u is pinned down by how much a firm's broker-use decisions in one year predict its decisions in other years, beyond what ΔV_{ft} and the observables already explain. The magnitude of σ_u relative to the total error variance (measured by $\rho = \sigma_u^2/(1+\sigma_u^2)$) indicates how much of the residual variation in broker use is driven by persistent firm-level factors versus idiosyncratic year-to-year shocks.

Results. Table 6 reports the estimated hassle cost coefficients from equation (17), expressed in dollars PMPM.

Table 6. Hassle Cost Estimates

Variable	Hassle Cost ($ PMPM)	Std. err.
Constant	125.58	13.49
Industry (NAICS):		
Construction	+27.97	19.15
Manufacturing	−2.24	8.84
Wholesale and retail trade	+10.40	9.46
Transportation and utilities	+39.05	20.29
Information	−1.24	10.13
Financial activities	−3.68	8.48
Professional and business services	+1.29	8.41
Education, health, and social services	+6.24	8.24
Leisure and hospitality	−7.30	10.27
Other services	+0.00	0.00
Firm Size:		
Per employee	−0.06	0.01
Region:		
Central NY	+21.19	13.75
Finger Lakes	−1.47	8.54
Mid Hudson	−2.30	8.52
Long Island	+5.88	8.59
Mohawk Valley	−3.81	12.74
New York City	+1.41	7.27
North County	−17.84	16.36
Southern Tier	+7.83	12.16
Western NY	−17.29	8.24
Broker Characteristics:		
Log number of clients	−20.47	1.86
Counties served	+2.12	0.30
Individual agent	+17.59	4.87
Inertia:		
Used broker in prior year	+23.18	2.25
Random Effect:		
$\hat{\sigma}_u$	1.91	0.15
$\hat{\rho} = \sigma_u^2/(1+\sigma_u^2)$	0.785	0.027

Notes: Estimates represent hassle costs per member per month. The omitted industry category is mining. The omitted region is the Capital Region. Inertia reports the PMPM contribution of lagged broker use ($\hat{\beta}_\ell/\hat{\beta}_V$). $\hat{\sigma}_u$ and $\hat{\rho}$ are the estimated standard deviation and variance share of the firm-level random effect, reported in native units (not rescaled to PMPM).

The random effect is large and highly significant: $\hat{\sigma}_u = 1.91$, implying $\hat{\rho} = 0.785$. That is, approximately 78.5% of the residual variance in broker use is attributable to persistent firm-level heterogeneity rather than idiosyncratic shocks. This result is consistent with employers selecting to use brokers for rea-

sons not captured simply by their plan menus and premiums: variation in firm risk aversion, preferences, etc.[22]

The Mundlak term $\overline{\Delta V}_f$ is significant and negative, indicating that firms with higher average ΔV across years are *less* likely to use brokers, conditional on within-year ΔV—consistent with a selection pattern where firms that would benefit most from brokers have already sorted into broker use, leaving those with high ΔV but low broker adoption as firms with low unobserved hassle costs.

Lagged broker use is strongly predictive of current use, with an implied PMPM contribution of approximately \$23. This captures inertia in broker relationships: firms that used a broker in the prior year are substantially more likely to continue doing so, consistent with switching costs or habit persistence. Whether this inertia component represents a "real" cost of transitioning away from broker services or a behavioral preference is ambiguous; we therefore report hassle costs both inclusive and exclusive of this term. Our preferred specification treats this as relevant for hassle costs, which we feel is a more conservative interpretation.

The average implied hassle cost across all firms in the estimation sample is approximately \$79.60 PMPM, with an average of \$82.05 PMPM among currently brokered firms and \$51.42 PMPM among currently non-brokered firms. Excluding the inertia component, the average search cost is approximately \$57 PMPM.

The heterogeneity across industries, regions, and firm sizes follows expected patterns. Transportation and utilities firms face the highest search costs (\$39 PMPM above the construction baseline), likely due to the complexity of risk profiles and regulatory requirements in those sectors. Information and leisure/hospitality firms face the lowest, consistent with simpler benefit structures. Hassle costs decline with firm size, at roughly \$0.06 PMPM per additional employee, reflecting economies of scale in HR capacity. Conditional on firm and broker observables, the geographic variation is modest, with Western New York and North Country firms facing substantially lower search costs than the Capital Region baseline, likely reflecting less complex insurance markets in rural areas.

Finally, our estimates imply that firms whose closest broker, on average, serves more employer clients face substantially lower hassle costs. This suggests that brokers with larger client bases bring more experience and market knowledge to the procurement process, reducing the administrative burden on employers. Conversely, firms more likely to be matched with individual agents rather than brokerage firms face substantially higher hassle costs, consistent with agencies offering greater institutional capacity. Firms matched with brokers serving more counties face slightly higher hassle costs, though this result is small in magnitude and likely reflects collinearity.

The magnitude of these estimates is economically plausible. For a 200-employee firm, the average search cost of \$80 PMPM translates to approximately \$192,000 per year. A benefits specialist in New York City earns between \$60,000 and \$80,000 annually, and a benefits manager approximately \$150,000. The implied search cost is thus equivalent to hiring roughly two to three benefits specialists or one to two benefits managers dedicated to insurance procurement.

[22] Without these unobseved persistent heterogeneity estimates, our model would load all this variation into implied hassle costs, leaving us with an average hassle cost estimate of \$160 pmpm, twice our current estimate.

6.4. *Brokers' Choice of Insurer Menu*

The key primitives in this stage are the per-insurer contracting costs $\{K_j\}_{j \in J}$, with total cost $K(m) = \sum_{j \in m} K_j$.

Estimation. Conducting empirical analysis on broker network formation presents a series of computational and econometric challenges. First, the number of possible network configurations increases exponentially with the number of insurers and thus suffers from the curse of dimensionality. Second, insurers are interdependent. That is, the revenue from the inclusion of an extra insurer depends on the identity of insurers in the existing portfolio. Instead of directly solving the industry equilibrium, we exploit the inequality restrictions implied by brokers' revealed preferences.

We assume that the total contracting costs of the broker menu m are linearly additive across the contracting costs of each included insurer. 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. Let m^* denote the observed broker menu, $m^- = m^* \setminus j$,

$$\Pi_b(m^*) - \Pi_b(m^-) = \Delta\Pi_{b,jm}(m^*, m^-) + \nu^1_{bjm} + \nu^2_{bjm} \geq K_j. \tag{19}$$

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, $m^+ = m^* \cup j$. This generates a lower bound for the contracting cost.

$$\Pi_b(m^+) - \Pi_b(m^*) = \Delta\Pi_{b,jm}(m^+, m^*) + \nu^1_{bjm} + \nu^2_{bjm} \leq K_j. \tag{20}$$

ν^1_{bjm} is the expectation error of brokers' expected revenue from a contract (Dickstein and Morales, 2013; Prager and Tilipman, 2025). This expectational error term is unknown to brokers when they make their network decisions. Including this term allows the model to rationalize cases in which it would otherwise predict that a broker should contract with an insurer, but the broker does not form the relationship.

ν^2_{bjm} is an error term that is known to brokers at the time they decide on contracts, but is not known to the econometrician. ν^2_{bjm} allows contracting costs to vary across markets and brokers in ways not otherwise captured by this parameterization. However, the presence of this term introduces a selection problem. Brokers that receive high draws of ν^2_{bjm} will be more likely to contract with insurer j.

Identification. To identify contracting costs, we require additional assumptions.

First, we assume that the residual profit difference ν^1_{bjm} satisfies $\mathbb{E}[\nu^1_{bjm}|\mathcal{I}] = 0$, where $\mathcal{I}$ is the brokers' information set when they decide on contracts. In other words, conditional on the information set, when the sample of inequalities grows large, the errors tend to zero in the limit. We thus stack the inequalities across brokers and estimate K_j using sample moment inequalities (Pakes et al., 2015). Similar assumptions have been made in Dickstein and Morales (2013) and Prager and Tilipman (2025).

Second, we assume that the contracting cost with a given insurer is identical across brokers and menus, $\nu^2_{bjm} = \nu^2_j, \forall b, m$. In other words, we rule out the possibility that network formation depends on private information of heterogeneity in contracting costs across brokers. This implicitly assumes away any structural errors that might bias the estimate of K_j (Eizenberg, 2014; Pakes, 2010); the estimated K_j thus absorbs the

constant structural-error component ν_j^2. Similar assumptions have been made in Ho (2009) and Nosko (2010).

The moment inequalities literature has proposed several alternative approaches to address the structural error ν_{bjm}^2. Ciliberto and Tamer (2009) estimate a distribution for the structural error term. This approach is computationally challenging in our setting, as the network formation problem is combinatorial rather than a binary-entry problem. Fan and Yang (2025) develop inequalities that allow the estimation of the distribution of sunk costs in settings with product positioning; however, their inequalities require the identification of the best and worst case situations in terms of rivals' actions. Identifying these in our setting is non-trivial, since the worst-case scenario may not be a full insurer menu, but a selective contract by rivals. Wollmann (2018) restricts how the structural error term may vary across observations. This approach is unattractive in our setting because it limits the ability to identify parameters that are common across insurers or years.

Third, K_j reflects the annual fixed costs of contracting incurred irrespective of whether an insurer had a contract with a broker in prior years. That is, we assume that the broker training and licensing to sell process is costly, even for existing contracts. This is motivated by the fact that brokers undertake the costly task of learning about changes in insurers' policies every year, then deciding whether to carry that insurer. This also facilitates our computation by abstracting from the dynamic decisions of long-term contracts.

