Stocking Under the Influence: Spillovers from Commercial Drug Coverage to Medicare Utilization*

Emma B. Dean, Josh Feng, Luca Maini†

May 3, 2026

Abstract

We show that variation in state-level commercial coverage of physician-administered drugs impacts Medicare patients' utilization: a 10pp increase in commercial exclusion rates leads to a 0.9-1.7pp reduction in share among Part B beneficiaries. Facilities, rather than physicians, drive prescribing variation, likely through the preferred stocking of products with better commercial coverage. Motivated by these findings, we calibrate a model of manufacturer competition over insurance coverage and facility stocking, and find that the stocking channel weakens the government's ability to influence Medicare utilization and that manufacturers can strategically use formulary coverage to soften price competition over facility stocking.

*We would like to thank Maura Coughlin, Ashvin Gandhi, Manuel Hermosilla, Daniel Kaliski, Pragya Kakani, Genevieve Kanter, Tim Layton, Jong Myeong Lim, Mark Shepard, Andrew Smith, and Doug Staiger for detailed comments and feedback. This paper benefitted from comments by audiences at the American European Health Economics Study Group at Catania University, the Annual Health Economics Conference at Stanford, ASHEcon, the Bates White Life Science Symposium, the Boston Conference on Markets and Competition, the Brookings Institution, the Dartmouth Institute, Florida State University, Indiana University, the International Industrial Organization Conference, the Southern Economic Association meetings, the Southeastern Health Economics Study Group, Tulane University, University of Massachusetts Amherst, University of Montreal, Utah Applied Economics Conference, Wayne State University, and Weill Cornell Medicine. This work was supported by the National Institute for Healthcare Management Foundation. We thank Steve Porcelli at MMIT for sharing their data and providing additional feedback. Katie Garrett provided excellent research assistance.

†Emma B. Dean: The Dartmouth Institute for Health Policy and Clinical Practice, emma.c.b.dean@dartmouth.edu. Josh Feng: University of Utah Eccles School of Business, josh.feng@eccles.utah.edu. Luca Maini: Harvard Medical School, maini@hcp.med.harvard.edu.

"Insurer decisions regarding reimbursement policies have a dramatic impact on which infliximab product will be stocked by healthcare providers such as hospitals and clinics. [...] there is almost no chance that providers will pay for a product that is not widely covered by insurers for fear of stocking a product that will not be reimbursed"

PFIZER INC. V. JOHNSON & JOHNSON AND JANSSEN BIOTECH, INC., COMPLAINT

In the United States, government-regulated health insurance markets like Medicaid and Medicare coexist alongside a large commercial market. Although health plans across these two markets often serve different patients, they contract with the same facilities and providers. As a result, when facilities and providers make decisions that could affect the quality or type of care (e.g., staffing and acquiring specific machines and medical supplies), they must consider the combined incentives arising from all payers. A theoretical literature (see, e.g., Glazer and McGuire, 2002) argues that this market structure would lead patients treated in the same facility to receive similar care regardless of payer—a concept called "commonality of care". An untested but policy-relevant implication of this theory is that the design of commercial insurance plans can affect the care received by patients covered by a public payer.

In this article, we provide direct evidence that commercial health insurance design affects the care of patients covered by government-sponsored insurance through the commonality of care channel, by showing that commercial coverage of physician-administered drugs affects the utilization of Traditional Medicare (TM) patients. Furthermore, although a growing literature shows that physician practice style can lead to these kinds of spillover effects across patients (see, e.g., Glied and Zivin, 2002; Baker, 2003; Baicker et al., 2013; Einav et al., 2020; Barnett et al., 2023), our analysis rules out changes in physician practice style as the primary mechanism. Instead, we show that variation in utilization is driven by facilities, which likely prefer to stock drug brands with broad coverage among commercial health insurance plans to unlock additional pricing concessions from manufacturers. In the final part of the article, we calibrate a model of competition between two drug manufacturers that incorporates both insurance formulary setting and facility stocking choices. In counterfactual analyses we show that Medicare reimbursement reforms have limited effects on prices and spending, as stocking remains primarily driven by commercial insurance.

Our analysis focuses on the coverage of physician-administered drugs, which are administered by a healthcare professional in outpatient facilities.[1]Commercial health plans often exclude certain drugs or impose additional access restrictions, such as requiring prior authorization from the insurer before dispensing a given medication (see, e.g., Anderson et al., 2022).[2] Conversely, TM beneficiaries receive coverage for physician-administered drugs under Medicare Part B, which covers virtually all drugs approved by the Food and Drug Administration.

[1]Examples of these drugs include chemotherapy infusions, intravenous drugs for autoimmune conditions, monoclonal antibody infusions for cancer and macular degeneration, and iron infusions for anemia.

[2]Physician-administered drugs can also be covered under medical benefits, but medical benefit policies are generally consistent with coverage restrictions in prescription drug formularies.

Although Medicare Part B covers all physician-administered drugs, TM beneficiaries can only receive drugs that are available at the facility where they seek care, which we argue depends on commercial coverage. Facilities typically stock a limited set of drugs, both to simplify logistics such as inventory management and storage space, and to facilitate concentrating prescriptions on a single brand, which unlocks volume-based discounts from manufacturers (Dean et al., 2024). In these cases, facilities may focus their prescribing efforts on drugs broadly covered by commercial insurance formularies to maximize the chance that patients will be able to receive the drug. This incentive can create a spillover effect from commercial coverage decisions to facility stocking choices, and introduce a channel for equilibrium outcomes in the commercial market to influence patients covered by government-sponsored insurance plans.

In our empirical analysis, we test the existence of this link between commercial coverage and TM utilization patterns in two samples of physician-administered drugs. Our "core" sample consists of all branded physician-administered drugs, which are divided into markets based on therapeutic substitutability. We also study a second, smaller sample of substitution classes defined by a reference biologic product and its biosimilar competitors—clinically equivalent but non-identical alternatives that enter the market after a reference biologic loses patent protection. Biosimilars are considered bioequivalent to their reference biologics, making them obvious targets for a volume-based or share-based contract at the facility level. This "biosimilar" sample also yields important insights in its own right, given the existing evidence of uneven uptake of biosimilars among Medicare patients (Socal et al., 2020; Dean et al., 2021) and recent legislation in the Inflation Reduction Act to address this by providing greater incentives to administer biosimilars.

Our analysis proceeds in three steps. In the first step, we compare variation in the commercial coverage of drugs to their market share among Part B patients at the state-year level. Although we find a strong correlation, a key limitation is that unobserved variation in state-level patient and physician preferences rather than insurance coverage could be driving the results. To address this challenge, we employ a shift-share instrumental variable research design that isolates changes in coverage driven by national-level plans that are offered to employers by insurance companies. These plans include standard drug coverage terms that are negotiated by intermediaries known as Pharmacy Benefit Managers (PBMs). Although some plans design their own formularies and rely only on PBMs for rebate negotiations, many plans use off-the-shelf formularies designed by PBMs, known as national formularies, available to all plans across the country.[3] We use data from MMIT Network Solutions to track coverage of drug brands across all commercial insurance plans and calculate variation in state-level exposure to each national formulary. We then use coverage changes on each national formulary to instrument for state-level changes in commercial exclusion rates, and account for inference issues by following the approach in Borusyak et al. (2022).

Our instrumental variables estimates show that higher brand coverage rates in commercial plans result in higher utilization rates among Part B beneficiaries. Among drugs in our core sample, a 10 percentage point increase in brand coverage by commercial plans leads to a 0.9 to 1.7 per-

[3] In 2019, 45.5 percent of individuals covered by commercial insurance are on a plan that adopts a national formulary.

centage point increase in that brands' market share among Part B beneficiaries. In our secondary sample of biologics and biosimilars, with substitution classes defined at the molecule level, a 10 percentage point increase in commercial coverage leads to a 0.9 to 3 percentage point increase in Part B market share.

In the second step of our analysis, we provide evidence that facility stocking decisions (rather than physician behavior) are likely the primary transmission mechanism for this observed spillover, by studying variation in the prescription patterns of physicians who simultaneously practice in different facilities. Our analysis is conceptually akin to physician "mover" studies (see, e.g., Finkelstein et al., 2021). Focusing on physicians prescribing products within a specific substitution class at two facilities in the same year, we compare the change in within-physician prescribing behavior across the two facilities to the change in the facility-level market shares. We find a broad range of estimates suggesting that, across substitution classes, facilities are responsible for between 20–100% of the variation in product market share. Facility factors tend to matter more in substitution classes containing less horizontally differentiated products (e.g., contrast dyes for imaging), and less so in broader substitution classes containing more differentiated products (e.g., chemotherapy agents). Coefficients from the biosimilar analysis are on the high end of this spectrum, with facility-level factors explaining between 80–90% of variation in prescribing behavior. We link this result to the previous step by repeating the shift-share estimation but splitting the core sample between drug classes whose facility coefficient is above and below the median. Consistent with our stocking hypothesis, we find a much stronger relationship between commercial coverage and Part B market share in the group of drug classes for which facilities drive more of the variation.

Finally, in the third step of our analysis, we develop a model of price competition between two manufacturers that incorporates both commercial formulary coverage choices and exclusive facility stocking, allowing us to study how the two layers interact in equilibrium. We calibrate the model using data from a market with one reference product and one biosimilar, which represents a valuable case study given existing evidence that biosimilar adoption has been slow despite biosimilars being cheaper than their reference products (see e.g., Hong et al., 2024). We find that the reference biologic has large, built-in advantages stemming from patient, facility and insurer preferences, which are the primary reasons it maintains a dominant market share despite being more expensive.

We then use the model to assess how prices, market shares, and spending would change under three counterfactual scenarios aimed at improving biosimilar adoption: (i) an increase in Medicare payments for biosimilar products, (ii) a requirement that commercial plans cover all drugs, and (iii) a biosimilar coverage mandate. The stocking channel limits Medicare's ability to influence utilization, because Medicare payment reform has a negligible effect the marginal stocking choice of facilities. That choice tends to be dominated by commercial insurers, whose reimbursements are much higher (Robinson et al., 2024). Commercial coverage mandates provide an advantage to biosimilars at the cost of higher commercial spending, but their overall effect is blunted by equilibrium responses in the stocking game, where the reference biologic competes more aggressively

on price and successfully retains its advantage. These dynamics also illustrate a broader pattern of this two-layer game, whereby restricting competition on formulary coverage intensifies it in the stocking game.

The primary contribution of this paper is providing empirical evidence for the "commonality" of care hypothesis (Glazer and McGuire, 2002). Existing work has documented similarities in the level of care patients receive when treated in the same facility (Dranove and White, 1998; Grabowski et al., 2008). We use plausibly exogenous geographic variation in commercial insurance—which drives effects in the Glazer and McGuire (2002) framework—and trace its impacts on TM patients.[4]

More broadly, commonality of care can also be interpreted as a special case of two more general spillover effects. The first are the spillovers between public and commercial health insurance. Most existing work focuses on the effect of Medicare and Medicaid insurance design on commercial equilibrium outcomes, partly because of the availability of policy changes and natural experiments (see, e.g., Duggan and Scott Morton, 2006; Clemens and Gottlieb, 2017; Feng et al., 2023) . Other work has focused on capacity constraints, under which increased coverage in one segment can lead to decreased utilization in the other (Garthwaite, 2012; Richards and Tello-Trillo, 2019; Gandhi, 2023; Hackmann et al., 2024). A final strand of the literature has shown that policy-driven change in physician practice style can generate spillovers across segments (see, e.g., Glied and Zivin, 2002; Baker, 2003; Baicker et al., 2013; Einav et al., 2020; Barnett et al., 2023). The channel we identify focuses on spillovers from private to public insurance, has the opposite prediction to a capacity constraints mechanism, and operates at the facility level rather than the physician level.

The second are the spillovers that arise in markets where consumers purchase goods from intermediaries rather than directly from producers. In these markets, producers compete for shelf space controlled by intermediaries. Because intermediaries make stocking decisions based on aggregate demand, access to products depends on the preferences of all consumers served by the same intermediary, generating spillovers across consumers. Empirical evidence of these demand spillovers has been documented in the infant formula market, where brands that win government procurement contracts for subsidized sales to low-income households subsequently see higher overall sales, consistent with the contract expanding retail shelf space for the winning brand (Huang and Perloff, 2014; Abito et al., 2024).[5] A key empirical challenge in these settings is

[4] Finkelstein (2007) and Clemens and Gottlieb (2014) also include results consistent with the existence of facility-driven spillovers from Medicare to commercial patients, though neither paper frames them as such. Finkelstein (2007) shows that the introduction of Medicare led to a shift in the supply of hospitals, and presents suggestive evidence of a spillover effect on non-elderly patients. Clemens and Gottlieb (2014) show that an increase in Medicare reimbursement rates led to an increase in the diffusion of MRI technology, which in turn could have lowered the cost of accessing this technology for commercially insured patients.

[5] Evidence from Bronnenberg et al. (2012), which studies the evolution of consumer good brand preferences of migrants and finds that "approximately 60 percent of the gap in purchases between the origin and destination state closes immediately when a consumer moves" (p. 2473), is also suggestive of a supply-side effect driven by this spillover channel. More broadly, the mechanism is conceptually analogous to the idea of "preference externalities" in Waldfogel (2003), where product offerings in markets with differentiated goods depend on the aggregate distribution of consumer preferences, and to the management literature on slotting allowances, in which manufacturers compete to have their products stocked by retailers with limited shelf capacity (see, e.g., Shaffer, 1991).

separating intermediary choices from consumer preferences, since both are jointly determined and respond to common shocks. The advantage of our setting is that we can cleanly identify plausibly exogenous shifts in "local" product demand (i.e., that of commercially-insured patients) and trace their effect on unaffected consumers (Medicare patients), thereby providing direct evidence of intermediary-driven spillovers. The model we develop also contributes to the broader empirical industrial organization literature on vertical contracting and intermediaries, particularly ones that influence choice sets (Crawford et al., 2018; Ho and Lee, 2019; Starc and Swanson, 2021; Ho and Lee, 2023; Feng and Maini, 2024). Our model extends the standard setup to include two interacting intermediary layers.

Finally, our work also adds to the literature on the structure of prescription drug markets in two ways. First, we contribute a novel instrument that can be used to study the effect of coverage changes in the commercial segment. Existing work has relied on variation in out-of-pocket costs arising from Medicare benefit design (see, e.g., Einav et al., 2018; Chandra et al., 2024) and idiosyncratic assignment of enrollees to Medicare Part D plans (Brot-Goldberg et al., 2023; Geruso et al., 2023). Our instrument identifies exogenous variation in drug coverage at the local market level, which can then be used to study the effect of coverage choices on other outcomes.[6] Second, we uncover a potential additional barrier to biosimilar adoption. Past work on biosimilar uptake has highlighted the aggressive response of incumbent biologics (Feng et al., 2024), the role of patent uncertainty (Van de Wiele et al., 2021; Hemphill and Sampat, 2022), the relative safety and efficacy of biosimilar drugs (Zhai et al., 2019), and the financial incentives faced by physicians and facilities (Scott Morton, 2021; Bond et al., 2023). We show that the stocking behavior of outpatient facilities can also slow adoption.