Under these assumptions, we incorporate the following inequality moments into the GMM estimation by penalizing violations of these conditions proportional to their magnitude.

$$\mathbb{E}_{b,m}[\Delta\Pi_{b,jm}(m^+, m^*) - K_j)] \leq 0, \quad \mathbb{E}_{b,m}[K_j - \Delta\Pi_{b,jm}(m^*, m^-))] \leq 0. \tag{21}$$

K_j are identified by differences in expected revenue between different menus.

Results. Table 7 reports the estimates.

Table 7. Contracting costs (rescaled to a monthly basis)

Insurer	Lower Bound	Upper Bound	Midpoint
Aetna	666	1496	1081
MVP Health	635	1213	924
Empire/Anthem	1403	1539	1471
Oxford/United	688	2024	1356
Healthnow	644	1038	841
Connecticut/Cigna	1367	2034	1700
HIP of NY/Emblem	417	585	501
Excellus	2219	2689	2454
CDPHP	1174	1282	1228
Independent Health	1437	1629	1533

6.5. Model Fit.

Before proceeding to the simulation results, Table A3 reports the model fit. It is reassuring that the simulated equilibrium characterizes the observed pricing strategies and market shares well.

7. Counterfactuals

7.1. Quantify Distortions

We conduct two partial equilibrium simulations to quantify the welfare effects of relevant distortions. We hold the employer-broker matches fixed as in the status quo, and focus on the employers who work with brokers. First, we remove commission-induced steering in employers' choices rules to measure the impact of broker steering. Second, we augment employers' choices sets to include the non-contracted insurers to measure the impacts of selection contracting. We report average per-enrollee prices of the resulting choice patterns, employee surplus ($\mathcal{W}_{fjm}$ in equation (7)), and employer surplus (defined in equation (12)). For tractability, the results reported are in year 2018, across all geographic market in NYS.

Before proceeding to the simulation results, Table A3 reports the model fit. It is reassuring that the simulated equilibrium characterizes the observed pricing strategies and market shares well.

Steering. Table 8 reports the simulated market outcomes under alternative broker incentive regimes. Column (1) shows the baseline case with steering. Columns (2) and (3) remove steering, holding prices fixed and allowing for price re-optimization, respectively. Three key patterns are outlined.

Table 8. Welfare Effects of Steering Distortions

	(1)	(2)	(3)
	Employers with brokers		
	With steering	No steering No repricing	No steering With repricing
Average menu sizes	4.99	4.99	4.99
Average employee takeup rates	64.99	67.00	68.57
Average employee surplus	105.66	112.13	114.64
Average employer surplus (no hassle costs)	237.49	243.30	260.52
Average premium per enrollee	8890	8721	8411
Average premium per employee	5731	5799	5732
Average markup	1.20	1.18	1.14
Average commission per enrollee	216	155	170
Average percentage commission	2.43	1.78	2.02

First, comparing Columns (1) and (2), we find that restoring the optimal choice rule, so that employees choose plans aligned with their preferences rather than broker steering, raises welfare substantially. Average employee surplus increases from 105.7 to 112.13, while employer surplus (net of hassle costs) rises from 237.5 to 243.30. These welfare gains reflect both improved plan matching and lower priced products being choosen. Average per-enrollee premiums decline slightly from \$8,890 to \$8,721, consistent with reduced form results.

Second, comparing Columns (2) and (3), we find that allowing insurers to reprice in the no steering scenario induces further decreases in average enrollee premiums from \$8,721 to \$8,411, a 5% reduction from the baseline. Absent steering, employers now place a relatively smaller weight on employee surplus, or in other words, for the same unit of employee welfare, the employers' willingness to pay drops. This disciplines insurers' pricing power and equilibrium price reduction. As a result, average employee surplus further increases to 114.64, while employer surplus (net of hassle costs) rises to 260.52.

Finally, the effects are heterogeneous across insurers, as shown in Table 9. Average per-enrollee premium drops by 7.8% for large insurers like United, while only 3% for regional insurers like MVP Health. Price declines are most pronounced among insurers with higher baseline commission rates and more agents contracted, that are more likely to benefit from broker steering. Eliminating steering make regional lower-priced insurers stronger competitors for national higher-priced insurers. This equilibrium impacts further enhance consumer welfare.

Table 9. Equilibrium Strategy under With or Without Steering Case

	(1)	(2)	(3)	(4)	(5)
	Market Share			Price	
	With steering	No steering No repricing	No steering	With Steering	No steering
Aetna	13.7	11.5	11.9	9577	9222
MVP Health	3.9	5.5	4.6	6438	6242
Empire/Anthem	10.3	8.0	8.3	9331	9083
Oxford/United	21.9	9.9	14.7	9884	9166
Healthnow	5.8	7.4	6.4	8322	8137
Connecticut/Cigna	12.8	29.7	23.2	8814	8644
HIP of NY/Emblem	2.5	0.7	0.7	8793	8334
Excellus	24.3	24.0	26.7	8155	7759
CDPHP	2.5	1.2	1.4	8459	8129
Independent Health	2.3	2.1	2.1	7744	7448

7.2. Quantifying Anti-competition.

Table 10 quantifies the impact to competitive-access effects by expanding the menu of available plans. Column (1) reports the menus in the baseline, while Column (2) reports the scenario where we manually augment all insurers that appear on the geographic market in that specific year to the choice set. Three patterns are outlined.

Table 10. Welfare Effects of competitive-access Distortions

	(1)	(2)
	Employers with brokers	
	Baseline	Full
Average menu sizes	4.99	6.65
Average employee takeup rates	63.67	64.98
Average employee surplus	102.90	105.50
Average employer surplus (no hassle costs)	235.70	259.43
Average premium per enrollee	8903	8789
Average premium per employee	5622	5667
Average markup	1.20	1.18
Average commission per enrollee	215	223
Average percentage commission	2.42	2.54

First, expanding the choice set leads to meaningful improvements in welfare and competitive discipline. Employee surplus rises modestly, from 102.9 to 105.5 with steering, reflecting the gains from better plan matching. Employer surplus increases even more sharply from 235.7 to 259.4 under steering, reflecting the gains from both increased product variety and lower prices.

Secoond, prices respond in the expected direction. Average enrollee premiums decline by roughly 1–2 percent following menu expansion. These declines suggest that broader choice intensifies price competition among insurers.

Finally, Table 11 presents insurer-level responses to menu expansion. When new or popular insurers are introduced, market shares shift considerably toward those entrants. For instance, Oxford/United's share rises from 22.5 to 32.6 percent, while Excellus and Healthnow—insurers less frequently chosen in the baseline menus—lose market share. These reallocations highlight the importance of consumer preferences and name recognition in shaping equilibrium enrollment patterns once employees face a broader set of options. The entry of well-known insurers intensifies competition, forcing incumbent plans to reprice aggressively to maintain enrollment.

Table 11. Equilibrium Strategy under Baseline Menu and Counterfactual Full Menu

	(1)	(2)	(3)	(4)
	Market Share		Price	
	Baseline	Full	Baseline	Full
Aetna	14.2	13.4	9573	9412
MVP Health	3.7	1.7	6445	6356
Empire/Anthem	10.0	9.5	9327	9341
Oxford/United	22.5	32.6	9882	9170
Healthnow	6.0	4.1	8329	8602
Connecticut/Cigna	12.4	13.0	8810	8456
HIP of NY/Emblem	2.4	3.0	8809	8631
Excellus	23.7	19.6	8155	8060
CDPHP	2.6	2.0	8463	8486
Independent Health	2.5	1.2	7749	7630

7.3. Commission Caps

Our second counterfactual analysis is motivated by existing policies that constrain intermediaries' compensation. For instance, 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. At the federal level, the Centers for Medicare and Medicaid Services (CMS) sets explicit maximum payments that carriers and Field Marketing Organizations (FMOs) may provide to agents and brokers for enrolling beneficiaries into Medicare Advantage (MA) plans, as specified under 42 C.F.R. §422.2274. These examples illustrate that regulators already intervene in broker compensation schemes in order to limit distortions, yet the welfare consequences of such restrictions remain unclear.

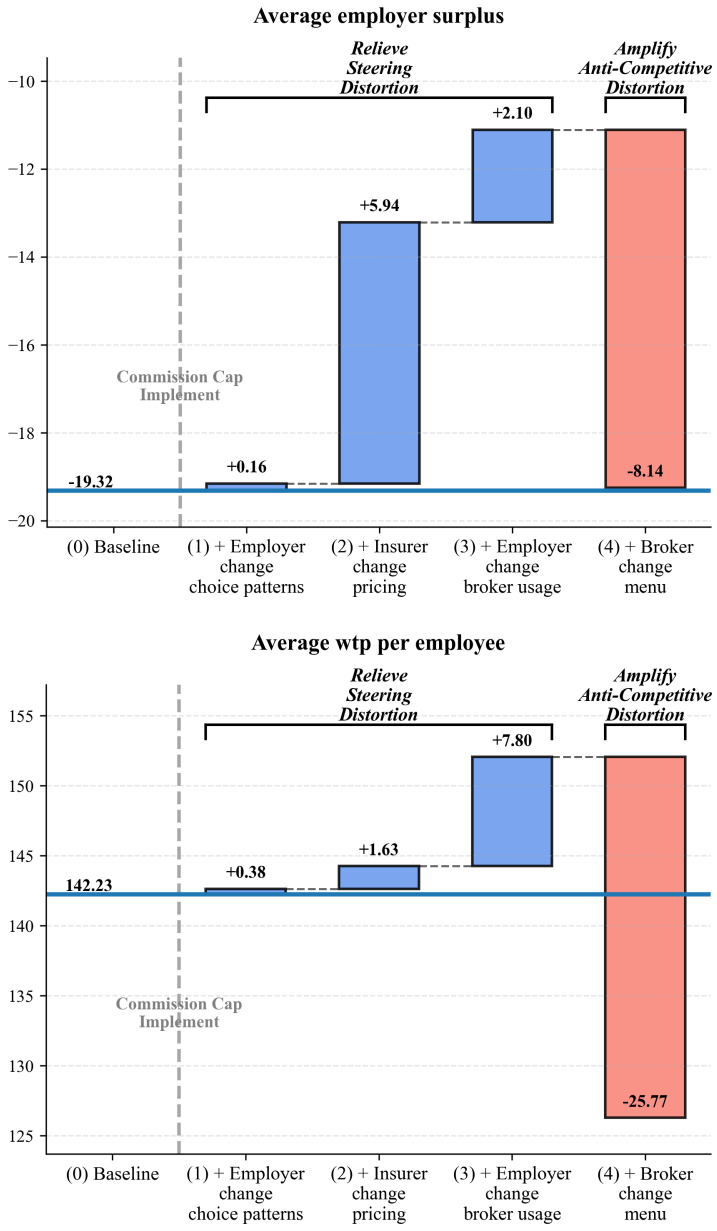
We use the model to evaluate a one-percentage-point cap on broker commissions, decomposing the equilibrium effects into distinct mechanisms.[23] Commission caps affect market outcomes through two competing forces identified in our conceptual framework. First, they weaken brokers' incentives to steer employers toward high-commission plans, improving allocative efficiency. Second, they reduce the prof-

[23]Unlike the partial-equilibrium analysis in the previous subsections, which pools across all geographic markets and restricts attention to employers with brokers, the commission cap counterfactual simulates a single geographic market (Finger Lakes, year 2019) and includes all employers—both brokered and non-brokered. Accordingly, employer surplus in this subsection includes hassle costs, which are relevant when the extensive margin of broker usage adjusts.

itability of broker-insurer relationships, potentially leading brokers to narrow their contracting networks and thereby softening insurer competition. To quantify the magnitude of each mechanism, we conduct a series of counterfactuals that sequentially allow different margins of adjustment.

Figure 9 displays the simulated effects on employer surplus as well as average wtp per employee under each counterfactual. Table A4 reports the complete set of equilibrium statistics for all market participants. The first counterfactual, denoted (0), represents the baseline scenario without any commission regulation. The subsequent counterfactuals apply the commission cap and sequentially allow employers, insurers, and brokers to adjust their behavior.

Figure 9. Decompose the welfare effects of commission cap at 1pp



Notes: Simulations reported for the year 2019, Finger Lakes region.

Our counterfactual (1) isolates the effect of the commission cap on employer choice patterns. We compare the baseline with a situation in which the cap mechanically truncates commissions above one percentage point, but insurers' prices, brokers' menu offerings, and employer-broker matches remain fixed at their baseline values. We allow only employers to adjust their plan choices in response to the new commission environment. Under this counterfactual, the cap weakens brokers' incentives to steer employers toward high-commission plans. As a result, employer choices become better aligned with their true preferences, improving allocative efficiency. Moving from counterfactual (0) to (1), employer surplus and average willingness to pay (wtp) per employee increase modestly, reflecting improved plan matching. Specifically, average employer surplus rises by approximately 0.9% relative to the baseline, while average WTP per employee increases by about 0.3%.

We then compute counterfactual (2), quantifying how insurer price responses amplify or attenuate the effects of reduced steering. We simulate a scenario where the commission cap affects both employer choice patterns and insurer pricing, but brokers' menu offerings and employer-broker matches remain fixed. We allow insurers to optimally reset prices given the new employer choice rules. Reduced steering lowers employers' effective willingness to pay for plan generosity, as the coefficient ρ_2 in equation (8) diminishes. This increased price sensitivity disciplines insurers' pricing power, leading to lower equilibrium premiums. Moving from counterfactual (1) to (2), average enrollee premiums decline by approximately 1.4% relative to the baseline. This reduction generates a substantial increase in average employer surplus—about 31% above baseline levels—while average wtp per employee rises more modestly, by roughly 1.1%.

Next, we compute counterfactual (3), measuring how employers' decisions to use brokers respond to the policy change. We simulate a scenario where the commission cap affects employer choices, insurer pricing, and employer-broker matching decisions, but brokers' contracting networks remain fixed. We allow employers to adjust their probability of using a broker given the new market conditions. Because brokers now distort choices less through steering, broker usage increases modestly. Across all employers, the broker utilization rate rises from 57.8% to 68.9%. This raises welfare through two channels: First, it reduces employer hassle costs, which decline from \$30.3 to \$22.3 per employee. Second, it lowers the share of employers selecting the non-offering option, from 16.9% in counterfactual (2) to 12.4% in counterfactual (3), thereby expanding employee access to insurance. As a result, moving from counterfactual (2) to (3), both average wtp per employee and average employer surplus rise further. Average employer surplus increases by about 11% relative to the baseline while average wtp per employee rises by 5.5%.

Taken together, counterfactuals (1) through (3) capture the relief of steering distortions induced by the commission cap. These adjustments generate welfare gains: average employer surplus rises by approximately 43% relative to the baseline and average wtp per employee increases by 6.9%, reflecting improved plan matching, reduced insurer markups, and expanded broker usage.

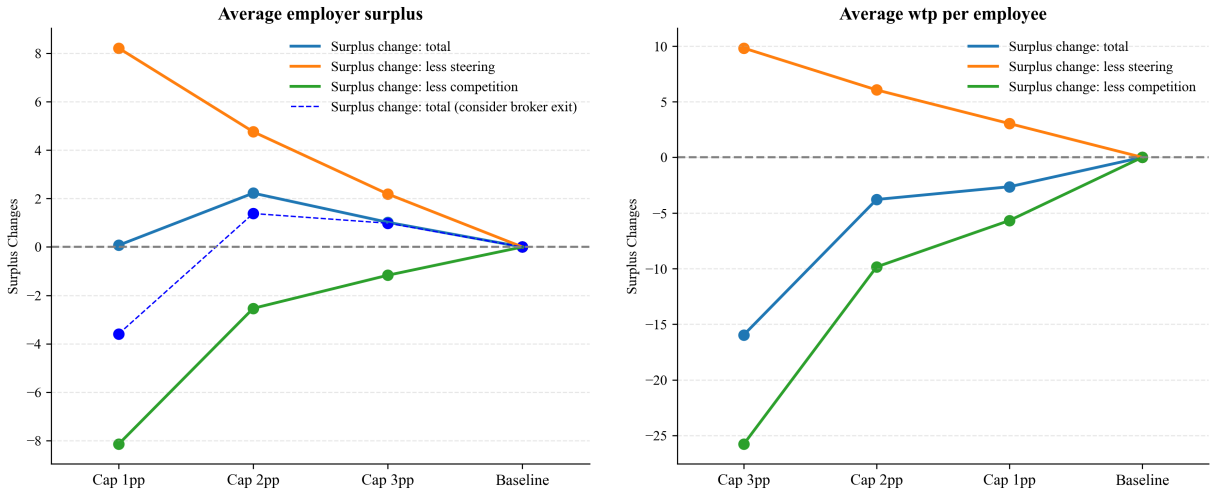
Finally, we compute counterfactual (4) to illustrate the full equilibrium effects, including brokers' contracting responses. This scenario corresponds to the complete equilibrium model, where the commission cap affects all margins: employer choices, insurer pricing, employer-broker matching, and broker-insurer contracting networks. The difference between counterfactuals (3) and (4) isolates the competitive-access effects of network contraction. The commission cap reduces the profitability of some insurer-broker rela-

tionships, leading brokers to drop insurers from their menus. This shrinkage in network breadth amplifies competitive-access distortions. With narrower choice sets, insurers face less competitive pressure, and prices rise in broker-mediated contracts. Moving from counterfactual (3) to (4), broker usage declines as the value of broker services diminishes, and more employers revert to not offering insurance. Employee welfare falls sharply, and employer surplus declines by an amount four times greater than the gains achieved in counterfactual (3).

The decomposition reveals a fundamental trade-off. Comparing counterfactuals (0) and (3) shows that commission caps improve allocative efficiency by mitigating broker-induced steering and reducing insurer market power, yielding a 43% gain in employer surplus. However, comparing counterfactuals (3) and (4) shows that commission caps also weaken broker incentives to maintain broad networks, leading to a reduction in employer surplus through competitive-access network contraction that is roughly comparable in magnitude to the gain. In net equilibrium, comparing counterfactuals (0) and (4), these two effects largely offset each other, and the commission cap generates a negligible increase in employer surplus of roughly 0.4%.

Figure 10 illustrates how employer surplus varies with the stringency of commission caps, decomposing the total effect into its constituent mechanisms. The horizontal axis measures the cap level in percentage points, with lower values corresponding to stricter regulation. The vertical axis shows the change in surplus: Panel A reports the average employer surplus, while Panel B presents the average willingness-to-pay per employee. The figure plots three series: the total change in surplus (solid blue line), the surplus change attributable to reduced steering (solid orange line), and the surplus change resulting from reduced competition due to network contraction (solid green line). In Panel A, we also plot the total surplus change after accounting for broker exits, shown as a dashed blue line.

Figure 10. The welfare effects of varying commission caps



Notes: Simulations reported for the year 2019, Finger Lakes region. Sample includes all employers. Employer surplus includes hassle costs.

Several patterns emerge from the figure. First, the steering-reduction benefits (orange line) increase monotonically as caps become stricter—lower commission caps more effectively curtail brokers' incentives

to steer employers toward high-commission plans. For example, tightening the commission cap from 3 percentage points to 1 percentage point increases the employer surplus gain from reduced steering from 2 to 8. Similarly, the gain in average willingness to pay per employee attributable to reduced steering rises from 3 to 10.