1 Institutional Background and Data

1.1 Commercial Insurance and Facility Stocking

When prescribing drugs, physicians consider clinical concerns such as efficacy, safety, and drug interactions. In the case of drugs administered in hospital outpatient departments and physician offices, they also need to account for two types of restrictions: those set by the patient's insurance plan and those set by facility stocking.

The tool that most health plans and facilities use to determine these restrictions are formularies. *Prescription drug formularies* are tiered lists of medications covered by health plans, where each tier assigns different patient cost-sharing levels and may impose utilization management restrictions such as prior authorization or step therapy requirements. The structure of these formularies is driven by rebate negotiations where manufacturers provide financial concessions in

[6] For example, researchers could study how pharmaceutical firms adjust advertising or detailing in response to local coverage changes, or whether access to specific drugs (e.g., weight-loss drugs or Hepatitis C treatments) affects health outcomes or generates spillovers to other markets such as food consumption or utilization of other medical services. While we focus on formulary exclusions, our approach could be adapted to study how prior authorization and other access restrictions influence patient outcomes.

exchange for favorable tier placement or formulary inclusion. These rebates are often contingent on competitive positioning, with manufacturers offering additional concessions when direct competitors are placed in less favorable tiers or excluded entirely (Senate Finance Committee, 2021).[7] As a result, it is not uncommon for prescription drug formularies to exclude drugs with close substitutes.

The negotiation process varies depending on drug type and organizational structure within the healthcare system. Retail and specialty drugs covered under pharmacy benefits are typically negotiated by intermediaries called Pharmacy Benefit Managers (PBMs), while physician-administered drugs under medical benefits may be negotiated directly by insurers or by PBMs.[8] Once rebate structures are established, employers can either develop custom formularies tailored to their specific needs or adopt standardized "national formularies" which are off-the-shelf plans offered to any employer nationwide. National formularies have gained significant traction among self-insured employers due to their administrative simplicity, making them particularly influential in the commercial insurance market.

Facility formularies determine the drugs that facilities stock and make available to physicians.[9] There are at least two reasons a facility would restrict its formulary. First, limiting the number of drugs in stock has logistical advantages, such as easier inventory management and reduced storage space. Second, facilities can unlock additional discounts from manufacturers through percentage-based guarantees, which often imply that competing products must either be excluded or their use strongly discouraged (Dean et al., 2024). According to interviews we conducted with hospital pharmacy representatives, contracts with percentage-based rebates and exclusion clauses are common in cases where close therapeutic substitutes are available.[10] These contracts offer significant price reductions to facilities that meet a pre-specified usage threshold for a given brand within a substitution class.

The interaction between insurance coverage and facility formularies can potentially generate spillover effects across payers. When facilities are choosing which drugs to stock, they will prioritize those with broad coverage on patients' health plans, because if too many patients are on health plans that do not cover their preferred drug, they would risk missing the threshold that triggers the discount. Because Traditional Medicare (TM) covers all drugs, commercial insurance plans will have an outsized impact on the facility formulary choice. We note here that plans in

[7] For instance, in a 2017 lawsuit, Pfizer sued Johnson & Johnson, the manufacturer of brand-name drug Remicade, alleging that J&J induced payers to sign contracts explicitly prohibiting coverage of Pfizer's biosimilar version, Inflectra, in order to prevent Inflectra from gaining market share (Pfizer v. Johnson & Johnson, 2018).

[8] Physician administered drugs were historically not covered by PBMs as they fell under the medical versus prescription drug benefit. However, PBMs are increasingly incorporating these drugs into their formularies through various arrangements (Parekh et al., 2020; National Alliance of Healthcare Purchaser Coalition, 2025).

[9] Hospitals may have separate inpatient and outpatient formularies. Our analysis focuses on drugs prescribed primarily in outpatient settings and therefore on hospital outpatient formularies.

[10] These discussions ($n = 11$) revealed that percentage- or volume-based contracts are increasingly common, but only for medication classes with competition between brands and can be negotiated directly with manufacturers or through intermediaries, such as group purchasing organizations. While specific contract details are considered trade secrets, respondents frequently cited a brand market share of approximately 80% within a class as a typical minimum threshold to qualify for discounts, with discounts commonly exceeding 10%.

Medicare Advantage (a growing segment of the Medicare market that is administered by private plans with more latitude to implement utilization management tools) are also required to cover all physician-administered Part B drugs. Thus, in our setting, the only relevant coverage exclusions come from commercial insurance, and the presence or growth of MA plans should not affect the outcomes we measure or induce shifts in Part B coverage.[11]

1.2 Data sources, sample selection, and substitution classes

Data. Our analysis focuses on the formulary coverage and utilization of physician-administered prescription drugs. To measure coverage, we rely on commercial insurance plan data from MMIT (Managed Markets Insight & Technology, 2020). The MMIT data reports the coverage status of a large set of branded drugs at the plan-year level for the vast majority of U.S. commercial health plans.[12] The data also includes enrollment numbers for each plan, which we use as weights in calculating state-level coverage statistics, and whether the plan adopted a specific national formulary or a custom formulary.

To measure utilization, we use multiple datasets from the Centers for Medicaid and Medicare Services (CMS), which cover Traditional Medicare patients only (and not Medicare Advantage). In our main analysis, we use publicly-available state-level annual spending data at the drug level (Centers for Medicare and Medicaid Services, 2022). In subsequent analyses requiring claim-level data, we use the 20 percent random sample of Medicare Part B fee-for-service claims data (combining the carrier and outpatient files) from 2015 to 2019 (Centers for Medicare and Medicaid Services, 2019a). In these datasets, physician-administered drugs are identified through a Healthcare Common Procedure Coding System (HCPCS) code. HCPCS codes identify drugs based on substance (i.e., molecule), dose, and form (e.g., injection), but not product name. Hence, in a few instances, multiple products can be assigned to the same HCPCS code. To match product names in MMIT to HCPCS codes in CMS data, we use a crosswalk from HCPCS code to drug name contained in the CMS Part B Spending Dashboard (Centers for Medicare and Medicaid Services, 2022).[13]

We also collect data on Average Sales Price (ASP, the price used by Medicare to reimburse hospitals), and detailing payments to physicians from OpenPayments data to use as a control in some of our regressions.[14]

[11] A limited exception began in 2018 when CMS started allowing MA plans to use non-exclusion tools like step therapy to help promote price competition (see: https://www.cms.gov/newsroom/fact-sheets/medicare-advantage-prior-authorization-and-step-therapy-part-b-drugs, last accessed May 1st, 2026). Although the introduction of these restrictions partially overlaps with our sample period, our preferred coverage measure captures only whether a drug is covered by a plan or not, and therefore is unaffected by these restrictions.

[12] Although we cannot verify MMIT's coverage directly, a comparison of MMIT's enrollment numbers with enrollment estimates from the American Community Survey suggests that MMIT covers approximately 90% of all commercial beneficiaries (Kakani et al., 2024).

[13] Drug naming conventions can differ slightly across datasets, so we match Dashboard names to MMIT names using a fuzzy string match algorithm.

[14] ASP is defined as the average price of a drug net of all rebates to payers (including both hospitals and health plans), and is set at the national level using a statutory formula. Commercial prices and rebates affect ASP mechanically through their influence on national average price, variation that is absorbed by our brand-year fixed effects and cannot generate the cross-market variation that identifies our estimates. We control for ASP in specifications that don't include

Sample definition. Our primary analysis focuses on a broad group of physician-administered drugs. We define drugs at the level of a "brand," i.e., any product associated to a specific molecule sold by a specific firm under a trademarked name. To identify brands that are primarily administered in physician offices and hospital outpatient departments (covered by Medicare Part B) rather than through retail pharmacies (covered by Medicare Part D), we rely on publicly available spending data for Medicare Part B and D from the CMS Dashboard. Our initial sample includes brands for which over 95% of Medicare spending comes from Medicare Part B. We exclude brands with generic competitors because Medicare claims do not distinguish between different versions of the same molecule sold by different firms, which makes it impossible to calculate product market shares. We also exclude a small number of instances where multiple brands are assigned to the same Healthcare Common Procedure Coding System (HCPCS) code, and brands for which we are missing other information. We provide more details on our selection criteria in Appendix A.1.

Definition of substitution classes. Our main outcome of interest is a brand's market share, which requires us to identify substitution classes for each of our brands of interest. To do so, we use manually-collected information on Anatomical Therapeutical Chemical (ATC) classification codes.[15] The ATC is a hierarchical classification system created by the WHO that is commonly used to define drug markets in economic research.[16] An ATC code consists of a five-level alphanumeric code. We define substitution classes using the first four components of the ATC code (ATC-4), which are standardized sets of products in the same therapeutic, pharmacological, and chemical subgroups.

Although ATC-4s provide reasonable, internationally standardized, and clinically relevant substitution classes, they are imperfect. Many drugs have multiple indications, including some that may be off-label (see, e.g., Bradford et al., 2018). Furthermore, there is no assurance that ATC-4s match the market definitions used in percentage-based contracts. Generally speaking, any mismatch between our definition of substitution class and what is actually used in drug procurement contracts will weaken the relationship between formulary coverage and utilization.

To address this issue, we also run a secondary analysis focused on a small number of well-defined classes: biologic drugs ("biologics") with biosimilar competitors. Biologics are complex molecules derived from organic material, such as insulin or monoclonal antibodies. Due to their complexity, biologics cannot be exactly reproduced like traditional, small-molecule drugs. Instead, potential competitors can create "biosimilars": highly similar versions of an existing reference biologic without clinically meaningful differences in safety and efficacy (Food and Drug Administration, 2021). We include four physician-administered biologic molecules with biosimilar competitors that launched before 2019 in our secondary analysis: filgrastim (a bone-marrow stimulant, brand name Neupogen), infliximab (an immunosuppressant, brand name Remicade),

brand-year fixed effects.

[15] We thank Pragya Kakani for kindly sharing this data with us.

[16] See, e.g., Dubois and Lasio (2018); Kakani et al. (2020); Maini and Pammolli (2023). For more details on the ATC classification, see https://www.who.int/tools/atc-ddd-toolkit/atc-classification, retrieved March 16, 2024.

Table 1: Summary Statistics

Panel A. Class-Level Statistics			
		Core sample	Biosimilar sample
Number of substitution classes		25	4
Number of drugs per class	mean	4.16	3
	SD	(4.93)	(0.82)
Market share of leading product/reference biologic (Part B)	mean	0.76	0.86
	SD	(0.23)	(0.21)
Yearly sales (part B, million USD)	mean	1,612	1,677
	SD	(3,752)	(1,390)

Panel B. Drug-State-Year-Level Statistics				
			Reference Biologics	Biosimilars
Number of observations		16,758	964	1,188
Market Share	mean	0.22	0.85	0.12
	SD	(0.34)	(0.27)	(0.23)
Formulary exclusion rate	mean	0.12	0.06	0.26
	SD	(0.18)	(0.09)	(0.26)
Detailing per 1000 enrollees (USD)	mean	8.17	2.73	0.33
	SD	(96.97)	(5.99)	(2.99)

Notes: This table presents separate summary statistics for the core and biosimilar samples. Panel A provides statistics at the market levels (i.e., an ATC-4 or a molecule), aggregating across states and years 2015–2019. Panel B provides summary statistics at the drug-state-year level for the main variables in our core analysis.

pegfilgrastim (a bone marrow stimulant, brand name Neulasta), and epoetin alpha (a treatment for anemia, brand name Procrit).[17]

Biologics and biosimilars represent a useful case study for three reasons. First, the FDA standard for biosimilar approval requires that there are no clinically meaningful differences between biosimilar medications and their reference brand, ensuring that these products are close substitutes and ideal candidates for percentage-based and exclusive contracts. Second, biologics and biosimilars are identified through separate HCPCS codes in Medicare claims. Third, biosimilar adoption features prominently in the policy debate surrounding drug spending, as it is widely

[17] We chose 2019 as a cutoff because our Medicare data ends in 2019. Epoetin alfa is also marketed under the name Epogen for the treatment of end-stage renal disease (ESRD). As usage for ESRD is limited and sample size is insufficient, we focus only on non-ESRD usage and the reference brand Procrit.

viewed as a key mechanism for reducing spending on biologic drugs.[18] By focusing on these markets, our analysis can provide policy-relevant insights as well as a clean test of our mechanism in a setting where clinical equivalence is guaranteed by regulation. Our broader analysis of Part B drugs then demonstrates that this mechanism extends beyond biologics to shape contracting patterns more broadly.

Summary statistics. Our final sample includes 104 brands in 25 ATC-4 classes, and 12 brands in 4 classes defined by biologic molecules. Appendix A.2 provides a full list of classes and brands.

Table 1 reports some basic summary statistics on these brands. Panel A shows that there are, on average, 4 brands per ATC-4 class, with significant variation in sales across classes. Markets defined by biologic molecules are smaller and more uniform. Panel B, column (1) shows that the average market share at the drug-state-year level is 0.22 with a long right tail and that the average exclusion rate for brands is about 12 percent. Panel B, columns (2) and (3) provide similar statistics but focus on the biosimilar sample. Biosimilars have lower market shares and higher exclusion rates than reference biologic products, confirming the narrative that commercial markets are favoring reference biologics over biosimilars.

2 Impact of Commercial Formulary Coverage on Utilization in Part B

2.1 Methodology

Our strategy compares a brand's Medicare Part B market share within its substitution class (defined by the ATC-4 code in the core sample and the molecule in the biosimilar sample) to the fraction of commercial lives whose formulary excludes it. Our analysis is at the state-year level. We estimate the following regression equation:

$$\text{Brand Market Share}_{jkt} = \lambda_t + \gamma_{jk} + \beta \text{ Fraction of uncovered lives}_{jkt} + \varepsilon_{jkt} \tag{1}$$

where j denotes brand, k denotes state, and t denotes year. As controls, we include brand-state fixed effects γ_{jk} and either year, λ_t, or brand-year fixed effects λ_{jt}. Estimates using year fixed effects reflect both panel variation across drugs in national averages and cross-sectional variation across states within a drug-year, whereas brand-year fixed effects isolate only the latter source of variation. Although brand-year fixed effects can control for potentially unobserved market-wide preference shocks that may bias estimates, they might also absorb useful identifying variation (e.g., idiosyncratic negotiation outcomes unrelated to unobserved demand shocks). Therefore, we report results for both specifications. Our coefficient of interest, β, represents the effect on Medicare Part B market share of a unit change in the fraction of uncovered commercial lives.