Second, the competition-reduction costs (green line) also grow in magnitude as caps become stringent, reflecting more pronounced network contraction when broker-insurer relationships become less profitable. As commissions decline, brokers have weaker incentives to sustain broad insurer networks, leading to diminished competition. For example, reducing the commission cap from 3 percentage points to 1 percentage point increases employer surplus losses from reduced competition from 1.5 to 8. Similarly, the decline in average wtp per employee attributable to weaker competition rises sharply, from 5.5 to 26. These patterns highlight how tighter caps can unintentionally erode competitive pressure in the market.

Third, and most importantly, the total effect (blue line) is non-monotonic. At a cap of approximately 3 percentage points, employer surplus increases modestly by about 1 relative to the unregulated baseline, as the benefits from reduced steering slightly outweigh the competition losses. Tightening the cap further to 2 percentage points initially amplifies this gain, generating a larger increase in total surplus for the average employer. However, once the cap is reduced to 1 percentage point, the intensified network contraction fully offsets these benefits, eliminating the earlier gains. When measured in terms of average willingness to pay per employee, stricter caps ultimately generate net welfare losses, as the adverse effects of reduced competition dominate the improvements from mitigating steering.

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.

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.

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.

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.

Chamberlain, Gary, "Panel data," *Handbook of Econometrics*, 1984, *2*, 1247–1318.

Ciliberto, Federico and Elie Tamer, "Market Structure and Multiple Equilibria in Airline Markets," *Econometrica*, 2009, *77* (6), 1791–1828.

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.

Fan, Ying and Chenyu Yang, "Estimating Discrete Games with Many Firms and Many Decisions: An Application to Merger and Product Variety," *Journal of Political Economy*, 2025, *133* (6), 1886–1931.

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.

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.

__ **and Michael Lettau**, "How elastic is the firm's demand for health insurance?," *Journal of Public Economics*, 2004, *88* (7-8), 1273–1293.

__ **, 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.

Ho, Katherine, "Insurer-Provider Networks in the Medical Care Market," *American Economic Review*, 2009, *99* (1), 393–430.

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.

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

Mundlak, Yair, "On the pooling of time series and cross section data," *Econometrica*, 1978, *46* (1), 69–85.

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.

Nosko, Chris, "Competition and Quality Choice in the CPU Market," *Manuscript, Harvard University*, 2010.

Pakes, Ariel, "Alternative Models for Moment Inequalities," *Econometrica*, 2010, *78* (6), 1783–1822.

__ , **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.

Petrin, Amil and Kenneth Train, "A control function approach to endogeneity in consumer choice models," *Journal of Marketing Research*, 2010, *47* (1), 3–13.

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), 000010151515380637 1083.

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*.

Wollmann, Thomas G, "Trucks Without Bailouts: Equilibrium Product Characteristics for Commercial Vehicles," *American Economic Review*, 2018, *108* (6), 1364–1406.

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 **50**

B Additional Details on Data **62**

B1 Panel Construction Using Form 5500 Filings 62

B2 Data from the Large National Representative Insurer 64

B3 Clarivate Data and Imputed Product Characteristics 64

C Theoretical Derivations and Microfoundation **67**

C1 Strategic Commissions and Equilibrium Existence 67

C2 Menu Formation Without Steering 67

C3 Menu Formation With Steering 67

C4 Welfare Decomposition 68

C5 Extensions 69

C6 Micro-foundation of the Steering Specification 69

D Additional Model and Estimation Details **76**

D1 Employee Surplus From Having Access to Employer-Sponsored Insurance 76

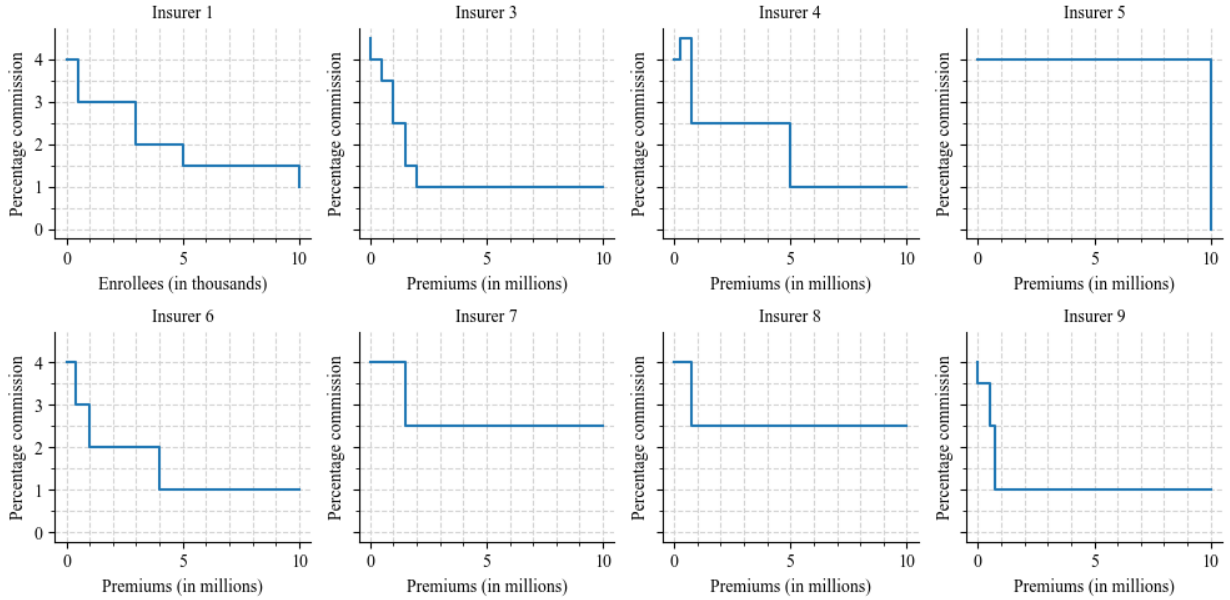
D2 Marginal Cost and the Enrolled Pool 79

D3 Distance is the Primary Factor in Broker-Employer Matches 81

D4 Multiplicity of Equilibria 82

A. Additional Tables and 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.

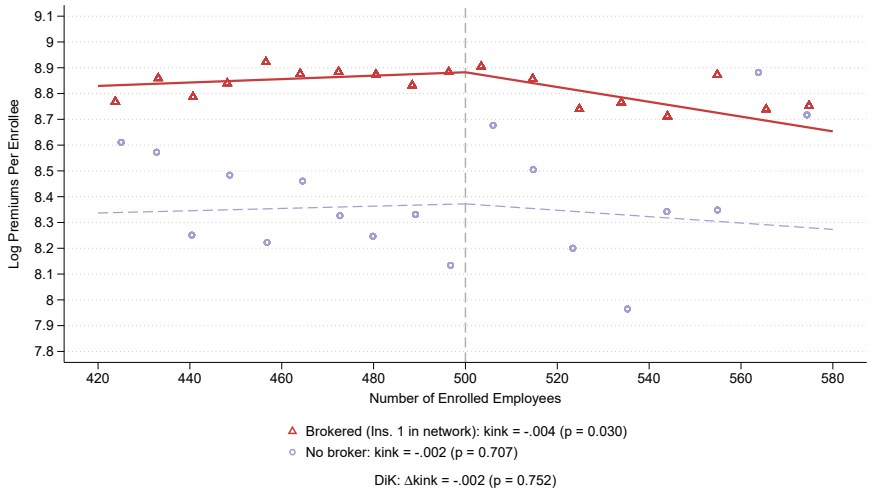
Table A1. RKD Robustness

	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			

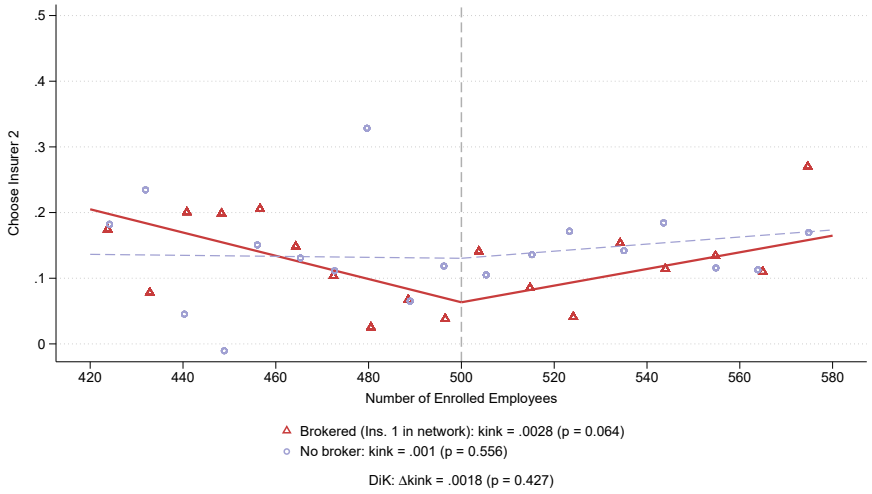
Notes: *Notes*: This table reports RKD kink estimates for the probability that an employer chooses Insurer 1 under alternative specifications. The sample comprises fully insured, single-carrier employers whose broker has Insurer 1 in network; the dependent variable and estimating equation are those of Figure 6. 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). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