[18] See for example, Marta Wosińska's interview on Tradeoff on January 26, 2023 (https://tradeoffs.org/2023/01/26/humira-biosimilar-drug-prices/, retrieved June 8, 2023).

The primary issue with interpreting the estimate from Equation 1 causally is that a correlation between commercial coverage and Part B utilization could reflect unobserved state-level preferences. Although our brand-state fixed effects γ_{jk} control for differences in the *average* preference for a given brand, unobserved *changes* in preferences over time may still simultaneously affect physician prescribing patterns and commercial formularies, many of which could be designed by health plans to match enrollee preferences.

To account for unobserved changes in state-level preferences over time, we propose a shift-share instrumental variable approach based on changes to drug coverage on *national* formularies—stock products available to all U.S. payers. Although some payers design custom formularies, many self-insured employers, who make up a significant fraction of the commercial insurance market, make use of national formularies. Between 2015 and 2019, 42 percent of commercial patients are enrolled in plans that adopted a national formulary. The biggest national formulary in each state covers about 10 percent of commercial lives, meaning that even a single coverage change could generate large effects.[19] We then use the differential exposure of states to national formularies and changes in coverage to specific brands by these formularies to generate variation unrelated to preferences.

We implement this shift-share style design by following the procedure laid out in Borusyak et al. (2022) (henceforth, BHJ).[20] BHJ show that the shift-share instrumental variable (SSIV) regression is algebraically equivalent to an instrumental variables regression that includes each shock as a separate instrument. Consistency and valid asymptotic inference with this estimator require three conditions. First, the shocks must be as-good-as-random with respect to the outcomes of exposed units to ensure consistency. Second, no single shock can dominate the variation in the instrument (many-shocks condition). Third, each observational unit must have limited exposure to any individual shock (share-granularity condition). The latter two conditions allow asymptotic inference.[21]

Our setting likely satisfies each of these three conditions. The primary concern is that demand shocks specific to states with high exposure to a given national formulary are driving that national formulary's coverage changes, which would threaten the validity of the random shock condition. However, most large national formularies operate in almost all states, and the largest exposure share to any single state averages about 0.16 across all formularies, making it unlikely that any one state would drive national-level changes.[22] The asymptotic inference conditions are also likely to hold in our setting: we observe 117 national formularies and each state is exposed to a small number of national formulary changes.

[19] For additional statistics on prevalence of national formularies, see Table A5 in Online Appendix B.

[20] Relative to the example application in BHJ, national formularies are analogous to industries and formulary coverage is analogous to industry-level import competition intensity.

[21] Note that none of these three conditions apply to exposure shares, which, under the BHJ approach, can be non-random.

[22] We confirm in robustness checks that large states are not driving our findings (Appendix B.5), and that plans with similar market shares across states are not more similar in coverage, suggesting that plans are not catering to state-level tastes (Appendix B.3).

Beyond these conditions, we address two practical considerations highlighted in BHJ. First, we have a panel data setting, so we use the recommended first-difference approach described in BHJ to allow for exposure shares changing over time. Second, not all plans in each state use a national formulary, so total exposure shares do not sum to one. We account for this by controlling for the total exposure shares.[23]

In addition, to avoid any issues with payers switching national plans in response to changes in national coverage, we use lagged exposure shares, as suggested in BHJ. Specifically, to ensure that our instrument is uncorrelated with preferences, we use *past* penetration (i.e., in the previous year) of national formularies to create a projected level of coverage measure. This approach still results in sufficient first-stage relevance, because, in any given year, 70–90 percent of plans keep the same national formulary as the year before.[24]

Formally, our modified specification is

$$\Delta\text{Brand Part B Market Share}_{jkt} = \beta \cdot \Delta\text{ Fraction of uncovered commercial lives}_{jkt} + \delta \cdot \sum_{f=1}^{N} \frac{lives_{kf,t-1}}{sumlives_{k,t-1}} + \lambda_t + \varepsilon_{jkt} \quad (2)$$

We instrument for Δ Fraction of uncovered commercial lives_{jkt} using the first difference instrument described in BHJ:

$$Z_{FD,jkt} = \sum_{f=1}^{N} \left(\frac{lives_{kf,t-1}}{sumlives_{k,t-1}} \times \Delta\mathbb{I}^{\text{excluded}}{}_{jft} \right)$$

where $lives_{kf,t-1}$ reflects the number of enrollees on national formulary f in state k and year $t-1$, $sumlives_{k,t-1}$is the total number of enrollees in state k in year $t-1$ (including enrollees in plans that do not use a national formulary), $\Delta\mathbb{I}^{\text{excluded}}{}_{jft}$ reflects whether national formulary f changed how it covered brand j between years $t-1$ and t. The second term in the estimating equation accounts for the fact that states have different total exposure to national formularies, following the procedure described in BHJ.

Variation in $Z_{FD,jkt}$ comes from two sources. The first is the time-varying market share of specific national formularies across states. The second source of variation is changes in the exclusion of brands across national formularies.[25] Overall, we identify 381 exclusion events and 727 inclusion events across all products in our core sample, and 30 exclusion events and 212 inclusion

[23] This is related to the broader idea of re-centering presented in Borusyak and Hull (2023). In our setting, the conversion from shocks (coverage changes) to independent variable (coverage level) is linear, so the "controlling for shares" approach is sufficient for dealing with non-random exposure to shocks.

[24] Similar approaches are commonly used in the economic literature to deal with the endogeneity of plan choice (see, e.g., Abaluck et al., 2018; Prager, 2020)

[25] As an illustrative example of how our instrument works, in 2019, two large PBMs (Express Scripts and CVS Caremark) added coverage of Udenyca (a biosimilar launched in 2018) to their national formularies, while a third (Optum Rx) did not. Because New Jersey had greater exposure to Optum Rx plans than Connecticut in 2018, both states started with the same coverage rate (44 percent) but diverged to 84 percent in New Jersey and 90 percent in Connecticut in 2019.

Table 2: Impact of Commercial Coverage on Brand Utilization—Core Sample

	(1)	(2)	(3)	(4)
Panel A. OLS Estimates				
Fraction Excluded	-0.091	-0.035	-0.014	-0.014
	(0.011)	(0.011)	(0.010)	(0.010)
Panel B. Shift-Share IV results				
Fraction Excluded	-0.174	-0.197	-0.093	-0.093
	(0.030)	(0.032)	(0.033)	(0.033)
Weak Identification *F*-stat. (Kleibergen–Paap)	418.4	405.7	129.9	129.9
Panel C. IV First Stage				
Uncovered on National Formulary	0.235	0.224	0.019	0.019
	(0.012)	(0.011)	(0.002)	(0.002)
R^2	0.330	0.322	0.0973	0.0973
Observations	16,114	14,363	16,114	16,114
Open Payments, ASP Controls	No	Yes	No	Yes
Fixed Effects	State-Brand, Year	State-Brand, Year	State-Brand, Brand-Year	State-Brand, Brand-Year

Notes: Estimates of Equation 2, using OLS and IV approaches. The data is from 2015 to 2019. The dependent variable is the brand-level market share for each brand in a given state-year. Reduced form standard errors are clustered at the state-brand level, and IV regressions are clustered at the brand-national formulary level. All brands are included in the calculation of market share, but only brands without generic competitors are included in regressions. Columns 2 and 4 include controls for detailing payments as reported on OpenPayments, and Column 2 also includes controls for ASP. Observations indicates number of brand-state-year observations. One clarification on fixed effects: our first difference IV specification implicitly accounts for state-brand fixed effects. We do not include an additional state-brand control in the first difference equation (but do include year or brand-year FEs).

events across products in our biosimilar sample.

Finally, we account for inference issues arising from correlated exposure shares. A key issue with our design is that if many states have similar exposure levels to the set of national formularies, a standard regression at the state-brand-year level would overstate the amount of variation in the underlying data. We follow Proposition 5 in BHJ and transform our panel into an equivalent national formulary-brand-year level dataset. We verify that the modified regression recovers an identical estimate of β as the simple state-level regression, and compute both heteroskedastic standard errors and standard errors clustered at the brand-national formulary level.[26]

[26] BHJ suggests clustering standard errors at the level of the underlying shock, which occurs at the national formulary level in our regressions. We cluster standard errors at the brand-national formulary level as opposed to clustering at the broader national formulary level across brands because there is likely negative correlation in shocks across brands at the national formulary level. For example, a new exclusion of a brand on a national formulary is likely to happen at the same time as the new inclusion of a competing brand. Consistent with this explanation, clustering standard errors at the broader national formulary level results in smaller estimated standard errors.

2.2 Results for the core sample

We report our results for the core sample of all Part B drugs in Table 2. Our first stage results in Panel C show a strong association between our instrument and coverage rates, confirming that our instrument has sufficient power. The OLS estimate suggests that a 10pp increase in the exclusion rate of a given brand is associated with a 0.9pp decline in market share among Medicare Part B beneficiaries (Column 1). When we rely solely on variation generated by national formularies through our IV approach, we find a similar effect (1.7pp). Robustness checks that control for marketing payments to physicians and ASP (Column 2) return almost identical coefficients. We find a significant but weaker effect (0.9pp decline) when including both state-brand and brand-year fixed effects (Columns 3 and 4). [27]

The instrumental variables estimates are economically significant. As an example, assume there are three competing drugs with equal market share, each earning 50 percent of revenue from Traditional Medicare. If a drug were excluded by a large PBM or insurer, reducing commercial coverage by 25pp, our most conservative estimate implies that it would lose about 2.3pp in TM market share, equivalent to a 3.5 percent loss in revenue *just from the spillover to Medicare Part B*.[28] The IV estimates are larger in magnitude than the OLS estimates, which could reflect commercial payers excluding popular drugs to reduce costs or differences in how much influence the marginal population (enrollees in plans that adopt national formularies) has on facility formularies.

2.3 Results for the biosimilar sample

Next, we run a secondary analysis focusing on biologic drugs with biosimilar competitors, with markets defined by the molecule of the reference biologic. As noted earlier, studying the biosimilars sample allows us to use well-defined substitution classes and focus on policy-relevant markets.

Results are shown in Table 3. Estimates are qualitatively consistent with our main results in 2, but the point estimates in specifications with only state-brand fixed effects are larger in magnitude, suggesting attenuation from misspecifications in the definition of substitution classes. In specifications controlling for both state-brand and brand-year fixed effects (Columns 3 and 4), the estimates are similar in magnitude to those in our main results (-0.09) but have substantially wider confidence intervals, as brand-year fixed effects absorb much of the longitudinal variation in biosimilar coverage, the main source of identifying variation in these markets.

[27] The smaller coefficient with brand-year fixed effects suggests that the panel variation across drugs is associated with larger impacts on market shares (i.e., drug A having a bigger change in nationwide coverage relative to drug B in a given year, driven by shifts in national formularies). While much of this variation may be driven by plausibly exogenous factors (e.g., idiosyncratic aspects of PBM negotiations), brand-year fixed effects provide insurance against confounding from unobserved national demand shocks that could correlate with the magnitude of coverage changes. We therefore treat the fixed-effects specification as our preferred, more conservative estimate.

[28] The loss in TM market share is calculated as 0.093 (the IV coefficient from column 4 of Table 2) times 0.25 (the hypothetical decrease in commercial coverage). The percentage revenue loss is calculated as 2.3pp (loss in TM market share) divided by 33pp (the original TM market share), multiplied by 0.5 (the fraction of revenue coming from TM).

Table 3: Impact of Commercial Coverage on Brand Utilization—Biosimilar Sample

	(1)	(2)	(3)	(4)
PANEL A. OLS ESTIMATES				
Fraction Excluded	-0.377 (0.023)	-0.522 (0.034)	0.004 (0.034)	0.005 (0.034)
Specification	OLS	OLS	OLS	OLS
PANEL B. SHIFT-SHARE IV RESULTS				
Fraction Excluded	-0.301 (0.033)	-0.291 (0.034)	-0.092 (0.078)	-0.093 (0.078)
Specification	2SLS	2SLS	2SLS	2SLS
Weak Identification *F*-stat. (Kleibergen–Paap)	526.1	558.1	39.39	39.28
PANEL C. IV FIRST STAGE				
Uncovered on National Formulary	0.262 (0.011)	0.260 (0.011)	0.015 (0.002)	0.015 (0.002)
Specification	OLS	OLS	OLS	OLS
R^2	0.417	0.415	0.149	0.149
Observations	2,142	1,809	2,142	2,142
Open Payments, ASP Controls	No	Yes	No	Yes
Fixed Effects	State-Brand, Year	State-Brand, Year	State-Brand, Brand-Year	State-Brand, Brand-Year

Notes: Estimates of Equation 2, using OLS and IV approaches. The data is at the state-year-brand level from 2015 to 2019. The dependent variable is the brand-level market share for each brand in a given state-year. Reduced form standard errors are clustered at the state-brand level, and IV regressions are clustered at the brand-national formulary level. Molecules are only included after their first biosimilar launch. Columns 2 and 4 include controls for detailing payments as reported on OpenPayments, and Column 2 also includes controls for ASP. Observations indicates number of brand-year-state observations.

2.4 Robustness checks

Appendix B.5 reports results from additional robustness checks, including: (i) traditional shift-share IV methodology, (ii) weighting by state population and market size, and (iii) excluding the four most populous states, whose preferences could be targeted in the design of national formularies (California, Texas, Florida, and New York). Estimates from these alternate specifications are qualitatively similar to our main analyses, although the magnitude of the effects varies across specifications, ranging from -0.081 to -0.196. We repeat the same robustness checks for the biosimilars sample in Appendix B.5. Most specifications confirm the overall results, though we find statistically insignificant results in all specifications that control for brand-year fixed effects.

Finally, we conduct an event-study analysis, which allows us to capture pre-trend dynamics and the timing of market share responses, at the cost of focusing on a much more limited group of events. We define coverage 'shocks' as the top 5% and bottom 5% of year-to-year state-level coverage changes for a given drug, measured as the change in the fraction of patients receiving

(or losing) access to the brand. We then compare states experiencing positive or negative shocks to states without large coverage changes for the same drug in the same year. The results, presented in Appendix B.7, show that positive shocks are followed by an immediate and significant increase in Medicare utilization, while negative shocks are followed by a smaller and insignificant decline.

3 The Role of Facilities and its Implications

3.1 Mechanism: the role of physicians versus facilities

Our conjecture is that facilities are the main drivers of the spillover effect from commercial coverage to Medicare utilization, through their choice of facility formularies. However, an alternative explanation could involve changes in physician behavior. For instance, physicians may develop a preference for a given drug if it is widely covered for their commercial patients, influencing their prescribing habits for all patients.

This distinction is relevant for two reasons. First, the "commonality of care" hypothesis is based on facility-level effects, which our previous results do not directly address. If facilities matter in our setting, this would motivate further exploration into other facility-level choices that could similarly shape health care choices of U.S. patients, such as staffing choices and capital investments. Second, changes in prescribing behavior attributable to facility practices suggest that facilities wield a direct influence on physician behavior and medication utilization, which matters for the modeling of treatment choice and competition between drug manufacturers.

To test whether physicians or facilities drive the relationship between commercial formulary coverage and Medicare Part B utilization, we compare the prescribing behavior of physicians that operate in different facilities in the same year using the methodology of physician "mover" studies (see, e.g., Finkelstein et al., 2021).[29] Formally, for each substitution class (either ATC-4 or molecule) c, we first identify the product with the highest market share, which we refer to as the "dominant" product. Then, for each physician i and facility f, we calculate the dominant product's average market share among claims by physician i when practicing in facility f. We denote this variable as $Y_{i,f,c}$. We also calculate the same market share for all other physicians prescribing within facility f,and denote it as $Y_{-i,f,c}$. Finally, we run the following regression for physicians i who prescribe across exactly two facilities, f_1 and f_2:

$$\left(Y_{i,f_1,c} - Y_{i,f_2,c}\right) = \alpha_c + \gamma_c \times \left(Y_{-i,f_1,c} - Y_{-i,f_2,c}\right) + \varepsilon_{i,c} \tag{3}$$

γ_c is specific to substitution class c, and measures how a physician changes their prescribing behavior to match the behavior of other physicians operating in that same facility. If variation in prescribing were entirely due to physician choices, γ_c would be zero. Conversely, if prescribing

[29] This approach is an indirect test, because we cannot measure physician and facility responses to a change in formulary coverage using our data. With more granular data, we could measure doctor-facility level responses to insurance coverage changes and decompose variation in that quantity. Instead, we decompose variation in the observational data.

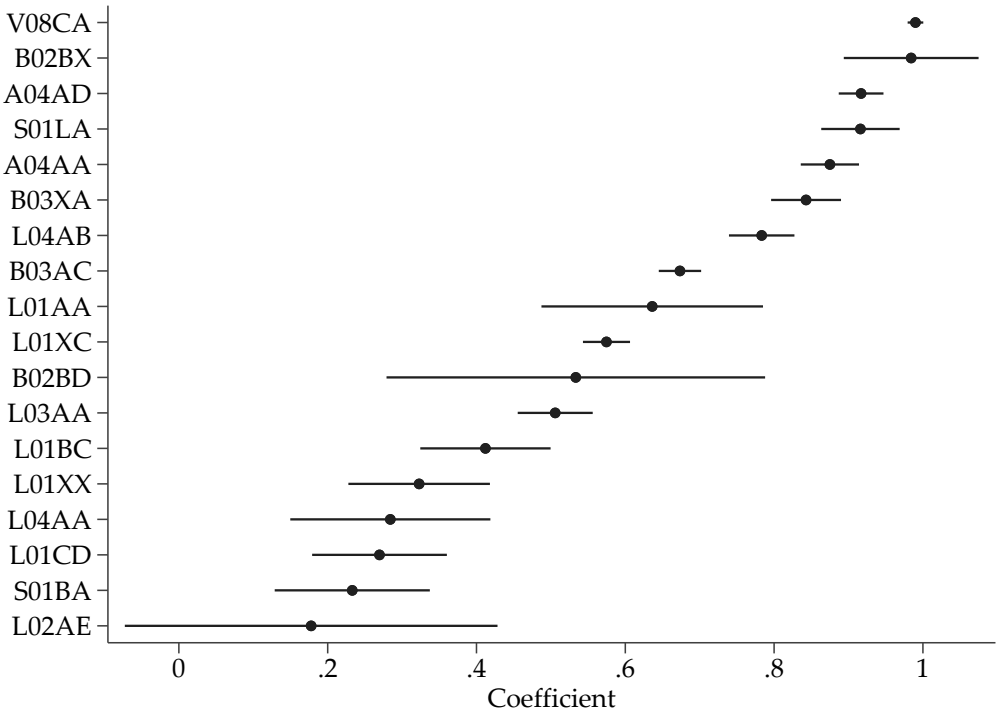


Figure 1: Change in Physician Brand Prescribing Across Facilities

Notes: Figure reports estimates based on Equation 3. The sample includes all physicians prescribing a given drug in two facilities, using data from the 20% Medicare Part B data in 2019.

decisions were entirely determined by facility-level factors, γ_c would be one.

Core Sample Results. Figure 1 presents regression slopes for each of our ATC-4 categories. Slopes range from 0.18 (95% confidence interval:—0.07 to 0.43) to 0.99 (95% confidence interval: 0.98 to 1.00). ATC-4s with lower slopes tend to have broader definitions (e.g., L01XX is "Other antineoplastic agents", a residual category for anticancer drugs) or include drugs with multiple indications (e.g., L02AE is "gonadotropin releasing hormone analogues," which includes drugs used to treat prostate cancer, but that also have on- and off-label indications in breast cancer, endometriosis, infertility, and precocious puberty). Conversely, higher slopes are associated with more narrowly defined categories, such as A04AD and A04AA (two classes of antinausea medication), or V08CA, (a group of "paramagnetic contrast media" used to enhance MRI images). Overall, our results show that facilities drive product choice in many ATC-4 categories. Importantly, our data includes facilities and physicians that prescribe a mix of products as well as those that almost exclusively prescribe a single brand. We include traditional "mover" graphs for each ATC-4 market (as well as markets in the biosimilar sample) in Appendix B.8, and report additional statistics about facility product-mix in Appendix B.11.

Biosimilar sample results. Results from the biosimilar sample mirror those in narrowly-defined core sample classes. Across all molecules, the facility appears to dominate prescribing decisions, with line slopes of 0.857 (95% confidence interval: 0.818–0.896) for filgrastim, 0.783 (95% CI: 0.739–0.827) for infliximab, 0.768 (95% CI: 0.719–0.817) for pegfilgrastim, and 0.843 (95% CI: 0.796–0.890) for epoetin alpha.

Robustness checks and alternative interpretations. We conduct several robustness checks to address potential alternative interpretations of our results. First, Appendix B.9 provides complementary evidence on the relative importance of facilities versus physicians using a variance decomposition approach adapted from Cooper et al. (2019) that allows us to include all physicians instead of restricting only to physicians practicing in multiple facilities. For each substitution class, we compare the explanatory power of drug-facility versus drug-physician fixed effects for predicting prescribing patterns. Even though physicians considerably outnumber facilities in our data, drug-facility fixed effects explain a comparable share of prescribing variation. The relative contribution of facilities implied by this approach is strongly correlated (0.63) with our main estimates across drug classes. Second, Appendix B.10 shows that effects are not driven by differences between hospital outpatient departments and physician offices, ruling out compositional differences across facility types driving our results. Third, Appendix B.11 examines within-facility brand concentration. In three-fourths of substitution classes, the majority of facilities administer a single brand rather than a mix of alternatives. By contrast, substitution classes in which facilities administer multiple brands within the same class are those where our mechanism analysis suggests a stronger physician role. If facilities were merely passive venues reflecting heterogeneous physician preferences, we would expect more within-facility brand mixing. Taken together, these patterns are consistent with our interpretation that facilities actively influence drug choice through formulary decisions.

A potential alternative explanation that we cannot easily rule out is that our findings reflect differences in patient composition across facilities rather than facility-level decision-making. Under this interpretation, physicians prescribe different drugs at different facilities simply because those facilities attract patients whose clinical profiles are a better match for those drugs. However, the explanation is at odds with our SSIV results. If prescribing were driven entirely by patient composition, we would not expect commercial coverage changes to affect utilization patterns among Part B beneficiaries, yet this is precisely what the first step of our analysis finds. Furthermore, any bias from patient composition differences should be largest in classes where products are less substitutable. Instead, our estimates are most pronounced in substitution classes with the most closely substitutable products (e.g., in our biosimilar sample), where patient characteristics should have little bearing on drug choice—the opposite of what this alternative explanation would predict.

3.2 Linking facility effects to coverage and utilization changes

The results presented in the previous section show that facilities influence prescription choices, but do not tie directly to the main analysis in Section 2. As a final empirical exercise, we show that the magnitude of the facility effect we just estimated (that is, γ_c from equation 3) is correlated with the magnitude of the effect of commercial coverage changes on Medicare utilization (that is, β from equation 2). To do so, we split our core sample in two groups based on whether γ_c is above or below the median estimate and re-estimate equation 2 separately for each group.[30]

We report our results in Table 4. Among drugs in substitution classes where facilities exert a weaker influence on prescription choices, the effect of commercial coverage on Part B utilization is approximately 40–60 percent smaller than our main estimate depending on the specification. Conversely, among drugs in substitution classes where facilities exert a stronger influence on prescription drug choices, the effect is 50–100 percent larger than our main estimate. These results support the idea that facilities are the primary mechanism driving the spillover effect from commercial coverage to Medicare utilization.

4 Implications for Spending, Competition, and Medicare

Our reduced-form analysis documents a spillover from commercial coverage to Medicare utilization, and traces it to a market organized around two interacting layers of intermediaries: insurers setting formularies and facilities making stocking decisions, with facilities responding to incentives from both Medicare and commercial payers. However, the reduced-form estimates cannot speak to how these two layers interact in equilibrium or how that interaction shapes prices and spending, because they reflect partial-equilibrium responses that hold competitors' decisions fixed.

To study equilibrium outcomes in this market, we develop a model of price competition between two manufacturers that incorporates both commercial formulary setting and exclusive facility stocking. The model is related to recent analyses of intermediaries that influence choice sets such as Ho and Lee, 2019; Starc and Swanson, 2021; Feng and Maini, 2024, but differs in ways that reflect both our empirical findings and institutional features of the market. Its key distinctive elements are: (i) two interacting layers of intermediaries that influence choice sets; (ii) facilities that respond to reimbursement incentives from both Medicare and commercial insurers; and (iii) reference pricing via Medicare reimbursement, which is based on the average net price across commercial payers.

We apply the model to markets with biosimilars, which have drawn significant attention from

[30] We split the sample in two—rather than perform class-specific regressions—because of sample size limitations. The analysis of facility effects relies on claims data, giving ample sample size for most drug classes. Conversely, our main analysis uses aggregated data at the state-brand-year level, which severely limits the sample size in most classes, and generates very noisy coefficients. We report estimates from a class-level analysis in Appendix A3 for completeness. We recover negative coefficients for 12 out of 18 classes. Two coefficients are exactly zero, likely due to limited shocks, and four are positive. The correlation between the class-level β and γ_c coefficients is very noisy.

Table 4: Impact of Commercial Coverage on Brand Utilization—Core Sample, split

	(1)	(2)	(3)	(4)
Panel A. Shift-Share IV results (Below Median)				
Fraction Excluded	-0.068	-0.082	-0.055	-0.055
	(0.011)	(0.013)	(0.022)	(0.022)
Specification	2SLS	2SLS	2SLS	2SLS
Weak Identification *F*-stat. (Kleibergen–Paap)	223.8	206.4	91.44	91.43
Observations	10,208	9,035	10,208	10,208
Panel B. Shift-Share IV results (Above Median)				
Fraction Excluded	-0.359	-0.366	-0.148	-0.148
	(0.067)	(0.069)	(0.098)	(0.098)
Specification	2SLS	2SLS	2SLS	2SLS
Weak Identification *F*-stat. (Kleibergen–Paap)	213.0	212.6	35.90	35.91
Observations	4,352	3,884	4,352	4,352
Open Payments, ASP Controls	No	Yes	No	Yes
Fixed Effects	state-brand, Year	state-brand, Year	state-brand, brand-year	state-brand, brand-year

Notes: Estimates of Equation 2 run on two different groups of core sample drugs separated based on whether the coefficient from Figure 1 is above or below the median. The data is at the state-year-brand level from 2015 to 2019. The dependent variable is the brand-level market share for each brand in a given state-year. Reduced form standard errors are clustered at the state-brand level, and IV regressions are clustered at the brand-national formulary level. All brands are included in the calculation of market share, but only brands without generic competitors are included in regressions. Columns 2 and 4 include controls for detailing payments as reported on OpenPayments, and Column 2 also includes controls for ASP. Observations indicates number of brand-year-state observations.

policymakers. A number of interventions have been targeted at reducing spending by spurring biosimilar adoption, including various Medicare payment reforms.[31] Within this application, we focus on equilibrium effects on prices and Medicare spending rather than on patient welfare. Because biosimilars are designed to be clinically interchangeable with their reference products, switching patients across them generates limited variation in match quality, so the welfare stakes from changes in choice sets are small relative to the spending stakes that motivate the policy debate.[32]

[31] Despite being cheaper, biosimilars are systematically less likely to be covered by commercial formularies than the reference products they compete with—a pattern the literature attributes to provider and patient preferences (Dalpoas et al., 2020; Feng et al., 2024), reimbursement policies that make it financially preferable for facilities to administer the reference product (Scott Morton, 2021; Bond et al., 2023), and rebate structures that can generate larger margins for PBMs on higher-priced reference products (Zhai et al., 2019).

[32] More generally, we find that exclusive stocking occurs most frequently in classes where products are close substitutes (e.g., biologics/biosimilars, MRI contrast dyes, anti-nausea medication), limiting its potential harm to patient welfare through restrictive choice sets.

4.1 A model of price-setting with formulary and stocking choices

We consider a vertical model where two single-product manufacturers compete to sell physician-administered drugs to patients through a large number of healthcare facilities. A single, representative insurer determines which products are covered. The model unfolds in three stages:

1. First, drug manufacturers simultaneously offer
   (a) per-unit rebates to a representative commercial insurer, and
   (b) per-unit exclusive stocking prices to facilities
2. After observing all prices, insurers and facilities simultaneously make formulary and stocking choices[33]
3. Finally, patients and physicians jointly choose drugs subject to the constraints imposed by formularies and stocking choices

We index drugs using $j \in \{1, 2\}$ (with 1 referring to the reference biologic and 2 referring to the biosimilar), patient-physician pairs using i, and facilities using h.

Throughout the model, we make several simplifying assumptions, motivated by tractability and data limitations. We discuss these assumptions and their implications in Appendix C.1.