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

(a). Premiums Per Enrollee

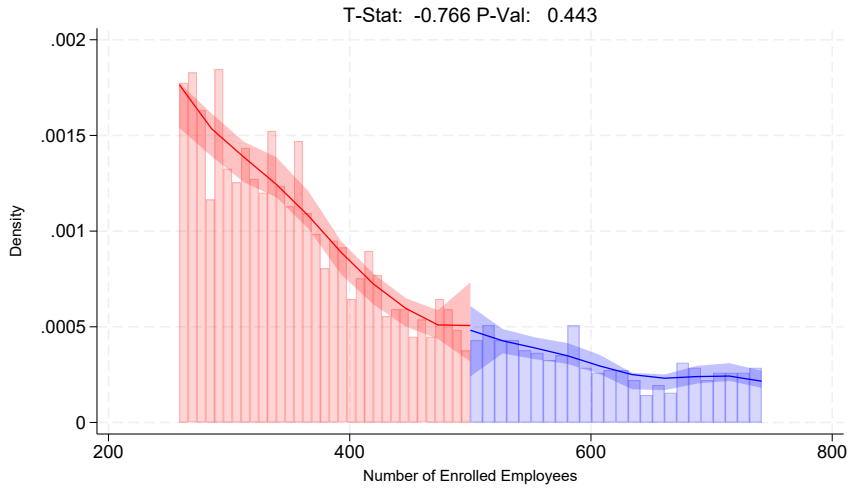


(b). Probability of Choosing Insurer 2



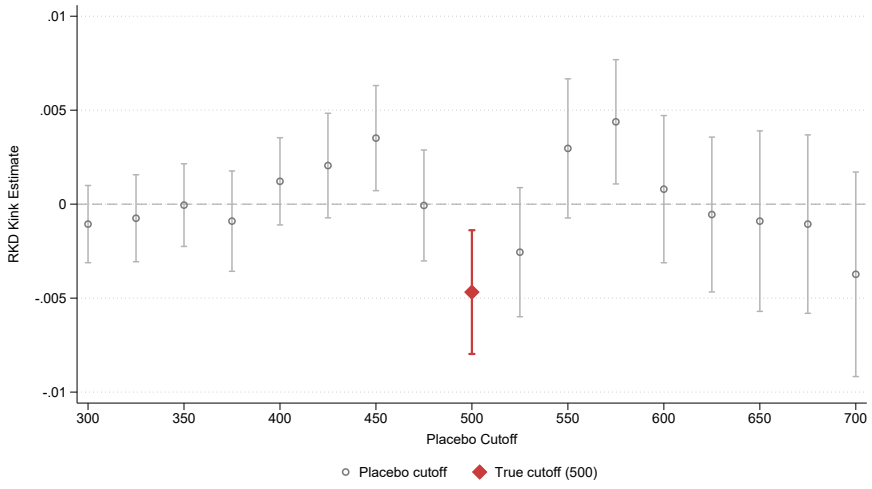
Notes: The figure reports difference-in-kinks estimates from equation (4). Panel (a) uses log premiums per enrollee as the outcome; panel (b) uses an indicator for choosing Insurer 2. The treatment group consists of brokered employers whose broker contracts with Insurer 1; the control group consists 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 6.

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.

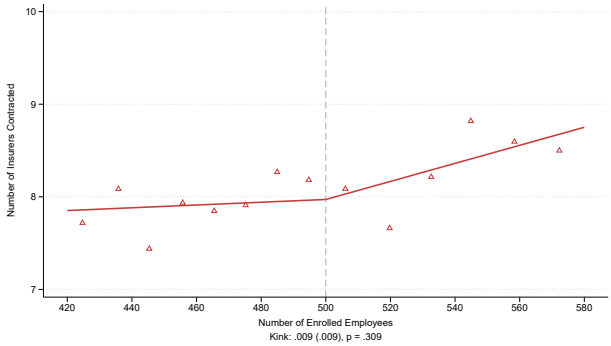
Figure A4. Placebo Cutoff Test



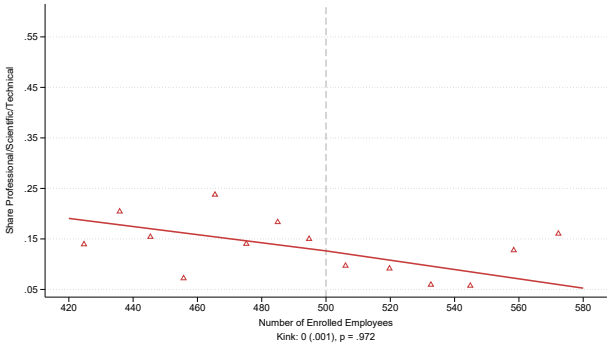
Notes: This figure reports RKD kink estimates from equation (3) at placebo cutoffs ranging from 300 to 700 enrolled employees, in increments of 25. The true cutoff at 500 (red diamond) is the commission kink for Insurer 1. Placebo cutoffs within one bandwidth (80 employees) of the true cutoff are sensitive 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

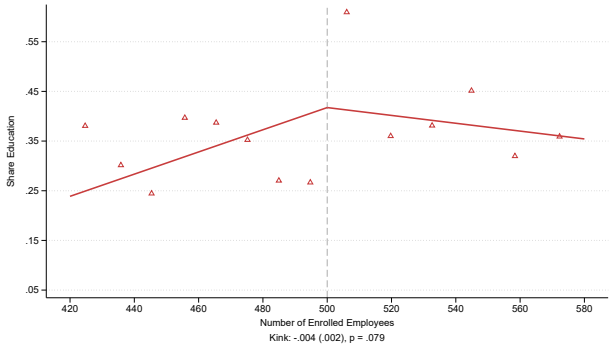
(a). Number of Insurers Contracted



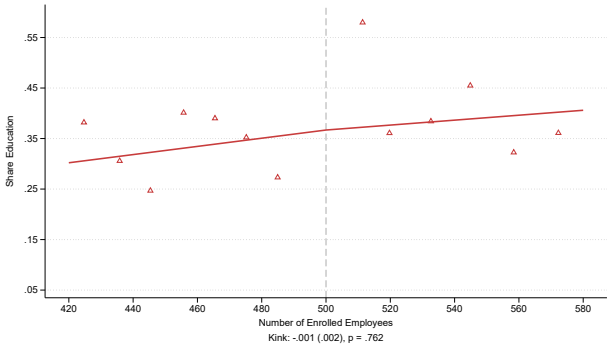
(b). Share Professional/Scientific/Technical



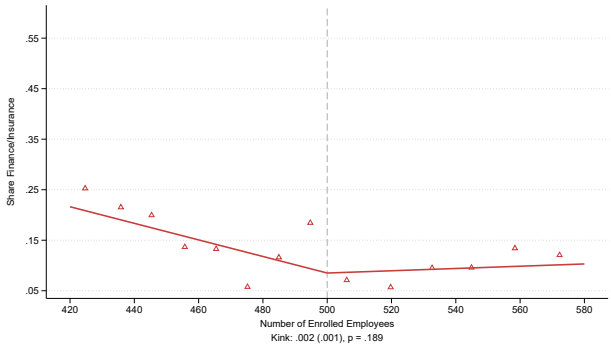
(c). Share Education



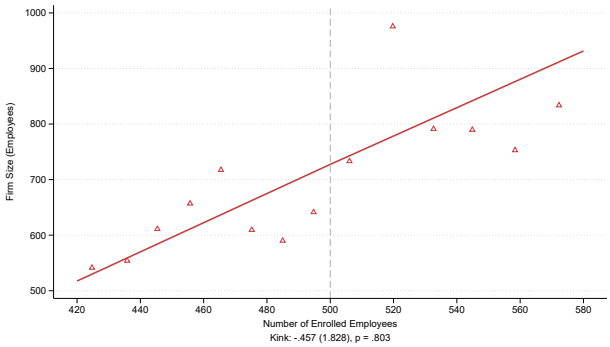
(d). Share Education: Donut (± 10)



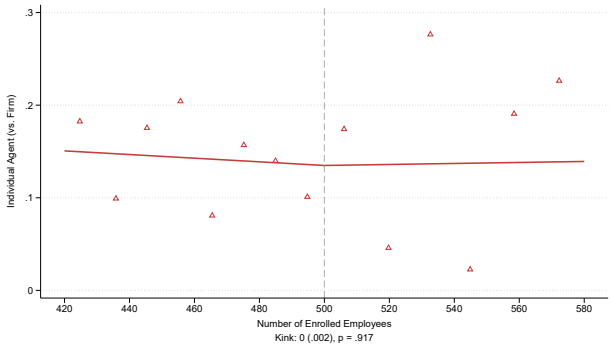
(e). Share Finance/Insurance



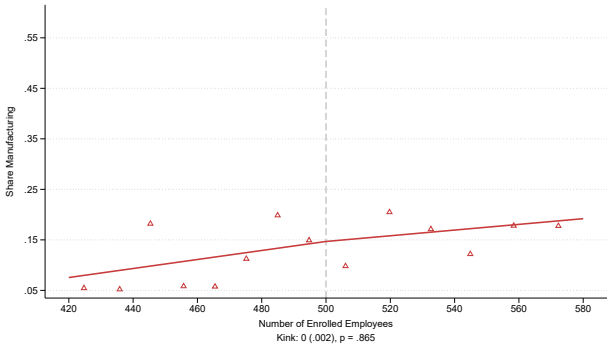
(f). Firm Size



(g). Individual Agent (vs. Firm)



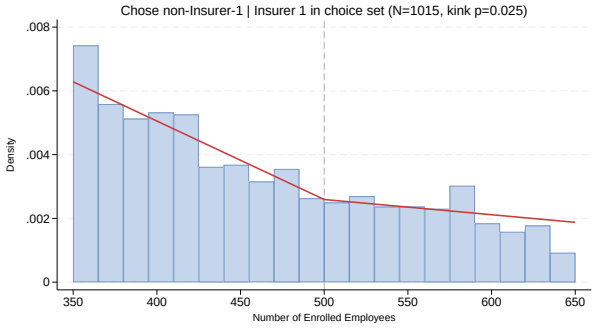
(h). Share Manufacturing



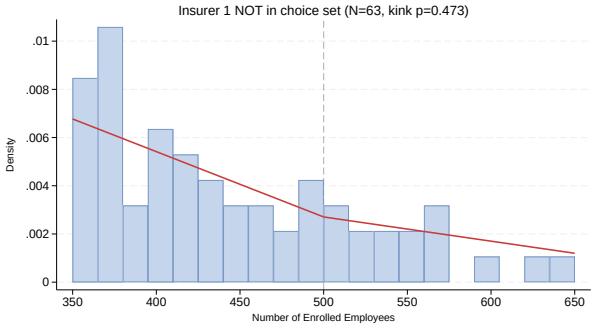
Notes: This figure tests whether predetermined covariates exhibit a kink at the commission cutoff of 500 enrolled employees. Each panel reports the RKD kink estimate from equation (3) with the covariate as the dependent variable. The specification matches the main analysis: local linear polynomial ($p = 1$), 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)