Patient and Physician Drug Choice. We assume that patients are exogenously assigned to facilities. Patient and physician choice utility for drugs is given by a simple logit random utility function:

$$u_{ij} = \delta_j - \eta p_j + \varepsilon_{ij} \tag{4}$$

where δ_j is a product fixed effect, η is price sensitivity, and ε_{ij} is an extreme-value type 1 random shock. Also note that p_j here represents the price the patient sees, inclusive of any potential added fees. We assume that there is no outside option, and that patients can only choose options that are covered by their insurance and stocked at the facility where they receive treatment. As a result, this choice actively comes into place only for Medicare patients and commercial patients using open formularies that seek care at facilities that stock both products. Given this setup, we denote the probability of choosing product j for such patients as

$$s_j^{\text{open}}(p_1, p_2) = \frac{\exp\left(\frac{\delta_j - \eta p_j}{\sigma_\varepsilon}\right)}{\exp\left(\frac{\delta_1 - \eta p_1}{\sigma_\varepsilon}\right) + \exp\left(\frac{\delta_2 - \eta p_2}{\sigma_\varepsilon}\right)} \tag{5}$$

where σ_ε is the scale parameter for ε_{ij}.

[33] The assumption of simultaneity in both coverage and stocking choices as well as rebates and exclusive prices are made for convenience, and can be relaxed at the cost of significantly lengthening the computational time required to solve the model.

Facility stocking choice and insurer coverage choice. The representative insurer and all facilities observe all drug prices and rebates before choosing formularies $F \in \mathcal{F} = \{\{1\},\{2\},\{1,2\}\}$ and stocking $S \in \mathcal{S} = \{\{1\},\{2\},\{1,2\}\}$. We denote the vectors of formulary arrangement and stocking market shares as $\Lambda = (\lambda_1, \lambda_2, \lambda_{12})$ and $\Gamma = (\gamma_1, \gamma_2, \gamma_{12})$. λ_F and γ_S represent, respectively, the fraction of patients with a given formulary configuration, and the fraction of facilities with a given stocking arrangement.

Three drug prices matter for these choices: the non-exclusive (or list) price p_j^{list} (paid by facilities that stock both products), the exclusive price p_j^{excl} (paid by facility that stock a product exclusively) and the average sales price p_j^{ASP} (ASP, which is what determines both Medicare and commercial reimbursement to hospitals). Per-unit rebates r_j from manufacturers to commercial insurers also matter.

We assume p_j^{list} is exogenous. This assumption is motivated by convenience (as solving the model with multiple prices becomes very complicated), but also grounded in reality, because list prices respond to a variety of incentives that go beyond the problem we consider. p_j^{excl} is set by manufacturers to maximize profits, together with r_j. Finally, p_j^{ASP} is determined mechanically by the following formula:

$$p_j^{\text{ASP}} = \frac{1}{Q_j}\left(Q_j^{\text{excl}} \times p_j^{\text{excl}} + Q_j^{\text{non-excl}} \times p_j^{\text{list}} - Q_j^{\text{comm}} \times r_j\right) \tag{6}$$

where Q_j^{excl} is the total quantity of drug j administered in facilities with exclusive stocking, $Q_j^{\text{non-excl}}$ is the quantity administered in facilities with non-exclusive stocking, $Q_j = Q_j^{\text{excl}} + Q_j^{\text{non-excl}}$, and Q_j^{comm} is quantity administered to commercially insured patients.

Formulary coverage. We model a commercial representative insurer acting on behalf of multiple atomistic plans that can choose separate formulary arrangements.[34] We assume that the insurer's objective function consists of a per-unit brand-specific benefit for each drug administration, denoted as ζ_j, and a reimbursement cost paid to facilities where the administration takes place. We assume that commercial insurers reimburse facilities at a pre-determined, exogenously set markup over the Medicare price (i.e., p_j^{ASP}). This assumption is necessary because we do not have access to data on commercial prices, but is also conceptually consistent with empirical evidence on how commercial prices behave (Clemens and Gottlieb, 2017).

Formally, the insurer's choice function is

$$\Pi_{ins}\left(F|\, p^{\text{ASP}}, r, \Gamma\right) = \begin{cases} \zeta_{12} + \sum_j s_j\left(\Gamma, p_1, p_2\right) \times \left(-\beta p_j^{\text{ASP}} + \rho r_j\right) + \nu_{12} & \text{if } F = \{1,2\} \\ \zeta_j - \beta p_j^{\text{ASP}} + \rho r_j + \nu_j & \text{if } F = \{j\} \end{cases} \tag{7}$$

[34] The main reason for modeling a representative insurer rather than multiple plans is to restrict manufacturers to offering a single rebate (for tractability reasons). Our model is equivalent to one where rebates are posted and single plans pick a formulary arrangement.

where

$$s_j(\Gamma, p_1, p_2) = \gamma_j + \gamma_{12} \times s_j^{open}(p_1, p_2) \tag{8}$$

is the overall market share of each drug, which depends on the stocking choices of facilities, $\beta > 1$ denotes the (exogenous) commercial markup factor over ASP, ρ represent a weight on rebates, to account for the fact that insurers may strictly prefer rebate dollars, ζ_j and ζ_{12} capture the insurer's benefit from offering a given formulary arrangement, and ν_{12} and ν_j are type 1 extreme value shocks.[35] Equation 7 yields markets shares for each formulary arrangement that are expressed through the logit probability formulas:

$$\lambda_F = \frac{\exp\left(\frac{\delta_F(p^{ASP}, r, \Gamma)}{\sigma_\nu}\right)}{\sum_{F' \in \mathcal{F}} \exp\left(\frac{\delta_{F'}(p^{ASP}, r, \Gamma)}{\sigma_\nu}\right)} \tag{9}$$

where $\delta_F(p^{ASP}, r, \Gamma)$ is the utility of each formulary arrangement excluding the logit shock, and σ_ν is the scale parameter for ν_F.

Facility stocking. Facilities differ in their payer mix. In other words, different facilities receive a different mix of patients with Medicare vs. commercial coverage, and with different formularies within the commercial segment.

We denote the fraction of Medicare patients as $\alpha < 1$. The distribution of formularies within the commercial segment depends on the shares in equation 9. Formally, the share of patients using each formulary arrangement F is given by $(1-\alpha)\lambda_F$. Given parameter vector $\theta = (\alpha, \lambda_1, \lambda_2, \lambda_{12})$, the distribution of payer mix across facilities is given by $G(\theta)$.

Each facility is therefore defined as a vector $\theta_h = (\alpha_h, \lambda_{1h}, \lambda_{2h}, \lambda_{12h})$ denoting its patient payer-mix.[36] Facility profits are given by

$$\Pi_h(S|P, r, \theta_h, \Lambda) = \begin{cases} \sum_j \left(\psi_j p_j^{ASP} - p_j^{list}\right) s_j^{open}\left(\psi_j p_1^{ASP}, \psi_j p_2^{ASP}\right) \alpha_h + \\ \quad (1-\alpha_h) \times \sum_j \left(\beta p_j^{ASP} - p_j^{list}\right) s_j^{open}\left(\beta p_1^{ASP}, \beta p_2^{ASP}\right) \lambda_{12h} + \\ \quad\quad (1-\alpha_h) \times \left[\left(\beta p_1^{ASP} - p_1^{list}\right)\lambda_{1h} + \left(\beta p_2^{ASP} - p_2^{list}\right)\lambda_{2h}\right] \\ \quad\quad\quad + \kappa_{12} + \omega_{12h} & \text{if } S = \{1,2\} \\ \left(\psi_j p_j^{ASP} - p_j^{excl}\right)\alpha_h + \\ \quad (1-\alpha_h)\left[\left(\beta p_j^{ASP} - p_j^{excl}\right)\left(\lambda_{jh} + \lambda_{12h}\right)\right] + \kappa_j + \omega_{jh} & \text{if } S = \{j\} \end{cases} \tag{10}$$

[35] This flexible approach to modeling insurer preferences (using weights for various components of the objective function) mirrors our previous work (Feng and Maini, 2024), and reflects the difficulty of identifying where insurer and pharmacy benefit manager profits come from, which in turn implies that different sources of revenue may be valued differently.

[36] We assume that all facilities are homogeneous in size for simplicity of exposition. Allowing facilities to vary by size is conceptually straightforward.

where $P = \left\{p_j^{\text{list}}, p_j^{\text{excl}}, p_j^{\text{ASP}}\right\}_{j\in\{1,2\}}$ is the vector of prices, ψ_j incorporates the administration fee markup that Medicare pays to facilities for drug j, κ_{12} and κ_j denote market-wide preferences for specific stocking arrangements, and ω_{12h} and ω_{jh} are iid extreme value type 1 shocks.[37] The probability that hospital h chooses S is given by

$$s^{\text{stock}}(S|\theta_h) = \frac{\exp\left(\frac{\delta_S(P,r,\theta_h,\Lambda)}{\sigma_\omega}\right)}{\sum_{S'\in\mathcal{S}}\exp\left(\frac{\delta_{S'}(P,r,\theta_h,\Lambda)}{\sigma_\omega}\right)} \tag{11}$$

where $\delta_S(P, r, \theta_h, \Lambda)$ is the utility of each stocking arrangement excluding the logit shock, and σ_ω is the scale parameter for ω_{Sh}. From equation 11, we can derive the expression for the overall share of each stocking arrangement S as

$$\gamma_S = \int s_h(S|\theta)\,dG(\theta) \tag{12}$$

Finally, we can express the total quantity of product j sold in exclusive and non-exclusive facilities as

$$Q_j^{\text{excl}} = \int \left[\alpha_h + (1-\alpha_h)\left(\lambda_{jh} + \lambda_{12h}\right)\right] \times s^{\text{stock}}(\{j\}|\theta)\,dG(\theta) \tag{13}$$

i.e., the total number of patients for whom drug j is covered at the facilities that choose exclusive stocking of drug j, and

$$Q_j^{\text{non-excl}} = \int \left[\alpha_h s_j^{\text{open}}\left(\psi_j p_1^{\text{ASP}}, \psi_j p_2^{\text{ASP}}\right) + (1-\alpha_h)\left(\lambda_{jh} + \lambda_{12h} s_j^{\text{open}}\left(\beta p_1^{\text{ASP}}, \beta p_2^{\text{ASP}}\right)\right)\right] \times s^{\text{stock}}(\{1,2\}|\theta)\,dG(\theta) \tag{14}$$

i.e., the fraction of patients that choose j in facilities that choose to stock both drugs. These quantities can then be plugged into equation 6 to obtain the ASP.

For any given sets of prices and rebates $\left\{p_j^{\text{list}}, p_j^{\text{excl}}, r_j\right\}_{j\in\{1,2\}}$, an equilibrium for the formulary choice and facility stocking game consists of vectors of formulary market shares $\Lambda = (\lambda_1, \lambda_2, \lambda_{12})$ and stocking choices $\Gamma = (\gamma_1, \gamma_2, \gamma_{12})$, such that $\lambda_1 + \lambda_2 + \lambda_{12} = 1$, $\gamma_1 + \gamma_2 + \gamma_{12} = 1$, and that Λ, Γ and prices $\left\{p_j^{\text{ASP}}\right\}_{j\in\{1,2\}}$ satisfy the system given by equations 6, 9, and 12.

Manufacturer price-setting. Manufacturers set rebates r_j and exclusive prices p_j^{excl} to maximize profits. Firm j's problem is given by

$$\max_{r_j, p_j^{\text{excl}}} \int \left[p_j^{\text{list}}\left(s_j^{\text{open}}\left(\psi_j p_1^{\text{ASP}}, \psi_j p_2^{\text{ASP}}\right)\alpha_h\right) + \left(p_j^{\text{list}} - r_j\right)(1-\alpha_h)\left(\lambda_{jh} + s_j^{\text{open}}\left(\beta p_1^{\text{ASP}}, \beta p_2^{\text{ASP}}\right)\lambda_{12h}\right)\right] \times s_h(\{1,2\}|\theta) + \left[p_j^{\text{excl}}\alpha_h + \left(p_j^{\text{excl}} - r_j\right)(1-\alpha_h)\left(\lambda_{jh} + \lambda_{12h}\right)\right] \times s_h(\{j\}|\theta)\,dG(\theta) \tag{15}$$

[37] We use the same ψ_j in the facility objective function as in the patient demand function for notational simplicity. In reality, the two markups differ slightly. See Appendix C.2 for more details.

where firms receive a payment of p_j^{list} or p_j^{excl} from facilities for each unit supplied and then send a rebate r_j to commercial payers only. The expectation is taken over the distribution of payer-mix, conditional on predicted stocking and formulary coverage choices.[38]

The equilibrium concept of this stage is Nash-in-prices. In other words, an equilibrium of this stage is a set of exlusive prices and rebates such that, holding fixed their competitor's strategy and taking into account how the formulary and stocking game would adjust, neither firm wants to deviate by setting a different rebate or exclusive price.

4.2 Model Calibration

We calibrate this model using the Medicare spending, claims, and commercial formulary coverage data we used in the reduced-form analysis, supplemented with additional data on drug prices and sales from a variety of additional sources we discuss below.[39] We focus on the market for epoetin alfa in 2019. This is the only biosimilar market in our sample with only two competing products (Procrit, the reference biologic, and Retacrit, the biosimilar), which simplifies the computational complexity of our estimation routine. Epoetin alfa is widely used to treat anemia, including chemotherapy-induced anemia.[40]

The unknown parameters in the model are $(\delta_j, \eta, \sigma_\varepsilon, \beta, \rho, \alpha, \zeta_F, \sigma_\nu, \kappa_S, \sigma_\omega)$. The distribution of payer-mix at the facility level, $G(\theta)$ is also unknown.

Because we do not model the outside option, we normalize δ_2, ζ_2, and κ_2 to 0 and σ_ε to 1. Hence, δ_1, ζ_1, and κ_1 should be interpreted as the revealed preference "quality" premium of the reference biologic relative to its biosimilar competitor.[41] We set the commercial markup over ASP to $\beta = 1.75$, which we calibrate using data from the Transparency in Coverage (TiC) legislation, provided by Serif Health (Serif Health, 2025). We set ψ_j to follow Medicare rules in 2019 for biologics and biosimilars (see Appendix C.2 for more details). We calculate α (the share of Medicare sales) as the ratio of Procrit and Retacrit Medicare sales over total Procrit and Retacrit US sales, which we obtain from SSR Health (SSR Health, 2019).[42] We also extract the list prices of both drugs ($p_1^{list} = 165$ for Procrit and $p_2^{list} = 110$ for Retacrit) from SSR Health. Finally, we assume that $G(\theta)$ follows a scaled Dirichlet distribution (Aitchison, 1982), with common shape parameter equal to 2 and a

[38] Notice that in this problem, θ is a function of r_j and p_j^{excl}, because it depends on the insurer's optimal formulary coverage choice. As a result, there is no easy way to write the first-order conditions analytically.

[39] While parts of our routine follow structural estimation methods, our exercise is best understood as a calibration, because we lack sufficient data to rigorously estimate key elements of the model. Appendix C, which details our estimation procedure, also discusses what additional data would be required for a full estimation exercise, and carefully lays out the assumptions that our routine makes to compensate for insufficient data.

[40] Notably, Procrit is manufactured by Amgen but marketed by Janssen for non-ESRD indications under a longstanding licensing agreement, while Amgen retains ESRD rights under the Epogen brand. In ESRD, epoetin alfa is administered routinely within dialysis facilities under a bundled payment system. Given low use in our data and the distinct structure of the dialysis market, we focus on Procrit and Retacrit use in non-ESRD settings.

[41] σ_ε is essentially a scale parameter for η (the price sensitivity parameter in equation 4). The logit choice utilities we specify for the insurer's problem (equation 7) and the facility problem (equation 10) do not include a price-sensitivity parameter, so we cannot normalize σ_ν and σ_ω in the same way.

[42] Epoetin alfa can also be sold as a retail drug under Part D. Because we do not observe overall U.S. sales separately for pharmacies and outpatient facilities, our calculation implicitly assumes that both Medicare and commercial sales have the same split of pharmacy versus outpatient facilities.

scaling vector equal to the proportion of commercial patients with a given formulary (which in turn depends on the outcome of the rebate game).

We estimate the remaining parameters by matching model estimates for $s_j^{open}(p_1, p_2)$, p_j^{ASP}, p_j^{excl}, r_j, λ_F, and γ_S. In total, we have 11 moments and 9 parameters.[43]

Our sample moments are estimated from a variety of sources, which we summarize in Table 5. We use Medicare claims data to infer stocking choices of facilities. Specifically, we assume that facilities with over than 95% of administration for a given molecule (either the reference product or the biologic) are stocking it exclusively. To match $s_j^{open}(p_1^{ASP}, p_2^{ASP})$, we use Medicare claims in facilities that we classify as open-stocking. We calculate the fraction of commercial patients using a given formulary using MMIT data.[44]. Data on ASP is publicly reported by CMS (Centers for Medicare and Medicaid Services, 2019b). There is no publicly available data reporting p_j^{excl}, so we use prices from the federal supply schedule (FSS) reported by the Drug Pricing Database of the Department of Veterans Affairs as a proxy (Department of Veterans Affairs, 2019). Finally, we also add a moments matching rebates r_j, which we back out from p_j^{excl} and p_j^{ASP} using equation 6.

Intuitively, the facility stocking share moments, γ_1 and γ_2, and the commercial formulary share moments λ_1 and λ_2 are most helpful in pinning down the facility-side preference parameters (κ_1, κ_{12}) and the insurer preference parameters (ζ_1, ζ_{12}) respectively. The logit scale parameters σ_ω and σ_ν, the rebate sensitivity ρ, and the demand-side price sensitivity η govern responsiveness to price, and therefore are structurally identified by the price moments. Finally, δ_1 (the demand preference parameter), is pinned down by the Medicare market share in open-stocking facilities.

To match the moments, we solve the model numerically using an iterative algorithm, and then minimize the squared difference between the model predictions and the data (see Appendix C for more details on the estimation algorithm). Conceptually our approach mimics a simulated method of moments.

Discussion of Model Fit. Table 5 compares the sample moments to their model-predicted counterparts. Although the model matches some moments reasonably well, there are a few notable discrepancies between its predictions and the observed data. Most prominently, the model has difficulty accounting for the reference product's dominance in the stocking game, in spite of several built-in advantages stemming from patient, insurer, and hospital preferences (as reflected in the estimated parameters δ_1, ζ_1, and κ_1 reported in Table 6). The model also implies that the

[43] We have two moment conditions each for exclusive prices, ASP, and rebates, two moment conditions each for λ_F and γ_S, and one moment condition for the market share of product j. Because ASP is a linear function of p_j^{excl}, and r_j (equation 6), two of our 11 moments are effectively collinear, meaning that our procedure is a just-identified method of moments. However, we include all 11 moments because we find that doing so is useful in guiding the optimization routine away from corner solutions where one or both firms send zero rebates to insurers. To avoid putting too much weight on the collinear moments, we downweight both the p_j^{excl} and r_j moments by a factor of 0.5. We keep a higher weight on the ASP moment because ASP is reported directly, whereas we can only infer p_j^{excl} from VA data. Although the estimation is just-identified, it does not achieve a zero mean-squared error because the model is more complicated than a simple system of linear equations, and may contain misspecifications that prevent a perfect fit. See Appendix C.1 for a list of model simplifications and their potential effect on our results.

[44] Since we cannot observe epoetin patients specifically, we simply calculate the overall fraction of enrolled lives.

Table 5: Sample moments and sources

Sample moment	Calculation method	Source	Moment weight	Sample Value	Model Prediction
λ_1	Fraction of enrolled lives in commercial plans whose formulary covers F	MMIT	1	0.179	0.199
λ_2			1	0.030	0.090
λ_{12}			0	0.791	0.711
γ_1	Fraction of patients served by facilities stocking S	CMS 20% sample	1	0.653	0.543
γ_2			1	0.101	0.203
γ_{12}			0	0.247	0.254
$s_1^{open}\left(p_1^{ASP}, p_2^{ASP}\right)$	Fraction of reference product administrations among Medicare patients visiting open stocking facilities	CMS 20% sample	1	0.559	0.499
p_1^{ASP}	Weighted average of quarterly ASP	CMS ASP files	1	\$110.00	\$116.73
p_2^{ASP}			1	\$96.00	\$90.15
p_1^{excl}	Dosage-adjusted average NDC-level price for brand j (converted to HCPCS-level price to align with ASP)	FSS from VA Drug Pricing Database	0.5	\$119.70	\$128.69
p_2^{excl}			0.5	\$89.20	\$80.91
r_1	Inferred from p_j^{ASP}, p_j^{list}, and p_j^{excl} using sample estimates of $Q_j^{non\text{-}excl}$ and Q_j^{excl} (see Appendix C.3 for details)	Mix of FSS, ASP, and CMS 20% sample	0.5	\$26.45	\$28.03
r_2			0.5	\$6.96	\$6.88

Notes: *List of sample moments, calculation methods, data sources, weights, sample values, and model predictions. Moments with zero weight are perfectly collinear with other moments and are reported for completeness, but not directly used in the estimation. Moments with 0.5 weights are derived from other moments but included in the estimation because they help the estimation algorithm converge faster and more reliably.*

Table 6: Model parameters

Parameter	Description	Estimate
Estimated parameters		
δ_1	Demand reference product preference	8.795
η	Demand price sensitivity	0.331
κ_1	Facility reference product preference	36.835
κ_{12}	Facility open-stocking preference	32.450
σ_ω	Facility logit scale	34.280
ζ_1	Insurer reference product preference	22.953
ζ_{12}	Insurer open-formulary preference	130.057
ρ	Insurer rebate sensitivity	3.807
σ_ν	Insurer logit scale	72.121
Fixed parameters		
α	Medicare market share	0.33
β	Insurer hospital payment multiplier	1.75
ψ_1	Reference product Medicare admin fee	1.04304 (1.06 for patients)
ψ_2	Biosimilar Medicare admin fee	$0.984 \times p_2^{ASP} + 0.059 \times p_1^{ASP}$ ($p_2^{ASP} + 0.06 \times p_1^{ASP}$ for patients)
$G(\cdot)$	Dirichlet concentration	2

Notes: *List of estimated and fixed model parameters.*

biosimilar should be competing more aggressively on price.

The key reason for this discrepancy is that, in the model, marginal gains from price competition over exclusive stocking are high, and imply that the biosimilar should pursue a more aggressive pricing strategy. Conceptually, the estimated parameters are balancing two opposing forces. On the one hand, an even bigger advantage for the reference product in the stocking game would better match the observed stocking shares. On the other hand, such a change would dampen the reference product's incentive to compete on price, which would push the equilibrium further away from the observed ASP and exclusive prices, while forcing the biosimilar into an even lower equilibrium price.

Hence, the key puzzle that our model cannot resolve is why the reference product can achieve such a dominant position while keeping prices so high. A potential explanation is the presence of 340B hospitals. The 340B Drug Pricing Program requires manufacturers to provide eligible facilities with deeply discounted, federally mandated ceiling prices. Evidence shows that these hospitals overwhelmingly opt to administer reference biologics, largely because 340B rebates generate higher margins on reference products than on biosimilars, despite biosimilars having lower average prices (Bond et al., 2023). The presence of 340B facilities might justify why the data shows a much higher share of facilities stocking the reference product exclusively. Moreover, because these facilities are not making decisions based on ASP, our model may overestimate the gains of the biosimilar from competing on price.

Discussion of parameter estimates. Net of these discrepancies, our estimates confirm that the reference biologic has a large built-in advantage over its biosimilar competitor. Despite the two products being clinically equivalent, we find evidence of an advantage in patient demand ($\delta_1 = 8.795$), facility stocking ($\kappa_1 = 36.835$), and insurance formulary coverage ($\zeta_1 = 22.953$). We also estimate that both insurers and hospitals value having both drugs available. This advantage is relatively stronger for the insurer, which prefers an open formulary to a formulary that only covers the originator, than for the hospital, which is virtually indifferent between stocking the incumbent and stocking both drugs. This pattern is reasonable, given that hospitals face increasing logistical costs when stocking a large number of products, whereas insurers may actually face lower costs when using an open formulary if doing so streamlines claim administration.[45]

Finally, we note that the estimate for insurer rebate sensitivity is very high, suggesting that insurers value \$1 of rebates as \$3.807. This estimate must be interpreted within the context of the logit scale of the insurer function. Essentially, given the large variance, a high rebate weight is needed to induce manufacturers to compete on rebates for formulary coverage.[46] Weighing rebates more than spending may make sense in the context of the market we analyze, which is dominated (in terms of enrolled lives) by self-insured employers that contract with insurers and PBMs to provide health plan benefits to employees. PBMs have often been accused of hiding manufacturer rebates from employers, which could explain why rebates are so valued slightly more in our setup.[47]

4.3 Counterfactuals

Using the calibrated model, we assess the equilibrium effects of potential policy changes aimed at favoring biosimilar adoption. We consider three possible scenarios: i) an increase in the Medicare administrative fee for biosimilars, mimicking a proposal that was enacted in 2022; ii) a mandate for open commercial insurance formularies; and iii) a biosimilar commercial insurance coverage mandate. Table 7 summarizes our results.

ASP multiplier. In the first counterfactual, we increase the administrative fee paid by Medicare for biosimilar administrations from 6% to 8%, mimicking an actual policy enacted in 2022 as part of the Inflation Reduction Act.[48]

[45] Payers spend approximately \$6.0 billion annually administering drug utilization management (Howell et al., 2021), suggesting that an open formulary could generate substantial administrative savings by, for example, minimizing the need for back and forth over denied claims if a patient tries to submit a claim for an excluded drug.

[46] Rebates also lower ASP, which in turn lowers facilities' profit margins. For this reason, the fact that rebates to insurers even exist in the first place is moderately surprising. Yet, even though these rebates are not observed directly, plenty of suggestive evidence confirms their existence (see, e.g., allegations in Pfizer Inc. v. Johnson and Johnson and Janssen Biotech, 2018).

[47] See, e.g., the allegations in an ongoing Federal Trade Commission administrative complaint against PBMs (Federal Trade Commission, 2024), or a recent class action lawsuit (Plumbers' Welfare Fund, Local 130 U.A. v. Express Scripts, Inc. et al, 2026).

[48] The actual fee change was slightly lower because of the ongoing Medicare sequester. Our counterfactual takes the sequester into account. See Appendix C.2 for details.

Table 7: Counterfactual simulation results

	Baseline	**Biosimilar Adm. Fee 8%**	**Open Formularies**	**Biosimilar Coverage Mandate**
Panel A: Prices and Rebates				
p_1^{excl}	\$128.69	\$130.80	\$104.08	\$106.87
p_2^{excl}	\$80.91	\$79.88	\$60.37	\$70.15
r_1	\$28.03	\$28.75	\$0	\$0
r_2	\$6.88	\$5.58	\$0	\$0
p_1^{ASP}	\$116.73	\$117.98	\$105.40	\$110.47
p_2^{ASP}	\$90.15	\$90.18	\$76.68	\$84.60
Panel B: Payer Spending and Manufacturer Profits				
Total spending	\$774,485,422	\$783,487,610	\$813,567,818	\$764,165,157
Medicare spending	\$210,522,192	\$214,246,462	\$187,896,132	\$199,343,774
Commercial insurer spending	\$563,957,621	\$569,241,148	\$625,671,685	\$564,821,383
Reference product revenue	\$381,154,316	\$390,268,680	\$391,014,656	\$316,433,913
Biosimilar revenue	\$178,860,291	\$176,078,307	\$145,627,916	\$196,432,786
Facility revenue	\$214,470,815	\$217,140,623	\$276,925,246	\$251,292,848
Panel C: Market Structure (Stocking, Insurance, and Medicare Shares)				
Share of formularies covering the reference product exclusively	0.199	0.206	0	0
Share of formularies covering the biosimilar exclusively	0.090	0.086	0	0.190
Share of formularies covering both	0.711	0.708	1	0.810
Share of facilities stocking the reference product exclusively	0.543	0.538	0.647	0.555
Share of facilities stocking the biosimilar exclusively	0.203	0.210	0.227	0.263
Share of facilities with open stocking	0.254	0.252	0.126	0.182
Reference product Medicare market share	0.645	0.657	0.662	0.657
Biosimilar Medicare market share	0.355	0.343	0.338	0.343
Reference product commercial market share	0.614	0.614	0.661	0.493
Biosimilar commercial market share	0.386	0.386	0.339	0.507

Notes: *Counterfactual estimation. All counterfactuals are run on the same set of facility draws. Total spending numbers are scaled by total Medicare spending non-ESRD administrations of epoetin alfa ($210,522,192).*

According to our simulations, this change has virtually no effect on the equilibrium. The fee increase is too small to affect the stocking choice of virtually any facility, as that choice is instead dominated by the more profitable reimbursements received on the commercial side.

Our results align closely with the estimated impact of the 2022 policy change (Verrilli et al., 2024), but stand in contrast with its stated goals, which were to increase access to and utilization of biosimilars and promote competition in the marketplace.[49] We argue that this expectation would have been accurate in a market without stocking. Absent stocking, hospitals would have both products available and choose between them patient-by-patient, so every drug administration would be exposed to the policy. With stocking, by contrast, only administrations at facilities on the margin of their stocking decision are affected. We find that this margin responds weakly to Medicare reimbursement changes because stocking choices are dominated by commercial reimbursement, which are much higher than Medicare's.