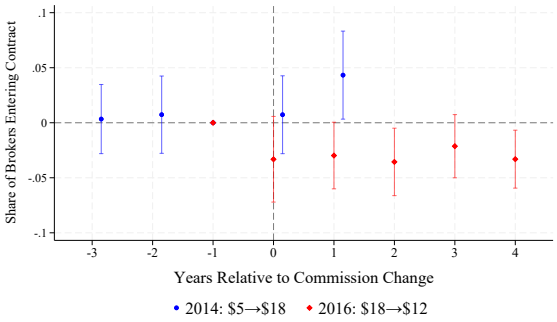
(b). Insurer 1 not in choice set (placebo)



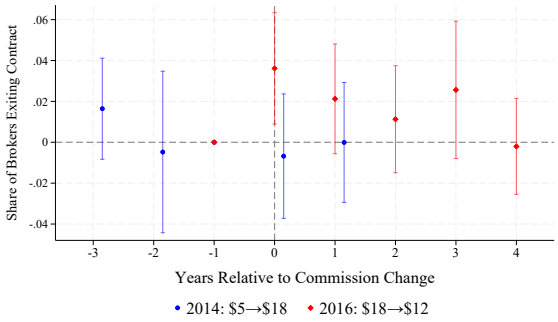
Notes: This figure gives a density-based view of the steering RKD. Panel (a) plots the density for brokered employers who have Insurer 1 in their choice set but chose a different carrier. Panel (b) plots the density for employers whose broker does not carry Insurer 1, and who therefore face no Insurer 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. Broker Network Response: Entry and Exit Margins

(a). Entry

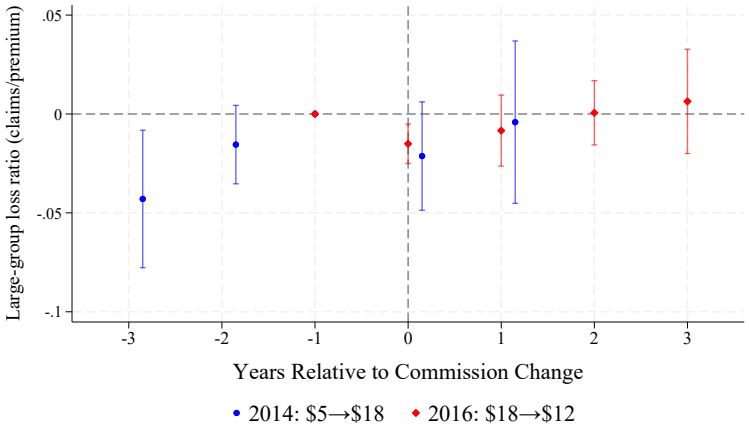


(b). Exit



Notes: This figure decomposes the broker-network event study in Figure 7 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 7. Specifications include insurer and county-by-year fixed effects. Standard errors are clustered at the county level, as in Figure 7, and bars denote 95% confidence intervals. Blue circles correspond to the 2014 commission increase, and red diamonds correspond to the 2016 commission decrease.

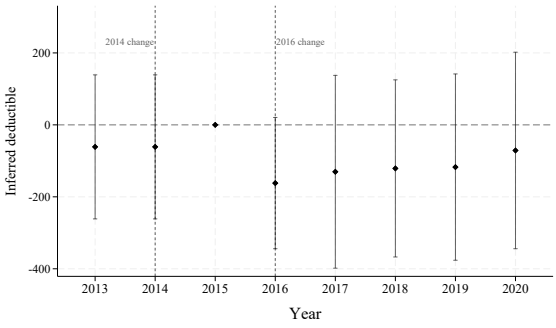
Figure A8. Placebo Event Study: Insurer 2's Large-Group Loss Ratio



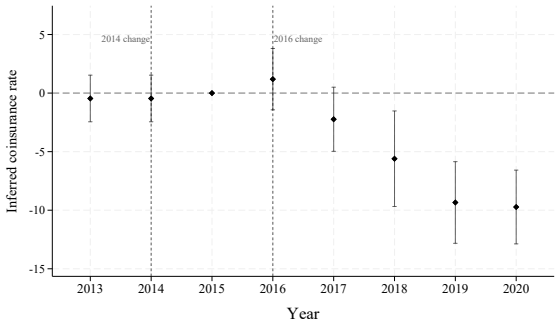
Notes: This figure presents stacked event-study estimates of the large-group medical loss ratio (claims divided by premiums), constructed from CMS Medical Loss Ratio (MLR) filings for New York and aggregated to the NAIC group-by-year level for 2011-2019. Blue circles correspond to the 2014 commission increase ($5 \rightarrow 18$), while red diamonds correspond to the 2016 commission decrease ($18 \rightarrow 12$). 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 A9. Placebo Event Studies: Insurer 2's Plan Design and Provider Network Breadth

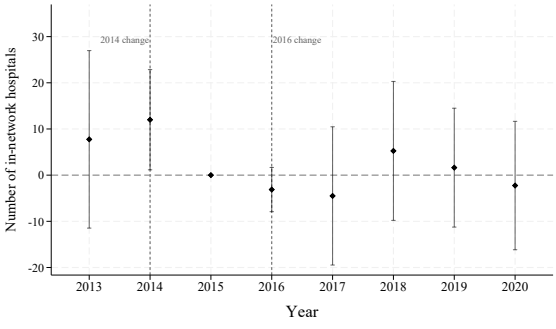
(a). Inferred deductible



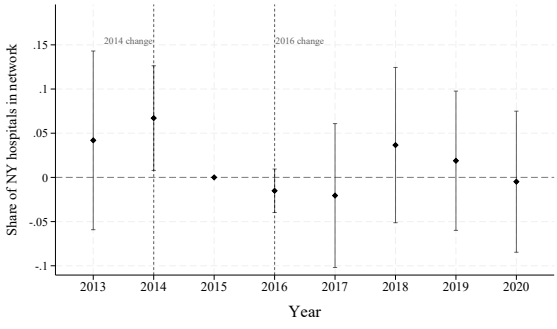
(b). Inferred coinsurance rate



(c). Number of in-network hospitals

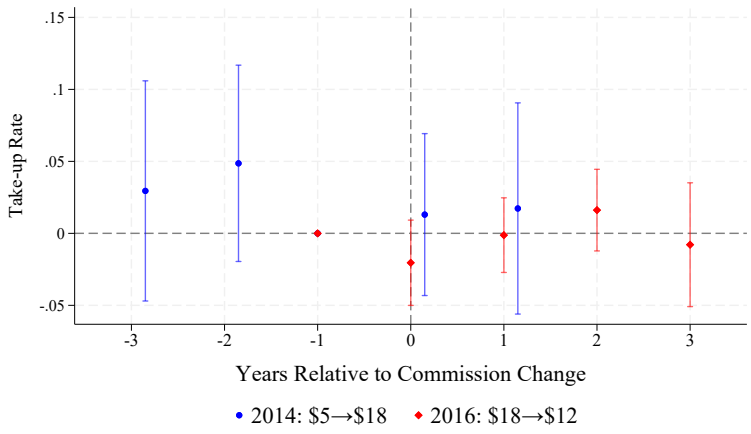


(d). Share of NY hospitals



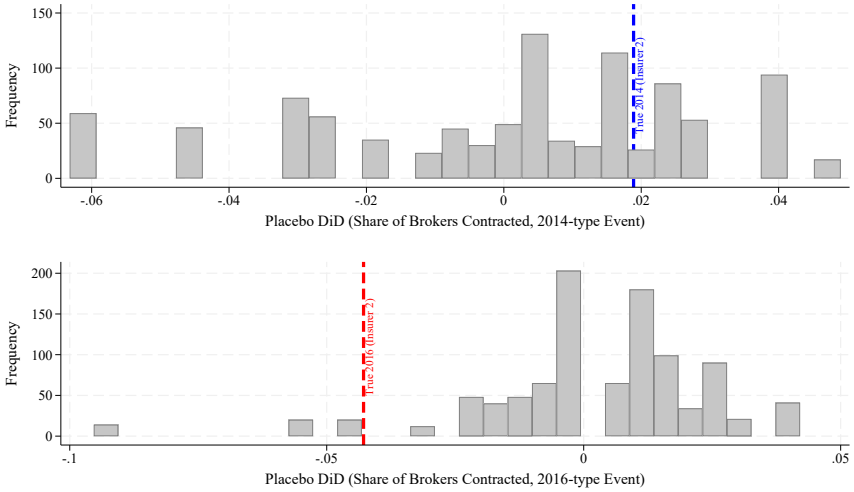
Notes: This figure presents stacked event-study estimates of plan characteristics, including the deductible (panel (a)) and coinsurance rate (panel (b)), inferred from claims data at the insurer-by-region-by-year level between 2013 and 2022, and the number of New York hospitals observed in-network for 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 at $t = -2$ and $t = 0$ correspond to calendar years 2014 and 2016, respectively. All specifications include insurer and region-by-year fixed effects. Standard errors are clustered at the region level, and bars denote 95% confidence intervals.

Figure A10. 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.