Open insurance formulary mandate. We next simulate two types of mandates on commercial insurance coverage intended to level the playing field between biologics and biosimilars. The first mandates that commercial plans cover all drugs. Both manufacturers receive an advantage from this policy, but the biosimilar should gain more, as it is more often excluded than the reference product in the baseline equilibrium. More subtly, a crucial implication of this policy is that it means that all patients can receive either drug, and, as a result, the downside of exclusive stocking (being unable to serve some patients) disappears, making open stocking less attractive.

Our simulations suggest that these effects produce an environment that greatly benefits the incumbent because of its built-in advantage in the stocking game. Both the reference product and the biosimilar compete more aggressively on prices charged to facilities, to the point that even without rebates (which are mechanically zero since the insurer does not make an active choice), the ASPs of both products fall below their baseline levels. However, the loss of rebates hurts commercial payers and leads to higher spending in that segment. Conversely, Medicare spending falls, as Medicare gains from the increased price competition without suffering any loss from rebates. Ultimately, the big winners under this counterfactual policy are facilities, which benefit from lower acquisition prices and the ability to shift to exclusive stocking.

Biosimilar insurance coverage mandate. In the last counterfactual, we attempt to give a clearer advantage to the biosimilar by eliminating $\mathsf{F} = \{1\}$ (i.e., the insurance formulary that covers the reference product exclusively) from the insurer's choice set.

Even though the insurer still has an active choice to make (either cover the biosimilar exclusively, or cover both drugs), we find that in this scenario, both manufacturers opt not to provide any rebates. However, both firms compete more aggressively for facility stocking, so that—like in the open formulary counterfactual—ASPs are again lower than baseline (though higher than in the open formulary scenario).

[49] See https://www.cms.gov/files/document/biosimilar-faqs.pdf, accessed April 22, 2026.

In the stocking game, lower acquisition prices from both manufacturers mean that the fraction of facilities with exclusive stocking goes up, with most of these facilities opting to stock the biosimilar. The commercial mandate implies that the biosimilar is utilized at a much higher rate among commercial patients (50.7% market share vs. 38.6% at baseline). However, among Medicare patients, biosimilar utilization actually falls (from 35.5% at baseline to 34.3%) because the reference product competes more aggressively on price. In terms of spending, we find that this policy slightly reduces Medicare (and overall) spending. The policy also produces a large transfer from the manufacturer of the reference biologic to the biosimilar manufacturer and facilities.

4.4 Discussion

Three main insights emerge from the counterfactuals.

First, the presence of stocking limits Medicare's ability to influence utilization among its own beneficiaries. Because facilities make stocking decisions based on a combination of public and private incentives, Medicare reimbursement rules can only affect utilization (i) at open-stocking facilities or (ii) at facilities on the margin of their stocking decision. Empirically, only a small fraction of facilities adopt open stocking, and stocking decisions on the margin are dominated by commercial incentives, since private insurer reimbursements are substantially higher than Medicare's. Medicare's reach over its own beneficiaries' care is therefore limited.

Second, restrictive formulary coverage confers market power that manufacturers can leverage to raise facility acquisition prices. This is most apparent in the open-formulary counterfactual, where uniform coverage triggers aggressive price competition in the stocking game: open formularies make facilities more homogeneous from the manufacturer's perspective, placing more facilities on the margin between stocking one product versus the other. The logic echoes classic results on competition and horizontal differentiation dating to Hotelling (1929), and we expect it to extend to other markets in which producers compete for retailer shelf space. In such settings, demand shifters such as advertising play a role analogous to formulary coverage: by skewing demand among specific consumers, they generate spatial heterogeneity and soften the competitive stocking game (Bagwell, 2007). The broader principle is that when manufacturers can compete simultaneously for formulary placement and facility stocking, these instruments may act as strategic substitutes: restricting competition on one dimension intensifies it on the other. An important implication is that coverage mandates may have a more muted impact on spending for physician-administered drugs than for retail prescription drugs, where the absence of a stocking layer leaves formulary policy as the sole margin of competition.

Third, the persistent dominance of the reference biologic reflects the implied preferences of patients, physicians, facilities, and insurers rather than features of market structure. Across all three counterfactuals, the policies we consider fail to meaningfully erode the reference biologic's market share. The reason is that this dominance is built into these preferences, rather than arising from how competition is organized. Tilting the competitive balance toward the biosimilar in any single layer is therefore insufficient, and additional regulation tends to raise spending rather than

lower it.

5 Conclusion

This paper documents a novel channel through which the design of commercial insurance shapes the care received by Medicare Part B beneficiaries, and uses a model to quantify the equilibrium consequences of that channel for drug pricing and payer spending. Using a shift-share instrumental variables design, we establish that commercial insurance coverage decisions impact care received by publicly insured patients: a ten-percentage-point increase in commercial exclusion of a physician-administered drug translates into a 0.9–1.7 percentage-point decline in its Traditional Medicare market share. We provide evidence that this spillover operates through facility stocking rather than physician practice style, confirming a key prediction of the commonality-of-care hypothesis first developed by Glazer and McGuire (2002).

The model incorporates the features uncovered in the empirical analysis, and helps us quantify the equilibrium impact of policies designed to shift Part B utilization toward biosimilars. Three lessons emerge from the model calibration exercise. First, the reference biologic's advantage is hard to undo through any single policy lever because it is rooted in simultaneous built-in preferences on the patient, facility, and insurer side. Second, restrictive commercial insurance formularies soften competition in facility stocking. By generating heterogeneity across facilities in the downside of committing to a single product, exclusive formularies pull some facilities away from the stocking margin and allow manufacturers to sustain higher acquisition prices. Mandating open commercial coverage therefore lowers acquisition prices, but it also eliminates insurers' ability to extract rebates in exchange for coverage, yielding a net increase in total payer spending. Third, policies that target only the Medicare side of the market, such as the 2022 Inflation Reduction Act's upward adjustment of the biosimilar administrative fee, have muted equilibrium effects, because the commercial side of facility revenue dominates most facilities' stocking incentives at the margin. Shifting Medicare utilization toward biosimilars would therefore likely require changing the terms on which commercial payers compete for facility stocking, not just the fees Medicare pays.

More broadly, our results place facilities on the list of intermediaries whose decisions shape the care patients receive, alongside physicians and insurers. Previous work has shown that facility-level choices, including imaging-equipment adoption (Clemens and Gottlieb, 2014) and nursing-home staffing (Hackmann, 2019), respond to Medicare and Medicaid payment reforms; our findings show that commercial coverage design can move the same levers, generating spillovers onto publicly insured patients who share those facilities.

Finally, the spillover we document is likely to become increasingly relevant. Commercial plans have, over time, shifted toward more restrictive benefit designs (Kyle et al., 2026), while Medicare Part B has retained a relatively open coverage structure. To the extent that commercial coverage continues becoming more restrictive, the channel we identify suggests that commercial formulary

design will increasingly shape the care delivered to Medicare beneficiaries, meaning that policies aimed at managing Medicare Part B spending will need to account for the structure of the commercial market.

References

Abaluck, Jason, Jonathan Gruber, and Ashley Swanson, "Prescription drug use under Medicare Part D: A linear model of nonlinear budget sets," *Journal of Public Economics*, August 2018, *164*, 106–138.

Abito, Jose Miguel, Kathleen Hui, Yuval Salant, and Kosuke Uetake, "The Mechanisms of the Demand Spillover in the WIC Program," 2024.

Acquatella, Angelique, Keith Marzilli Ericson, and Amanda Starc, "Lagged-price reimbursement contracts: The impact of medicare Part B on pharmaceutical price growth," *Journal of Public Economics*, April 2026, *256*, 105595.

Aitchison, J., "The Statistical Analysis of Compositional Data," *Journal of the Royal Statistical Society: Series B (Methodological)*, January 1982, *44* (2), 139–160.

Anderson, Kelly E., G. Caleb Alexander, Carolyne Ma, Sydney M. Dy, and Aditi P. Sen, "Medicare Advantage coverage restrictions for the costliest physician-administered drugs," *The American Journal of Managed Care*, July 2022, *28* (7), e255–e262.

Bagwell, Kyle, "The Economic Analysis of Advertising," in "Handbook of Industrial Organization," Vol. 3, Elsevier, January 2007, pp. 1701–1844.

Baicker, Katherine, Michael E. Chernew, and Jacob A. Robbins, "The spillover effects of Medicare managed care: Medicare Advantage and hospital utilization," *Journal of Health Economics*, December 2013, *32* (6), 1289–1300.

Baker, Laurence C., "Managed care spillover effects," *Annual Review of Public Health*, 2003, *24*, 435–56.

Barnett, Michael L., Andrew Olenski, and Adam Sacarny, "Common Practice: Spillovers from Medicare on Private Health Care," *American Economic Journal: Economic Policy*, August 2023, *15* (3), 65–88.

Bond, Amelia M., Emma B. Dean, and Sunita M. Desai, "The Role of Financial Incentives in Biosimilar Uptake in Medicare: Evidence from the 340B Program," *Health Affairs.*, May 2023, *42* (5), 632–641.

Borusyak, Kirill and Peter Hull, "Nonrandom Exposure to Exogenous Shocks," *Econometrica*, 2023, *91* (6), 2155–2185. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA19367.

_ **, _ , and Xavier Jaravel**, "Quasi-Experimental Shift-Share Research Designs," *The Review of Economic Studies*, January 2022, *89* (1), 181–213.

Bradford, David W., John L. Turner, and Jonathan W. Williams, "Off-Label Use Of Pharmaceuticals: A Detection Controlled Estimation Approach," *The Journal of Industrial Economics*, 2018, *66* (4), 866–903.

Bronnenberg, Bart J., Jean-Pierre H. Dube, and Matthew Gentzkow, "The Evolution of Brand Preferences: Evidence from Consumer Migration," *American Economic Review*, May 2012, *102* (6), 2472–2508.

Brot-Goldberg, Zarek C., Samantha Burn, Timothy Layton, and Boris Vabson, "Rationing Medicine Through Bureaucracy: Authorization Restrictions in Medicare," January 2023.

Callaway, Brantly and Pedro H. C. Sant'Anna, "Difference-in-Differences with multiple time periods," *Journal of Econometrics*, December 2021, *225* (2), 200–230.

Centers for Medicare and Medicaid Services, "20% Research Identifiable Fee-for-Service Claims Database," 2019.

_ , "ASP Pricing Files," 2019.

_ , "CMS Part B Spending by Drug," 2022.

Chandra, Amitabh, Evan Flack, and Ziad Obermeyer, "The Health Costs of Cost Sharing," *The Quarterly Journal of Economics*, May 2024, *139* (4), 2037–2082.

Che, Catherine, "Do Middlemen Raise Drug Costs? Countervailing Market Power Meets Agency Frictions," 2025.

Clemens, Jeffrey and Joshua D. Gottlieb, "Do Physicians' Financial Incentives Affect Medical Treatment and Patient Health?," *American Economic Review*, April 2014, *104* (4), 1320–1349.

_ **and** _ , "In the Shadow of a Giant: Medicare's Influence on Private Physician Payments," *Journal of Political Economy*, February 2017, *125* (1), 1–39.

Conti, Rena M., Brigham Frandsen, Michael L. Powell, and James B. Rebitzer, "Common Agent or Double Agent? Pharmacy Benefit Managers in the Prescription Drug Market," June 2022.

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

Crawford, Gregory S., Robin S. Lee, Michael D. Whinston, and Ali Yurukoglu, "The Welfare Effects of Vertical Integration in Multichannel Television Markets," *Econometrica*, 2018, *86* (3), 891–954. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA14031.

Dalpoas, Stacy Elder, Mariana Socal, Celia Proctor, and Kenneth M Shermock, "Barriers to biosimilar utilization in the United States," *American Journal of Health-System Pharmacy*, November 2020, *77* (23), 2006–2014.

de Wiele, Victor L Van, Aaron S Kesselheim, and Ameet Sarpatwari, "Barriers to US biosimilar market growth: Lessons from biosimilar patent litigation," *Health Affairs*, August 2021, *40* (8), 1198–1205.

Dean, Emma Boswell, Phyllis Johnson, and Amelia M. Bond, "Physician, Practice, and Patient Characteristics Associated With Biosimilar Use in Medicare Recipients," *JAMA Network Open*, January 2021, *4* (1), e2034776–e2034776.

_ , Reekarl Pierre, Samuel Carter, and Amelia M Bond, "Role of supply chain intermediaries in steering hospital product choice: Group Purchasing Organizations and biosimilars," *Health Affairs Scholar*, May 2024, *2* (6), qxae067.

Department of Veterans Affairs, "Historical VA Pharmaceutical Prices," 2019.

Dranove, David and William D. White, "Medicaid-dependent hospitals and their patients: how have they fared?," *Health Services Research*, June 1998, *33* (2 Pt 1), 163–185.

Dubois, Pierre and Laura Lasio, "Identifying industry margins with price constraints: Structural estimation on pharmaceuticals," *American Economic Review*, 2018, *108* (12), 3685–3724.

Duggan, Mark and Fiona M. Scott Morton, "The Distortionary Effects of Government Procurement: Evidence from Medicaid Prescription Drug Purchasing," *Quarterly Journal of Economics*, 2006, *121* (1), 1–30.

Einav, Liran, Amy Finkelstein, and Maria Polyakova, "Private Provision of Social Insurance: Drug-Specific Price Elasticities and Cost Sharing in Medicare Part D," *American Economic Journal: Economic Policy*, August 2018, *10* (3), 122–153.

_ , _ , Yunan Ji, and Neale Mahoney, "Randomized trial shows healthcare payment reform has equal-sized spillover effects on patients not targeted by reform," *Proceedings of the National Academy of Sciences of the United States of America*, August 2020, *117* (32), 18939–18947.

Federal Trade Commission, "Caremark Rx, LLC; Zinc Health Services, LLC; Express Scripts, Inc.; Evernorth Health, Inc.; Medco Health Services, Inc.; Ascent Health Services LLC; OptumRx, Inc.; OptumRx Holdings, LLC; and Emisar Pharma Services LLC.," November 2024.

Feng, Josh and Luca Maini, "Demand Inertia and the Hidden Impact of Pharmacy Benefit Managers," *Management Science*, 2024, *70* (12), 8940–8961.

_ and _ , "Insurance Design and the Passthrough of Nominal Drug Prices," 2025.

_ , **Thomas Hwang, and Luca Maini**, "Profiting from Most-Favored Customer Procurement Rules: Evidence from Medicaid," *American Economic Journal: Economic Policy*, May 2023, *15* (2), 166–197.

_ , _ , **Jacob Klimek, and Luca Maini**, "Biosimilar Entry and the Pricing of Biologic Drugs," November 2024.

Finkelstein, Amy, "The Aggregate Effects of Health Insurance: Evidence from the Introduction of Medicare*," *The Quarterly Journal of Economics*, February 2007, *122* (1), 1–37.