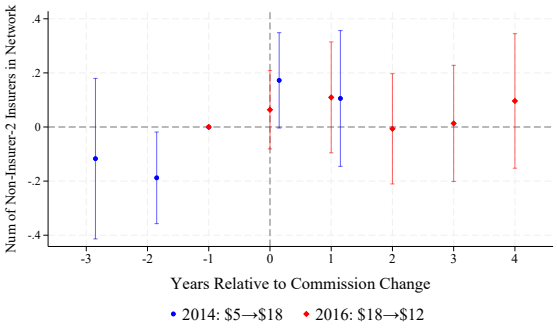
Figure A11. Broker-contracting Placebo: Insurer 2 vs. Placebo Insurers



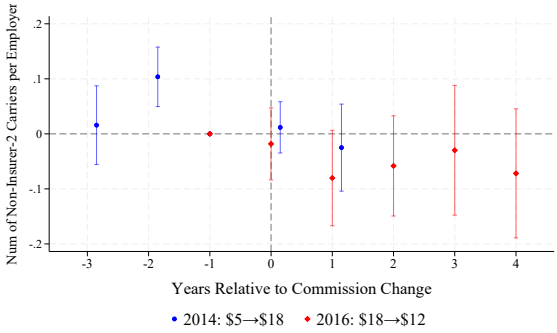
Notes: This figure reports the null distribution of 1,000 placebo difference-in-differences (DiD) estimates against Insurer 2's true estimate (vertical line). For each replication, we draw one placebo insurer uniformly from the nine non-Insurer-2 carriers and two ordered placebo event years (gap$\geq$2 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, year, and county fixed effects and standard errors clustered by county. The true estimates are Insurer 2's pooled DiD coefficients from the actual commission changes.

Figure A12. Placebo Event Studies: Treated Brokers' Choice Sets and Client Composition

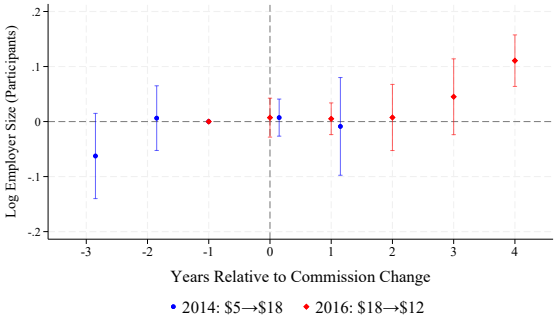
(a). Non-Insurer-2 insurers per broker



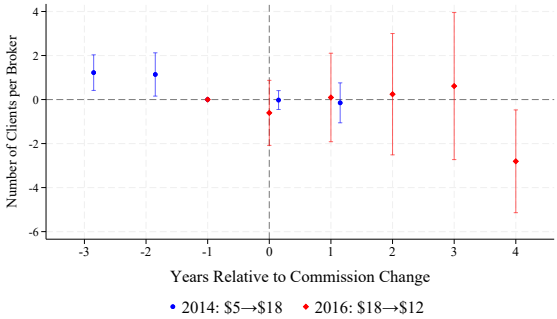
(b). Non-Insurer-2 carriers per employer



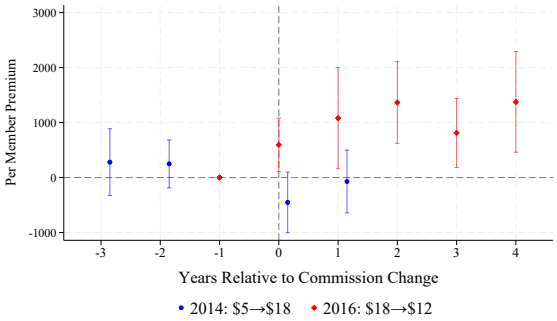
(c). Log employer size



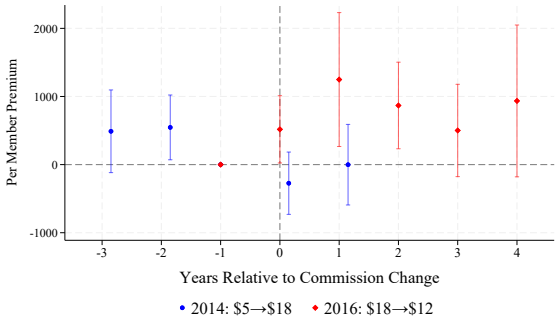
(d). Clients per broker



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, county, and year fixed effects, as well as employer and insurer fixed effects where applicable. Standard errors are clustered at the county level, and error bars denote 95% confidence intervals.



(a). Baseline

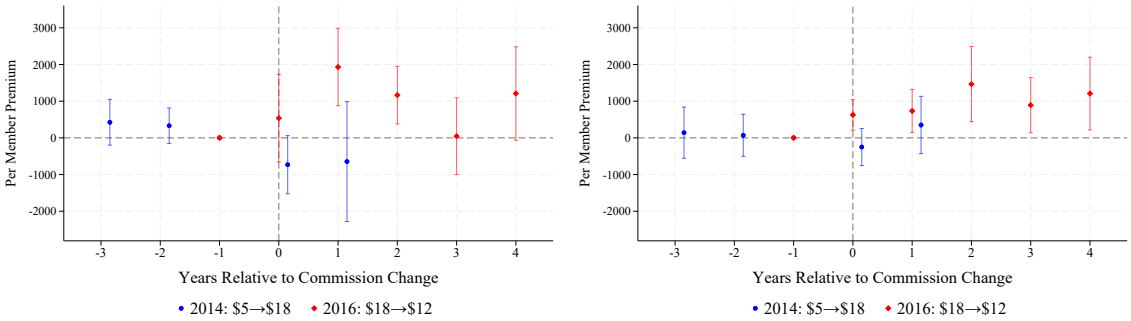


(b). Stayer pairs only

Figure A13. Rival-Premium Response: Baseline Versus Stayer Employer-Insurer Pairs

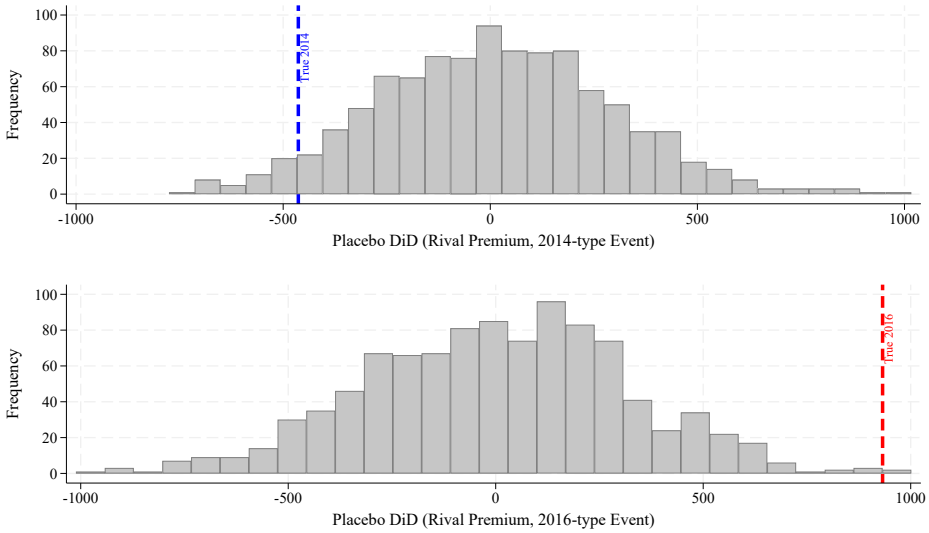
Notes: This figure replicates the rival-premium event study shown in Figure 8. Panel A reproduces the baseline specification using per-member premiums, winsorized at the 1st and 99th percentiles, for fully insured employers with at least 100 participants; Insurer 2 is excluded from the sample. Panel 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 broker, employer, insurer, county, and year fixed effects. Standard errors are clustered at the county level, as in Figure **??**, and bars denote 95% confidence intervals.

Figure A14. Rival-Premium Response by Broker-Network Overlap with Insurer 2



Notes: This figure decomposes the rival-premium event study in Figure 8 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 A13.

Figure A15. 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 and 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.

Table A2. Triple Difference: Rival Premiums Response by Broker-network Overlap

	(1) 2014	(2) 2016	(3) Pooled
Post × Exposed × High-overlap	-943.7***	-69.2	263.4*
	(336.0)	(324.5)	(135.6)
Post × High-overlap	-90.9	223.9	56.7
	(189.5)	(235.5)	(146.6)
Exposed × High-overlap	2700.4***	-145.2	-1404.1***
	(861.5)	(749.3)	(431.5)
Post × Exposed	-225.6	920.6***	597.7**
	(203.8)	(301.9)	(224.7)
N	4,257	5,380	9,637
Adj. R^2	0.711	0.699	0.905

Notes: This table formalizes Figure A14 using pooled regressions. The sample consists of fully insured employer-broker-carrier-year observations, excluding Insurer 2. *Exposed* identifies the treated broker cohort, as defined in Figure 8. *High-overlap* equals one for rival insurers whose pre-period share of brokers that also contracted with Insurer 2 exceeds the sample median. Column (3) pools the two event windows and sign-inverts the 2014 outcome, so that a positive triple-difference coefficient in very specification indicates that the premium response is concentrated among close-substitute rivals when Insurer 2 decreases commissions. Standard errors are clustered at the county level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table A3. Observed versus Simulated Equilibrium

	(1)	(2)	(3)	(4)	(5)
	Market Share			Price	
	Observed	Observed price Simulated choice	Simulated price Simulated choice	Observed	Simulated
Employers with brokers					
Aetna	13.4	14.2	14.2	9621	9573
MVP Health	4.8	3.7	3.7	6974	6445
Empire/Anthem	8.5	10.0	10.0	9651	9327
Oxford/United	24.3	22.5	22.5	9988	9882
Healthnow	6.3	5.9	6.0	8331	8329
Connecticut/Cigna	12.4	12.4	12.4	8955	8810
HIP of NY/Emblem	2.8	2.4	2.4	8809	8809
Excellus	22.5	23.7	23.7	8363	8155
CDPHP	2.6	2.6	2.6	8463	8463
Independent Health	2.3	2.5	2.5	7753	7749
Employers without brokers					
Uninsured	44.0	17.6	22.1	0	0
Aetna	8.2	8.9	11.2	9522	9500
MVP Health	1.6	0.4	0.4	6512	6566
Empire/Anthem	4.9	17.1	10.0	16713	9628
Oxford/United	15.2	22.4	29.4	9302	9221
Healthnow	5.4	2.3	3.0	8815	8852
Connecticut/Cigna	7.6	6.0	7.1	8849	8939
HIP of NY/Emblem	1.6	15.7	3.6	18681	9148
Excellus	2.7	8.1	10.9	8476	8421
CDPHP	2.2	1.1	1.3	8489	8569
Independent Health	6.5	0.5	0.8	7513	7553

Table A4. The welfare effects of commission cap at 1pp

	CF0	CF1	CF2	CF3	CF4
Overall Market					
Broker usage rate	0.578	0.578	0.578	0.689	0.257
Average choice set size	5.572	5.572	5.572	5.075	7.233
Average worker takeup rate	0.629	0.630	0.636	0.670	0.557
Average price per employee	405.910	407.711	403.715	427.747	345.217
Average WTP per employee	142.225	142.604	144.239	152.042	126.272
Average non-offering rate	0.169	0.169	0.169	0.124	0.279
Average price per enrollee	427.763	429.098	423.161	424.636	416.734
Average cost per enrollee	378.400	380.218	378.939	379.159	376.760
Average markup	1.135	1.133	1.120	1.124	1.109
Average commission percentage	0.019	0.006	0.006	0.007	0.003
Average employer choice value	10.979	11.143	17.087	11.143	34.202
Average hassle costs	30.301	30.301	30.301	22.252	53.452
Average employer surplus	−19.322	−19.157	−13.214	−11.110	−19.250
Among firms that use brokers					
Average choice set size	3.798	3.798	3.798	3.755	5.016
Average worker takeup rate	0.747	0.748	0.759	0.761	0.770
Average price per employee	486.821	489.938	483.023	487.847	478.554
Average WTP per employee	167.602	168.256	171.086	172.315	174.894
Average non-offering rate	0.000	0.000	0.000	0.000	0.000
Average price per enrollee	429.209	431.518	421.243	424.206	413.012
Average cost per enrollee	371.500	374.647	372.433	375.383	368.887
Average markup	1.161	1.157	1.135	1.134	1.123
Average commission percentage	0.033	0.010	0.010	0.010	0.010
Average employer choice value	−10.867	−10.583	−0.296	−2.309	6.308
Average hassle costs	0.000	0.000	0.000	0.000	0.000
Average employer surplus	−10.867	−10.583	−0.296	−2.309	6.308
Among firms that do not use brokers					
Average choice set size	8.000	8.000	8.000	8.000	8.000
Average worker takeup rate	0.468	0.468	0.468	0.467	0.483
Average price per employee	295.165	295.165	295.165	294.507	299.094
Average WTP per employee	107.493	107.493	107.493	107.098	109.453
Average non-offering rate	0.400	0.400	0.400	0.400	0.375
Average price per enrollee	420.041	420.041	420.041	419.625	412.360
Average cost per enrollee	381.831	381.831	381.831	381.273	373.464
Average markup	1.103	1.103	1.103	1.103	1.107
Average commission percentage	0.000	0.000	0.000	0.000	0.000
Average employer choice value	42.866	42.866	42.866	43.038	45.883
Average hassle costs	71.556	71.556	71.556	71.369	71.629
Average employer surplus	−28.689	−28.689	−28.689	−28.331	−25.747

Notes: Simulations reported for year 2019 finger lakes region. CF0: baseline. CF1: allow change choice patterns. CF2: allow change insurer pricing. CF3: allow change broker usage decision. CF4: alloq change broker eqlm menu offering

Table A5. Employer Surplus across Geographic Markets

	Employer Surplus			Δ Surplus		Comm.	corr(Δsurplus,
	CF0	CF1	CF2	CF1−CF0	CF2−CF1	CF0 (%)	Δcomm)
Capital Region	26.46	30.99	28.82	4.53	−2.17	3.00	0.91
Central NY	−13.54	−12.30	−17.12	1.24	−4.82	2.87	0.95
Finger Lakes	16.42	18.02	15.29	1.60	−2.73	2.28	0.97
Mid Hudson	62.13	64.96	60.77	2.83	−4.19	3.05	0.93
Long Island	51.70	54.36	50.06	2.66	−4.29	3.10	0.89
Mohawk Valley	31.06	33.88	31.78	2.82	−2.10	2.80	0.97
New York City	77.84	81.68	78.12	3.84	−3.56	2.93	0.90
North County	24.61	25.85	23.50	1.24	−2.35	1.89	0.95
Southern Tier	21.86	22.63	18.99	0.78	−3.64	2.11	0.92
Western NY	25.64	29.60	27.41	3.96	−2.18	3.04	0.77
All Regions	52.50	55.69	52.25	3.19	−3.44	2.87	0.94

Notes: CF0: baseline with commission steering and status-quo equilibrium prices. CF1: remove steering ($\rho_2 = 0$), hold prices fixed. CF2: remove steering and allow insurers to re-optimize prices. Employer surplus in $/employee/month, excluding hassle costs. Comm. CF0: baseline enrollment-weighted average commission rate. corr(Δsurplus, Δcomm): within-region employer-level weighted correlation between surplus gain (CF1 − CF0) and commission drop. All Regions row: enrollment-weighted averages across regions; pooled employer-level correlations. Simulations for year 2018, among employers with brokers.

Table A6. Employee WTP across Geographic Markets

	Employee WTP			Δ WTP		Correlations		
	CF0	CF1	CF2	CF1−CF0	CF2−CF1	comm, wtp	Δwtp, Δcomm	Δwtp, Δp
Capital Region	198.70	195.10	195.61	−3.61	0.51	0.61	−0.41	0.73
Central NY	174.03	174.40	173.11	0.37	−1.29	0.08	0.69	−0.46
Finger Lakes	167.08	166.31	165.87	−0.77	−0.45	−0.05	0.02	−0.75
Mid Hudson	180.87	180.96	179.85	0.09	−1.11	0.13	0.07	0.53
Long Island	182.77	183.32	182.24	0.55	−1.08	−0.49	0.28	0.41
Mohawk Valley	177.52	174.85	175.05	−2.67	0.20	0.30	−0.82	0.25
New York City	194.68	194.44	193.80	−0.24	−0.63	−0.50	0.61	−0.64
North County	180.49	178.48	178.27	−2.00	−0.21	0.12	−0.76	−0.67
Southern Tier	166.41	167.11	166.24	0.71	−0.87	−0.63	0.79	0.75
Western NY	184.26	181.57	181.66	−2.68	0.08	0.17	−0.10	−0.00
All Regions	186.56	185.96	185.37	−0.59	−0.59	0.18	−0.07	0.27

Notes: CF0: baseline with commission steering and status-quo equilibrium prices. CF1: remove steering ($\rho_2 = 0$), hold prices fixed. CF2: remove steering and allow insurers to re-optimize prices. Employee WTP in $/employee/month. Correlations are within-region employer-level weighted correlations. comm, wtp: corr(commission_{CF0}, wtp_{CF0}). Δwtp, Δcomm: corr(Δwtp_{CF1}, Δcomm). Δwtp, Δp: corr(Δwtp_{CF1}, Δprice_{CF1}). All Regions row: enrollment-weighted averages across regions; pooled employer-level correlations. Simulations for year 2018.

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.

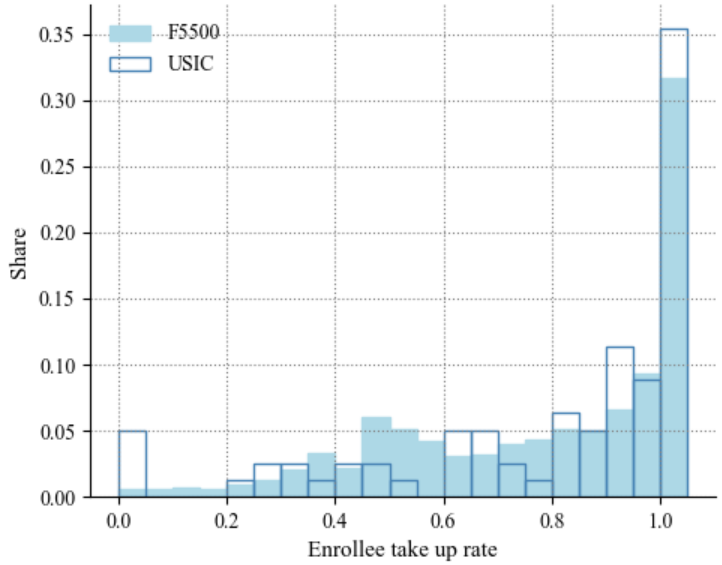
First, we begin with Schedule A filings of Form 5500, which records employers' benefit contracts. Each filing reports the carrier, premiums paid, and the number of covered enrollees (employees and dependents) for each plan. Second, we restrict attention to health insurance filings. Other benefit types, 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 A16. 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 A16 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 A7 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 A8 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 A7. 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 A8. 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 A9.

Table A9. 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 3b 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_{fjm}^E 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 (15)). 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 A10 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 A10. 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 A11 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 A11. 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 A12.

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 A12. 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 A13.

Table A13. 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 A13 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 A14 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 A14. 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.