_ , **Matthew Gentzkow, and Heidi Williams**, "Place-Based Drivers of Mortality: Evidence from Migration," *American Economic Review*, August 2021, *111* (8), 2697–2735.

Food and Drug Administration, "Biosimilar and Interchangeable Biologics: More Treatment Choices," Technical Report 2021.

Gandhi, Ashvin, "Picking Your Patients: Selective Admissions in the Nursing Home Industry," April 2023.

Gardenghi, Chiara, "Bundled Loyalty Discounts in Healthcare: Evidence from Physician-Administered Vaccines," 2024.

Garthwaite, Craig L., "The Doctor Might See You Now: The Supply Side Effects of Public Health Insurance Expansions," *American Economic Journal: Economic Policy*, April 2012, *4* (3), 190–215.

Geruso, Michael, Timothy J. Layton, and Jacob Wallace, "What Difference Does a Health Plan Make? Evidence from Random Plan Assignment in Medicaid," *American Economic Journal: Applied Economics*, July 2023, *15* (3), 341–79.

Glazer, Jacob and Thomas G. McGuire, "Multiple payers, commonality and free-riding in health care: Medicare and private payers," *Journal of Health Economics*, November 2002, *21* (6), 1049–1069.

Glied, Sherry and Joshua Graff Zivin, "How do doctors behave when some (but not all) of their patients are in managed care?," *Journal of Health Economics*, March 2002, *21* (2), 337–353.

Grabowski, David C., Jonathan Gruber, and Joseph J. Angelelli, "Nursing Home Quality as a Common Good," *The Review of Economics and Statistics*, November 2008, *90* (4), 754–764.

Hackmann, Martin B., "Incentivizing Better Quality of Care: The Role of Medicaid and Competition in the Nursing Home Industry," *American Economic Review*, May 2019, *109* (5), 1684–1716.

_ , **R. Vincent Pohl, and Nicolas R. Ziebarth**, "Patient versus Provider Incentives in Long-Term Care," *American Economic Journal: Applied Economics*, July 2024, *16* (3), 178–218.

Hemphill, C Scott and Bhaven N Sampat, "Fixing the FDA's orange book," *Health Affairs*, June 2022, *41* (6), 797–800.

Hermosilla, Manuel, Kelly Anderson, and Joseph Levy, "Biosimilar Turf Wars," 2025.

Ho, Kate and Robin Lee, "Contracting over Rebates: Formulary Design and Pharmaceutical Spending," 2023.

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

_ **and** _ , "Equilibrium Provider Networks: Bargaining and Exclusion in Health Care Markets," *American Economic Review*, February 2019, *109* (2), 473–522.

Hong, Dongzhe, Aaron S. Kesselheim, Ameet Sarpatwari, and Benjamin N. Rome, "Biosimilar Uptake In The US: Patient And Prescriber Factors," *Health Affairs*, August 2024, *43* (8), 1159–1164.

Hotelling, Harold, "Stability in Competition," *The Economic Journal*, March 1929, *39* (153), 41–57.

Howell, Scott, Perry T. Yin, and James C. Robinson, "Quantifying The Economic Burden Of Drug Utilization Management On Payers, Manufacturers, Physicians, And Patients," *Health Affairs*, August 2021, *40* (8), 1206–1214.

Huang, Rui and Jeffrey M. Perloff, "WIC Contract Spillover Effects," *Review of Industrial Organization*, February 2014, *44* (1), 49–71. Company: Springer Distributor: Springer Institution: Springer Label: Springer Number: 1.

Kakani, Pragya, Michael Anne Kyle, Amitabh Chandra, and Luca Maini, "Medicare Part D Protected-Class Policy Is Associated With Lower Drug Rebates," *Health Affairs*, October 2024, *43* (10), 1420–1427.

_ **, Michael Chernew, and Amitabh Chandra**, "Rebates in the Pharmaceutical Industry: Evidence from Medicines Sold in Retail Pharmacies in the U.S.," Technical Report w26846, National Bureau of Economic Research, Cambridge, MA March 2020.

Kyle, Michael Anne, Pragya Kakani, Samantha L Burn, and Luca Maini, "Trends in prescription drug coverage restrictions in Medicare, Medicaid, and commercial insurance plans, 2011-2019," *Health Affairs Scholar*, March 2026, *4* (3), qxag059.

Lewis, Matthew S. and Kevin E. Pflum, "Diagnosing Hospital System Bargaining Power in Managed Care Networks," *American Economic Journal: Economic Policy*, February 2015, *7* (1), 243–274.

Maini, Luca and Fabio Pammolli, "Reference Pricing as a Deterrent to Entry: Evidence from the European Pharmaceutical Market," *American Economic Journal: Microeconomics*, May 2023, *15* (2), 345–383.

Managed Markets Insight & Technology, "Formulary Insights," 2020.

Morton, Fiona Scott, "Paying for Biologic PADs in Medicare Part B: 1% Steps for Health Care Reform," Technical Report 2021.

National Alliance of Healthcare Purchaser Coalition, "Biosimilars Employer Toolkit," Technical Report October 2025.

Parekh, Natasha, Sarah Papa, Alek Drnach, Laura Spiegel, Yan Huang, Chronis Manolis, and Chester B. Good, "Effect of Carving in Pharmacy Benefits on Utilization and Costs," *Journal of Managed Care & Specialty Pharmacy*, October 2020, *26* (10), 1317–1324.

Pfizer v. Johnson & Johnson, "Pfizer Inc. v. Johnson & Johnson & Janssen Biotech, Inc.," August 2018.
Pfizer Inc. v. Johnson and Johnson and Janssen Biotech

***Pfizer Inc. v. Johnson and Johnson and Janssen Biotech**, August 2018.*
Plumbers' Welfare Fund, Local 130 U.A. v. Express Scripts, Inc. et al

Plumbers' Welfare Fund, Local 130 U.A. v. Express Scripts, Inc. et al, February 2026.

Prager, Elena, "Healthcare Demand under Simple Prices: Evidence from Tiered Hospital Networks," *American Economic Journal: Applied Economics*, October 2020, *12* (4), 196–223.

Richards, Michael R. and D. Sebastian Tello-Trillo, "Public Spillovers from Private Insurance Contracting: Physician Responses to Managed Care," *American Economic Journal: Economic Policy*, November 2019, *11* (4), 375–403.

Robinson, James C., Christopher Whaley, and Sanket S. Dhruva, "Hospital Prices for Physician-Administered Drugs for Patients with Private Insurance," *New England Journal of Medicine*, January 2024, *390* (4), 338–345. _eprint: https://www.nejm.org/doi/pdf/10.1056/NEJMsa2306609.

Senate Finance Committee, "Insulin: Examining the Factors Driving the Rising Cost of a Century Old Drug," Technical Report 2021.

Serif Health, "Our Data," September 2025.

Shaffer, Greg, "Slotting Allowances and Resale Price Maintenance: A Comparison of Facilitating Practices," *The RAND Journal of Economics*, 1991, *22* (1), 120–135.

Socal, Mariana P., Kelly E. Anderson, Aditi Sen, Ge Bai, and Gerard F. Anderson, "Biosimilar uptake in Medicare Part B varied across hospital outpatient departments and physician practices: the case of filgrastim," *Value in Health*, April 2020, *23* (4), 481–486.

SSR Health, "US Brand Rx Net Pricing Tool," 2019.

Starc, Amanda and Ashley Swanson, "Preferred Pharmacy Networks and Drug Costs," *American Economic Journal: Economic Policy*, August 2021, *13* (3), 406–446.

Verrilli, Caitlin, Max Vargas, and Matias Junghahn, "Evaluating the IRA's Impact on Medicare Part B Biosimilar Reimbursement," Technical Report, Certara 2024.

Waldfogel, Joel, "Preference Externalities: An Empirical Study of Who Benefits Whom in Differentiated-Product Markets," *The RAND Journal of Economics*, 2003, *34* (3), 557–568.

Whaley, Christopher, Nandita Radhakrishnan, Michael Richards, Kosali Simon, and Benjamin Chartock, "Understanding health care price variation: evidence from Transparency-in-Coverage data," *Health Affairs Scholar*, January 2025, *3* (2), qxaf011.

Zhai, Mike Z., Ameet Sarpatwari, and Aaron S. Kesselheim, "Why Are Biosimilars Not Living up to Their Promise in the US?," *AMA Journal of Ethics*, August 2019, *21* (8), 668–678.

Online Appendix

A Sample Selection and Additional Summary Statistics

A.1 Sample selection criteria

Main sample

We construct our core sample of Part B drugs using the following steps:

1. **Identify drugs launched before 2020 and primarily administered by physicians:** We use Medicare spending data to identify drugs whose Medicare sales occur predominantly in Medicare Part B (as opposed to Part D, which covers prescriptions from retail and mail pharmacies). We restrict our sample to drugs with at least 95 percent of sales through Part B.

   Sample size: 163 brands, 60 ATC-4s

2. **Drop products in ATC-4 classes where all products lose exclusivity before January 1st, 2015:** our analysis begins in 2015 and we don't have coverage data for generic products.

   Sample size: 144 brands, 46 ATC-4s

3. **Drop products with missing data:** we don't have coverage data for "ZOFRAN."

   Sample size: 143 brands, 46 ATC-4s

4. **Restrict to drugs with a unique HCPCS match, and to ATC-4s with multiple products:** Part B claims are identified by HCPCS code, so we need a unique product-to-HCPCS match to accurately assign market shares (the same product can match to multiple HCPCS codes, but the reverse cannot happen). ATC-4 classes cannot be used in the analysis if they have a single product.

   Sample size: 117 brands and 25 ATC-4s

5. **Drop brands after they lose exclusivity:** We use all brands to determine within ATC-4 market share. However, we drop brands that have lost exclusivity from our main regressions as we cannot separate out usage of branded and generic products in the Medicare data using HCPCS codes.

   Sample size: 104 brands in 25 ATC-4 categories

Mechanism-analysis restrictions

6. **Drop ATC codes with too few observations:** our mechanism analysis is restricted to physicians who operate in two facilities in the same year (about 10% of physicians). As a result, several ATC codes do not have sufficient sample size or variation to perform the mechanism analysis. We drop those ATC-4 codes.

Sample size: 88 brands in 18 ATC-4 categories

Biosimilar sample

Our biosimilar sample consists of the four physician-administered molecules where biosimilars launched prior to 2019: filgrastim, infliximab, pegfilgrastim, and epoetin alfa. We include molecules in our analysis only after launch of their first biosimilar product.

Notes on specific brands: Granix's (biosimilar version of filgrastim) manufacturer submitted its regulatory application to the FDA prior to the availability of the Biologics Price Competition and Innovation Act, thus is approved under the traditional biologic pathway. Inflectra launched in November 2016, but was not given a Medicare billing code until January 2017, thus our analysis considers it to have launched in 2017. Two additional biosimilar—Avsola (infliximab, launched 12/2019) and Ziextenzo (pegfilgrastim, launched 11/2019)—are not included in our sample as they launched at the end of our timeframe.

A.2 List of all products included in the analysis

Table A1 reports all products in the core sample and their substitution class. Table A3 reports all products in the biosimilar sample, their molecule, and indication.

Table A1: List of included drug classes and characteristics, core Sample

ATC-4	Description	List of included products
A04AA	Serotonin (5HT3) antagonists	Aloxi, Kyytil[a], Sustol
A04AD	Other antiemetics	Cinvanti, Emend
B01AC[b]	Platelet aggregation inhibitors excl. heparin.	Integrilin, Remodulin, Reopro, Tyvaso, Ventavis
B02BD	Blood coagulation factors	Adynovate, Alprolix, Benefix, Eloctate, Feiba, Hemofil, Jivi, Kogenate/FS/Kovaltry, Novoeight, Novoseven/RT, Xyntha/Refacto
B02BX	Other systemic hemostatics	Hemlibra, Nplate
B03AC	Iron, parenteral preparations	Feraheme, Ferrlecit[a], Infed, Injectafer, Venofer
B03XA	Other antianemic preparations	Procrit, Retacrit
G03BA[b]	3-oxoandrosten (4) derivatives	Aveed, Delatestryl
L01AA	Nitrogen mustard analogues	Belrapzo, Bendeka, Treanda
L01BC	Pyrimidine analogues	Dacogen, Gemzar[a], Vidaza[a], Xeloda[a]
L01CD	Taxanes	Abraxane, Jevtana, Taxotere[a]
L01DB[b]	Anthracyclines and related substances	Doxil[a],Valstar
L01XC	Monoclonal antibodies	Adcetris, Arzerra, Avastin, Blincyto, Cyramza, Darzalex, Empliciti, Erbitux, Gazyva, Herceptin, Imfinzi, Kadcyla, Keytruda, Libtayo, Mvasi, Opdivo, Perjeta, Polivy, Portrazza, Rituxan, Tecentriq, Truxima, Vectibix, Yervoy
L01XX	Other antineoplastic agents	Camptosar[a], Elzonris, Erwinaze, Halaven, Hycamtin[a], Istodax, Kyprolis, Trisenox, Velcade, Zaltrap
L01XY[b]	Combinations of antineoplastic agents	Kanjinti, Vyxeos
L02AE	Gonadotropin releasing hormone analogues	Lupron, Vantas, Zoladex
L03AA	Colony stimulating factors	Fulphila, Granix, Neulasta, Neupogen, Nivestym, Udenyca, Zarxio
L03AX[b]	Other immunostimulants	Mozobil, Provenge
L04AA	Selective immunosuppressants	Cellcept[a], Lemtrada, Myfortic[a], Thymoglobulin, Tysabri
L04AB	Tumor necrosis factor alpha (TNF-α) inhibitors	Inflectra, Remicade, Renflexis
R03AC[b]	Selective β-2-adrenoreceptor agonists	Accuneb[a], Brovana, Perforomist
S01BA	Corticosteroids, plain	Ozurdex, Retisert, Yutiq
S01LA	Antineovascularisation agents	Beovu, Eylea, Lucentis, Macugen, Visudyne
V03AF[b]	Detoxifying agents for antineoplastic treatment	Elitek, Ethyol[a], Totect
V08CA	Paramagnetic contrast media	Gadavist, Magnevist

[a] Indicates product is included in market share calculations but excluded from the analysis.
[b] Indicates ATC-4 is included in the regression analysis but not in the mechanism analysis.

Table A3: List of included drug classes and characteristics, Biosimilar Sample

Molecule	Indication	List of included products and years
filgrastim	treatment of low white blood cell count (short-acting)	Neupogen (2015-2019), Granix (2015-2019), Zarxio (2015-2019), Nivestym (2018-2019)
infliximab	treatment of autoimmune diseases	Remicade (2017-2019), Inflectra (2017-2019), Renflexis (2017-2019)
pegfilgrastim	treatment of low white blood cell count (long-acting)	Neulasta (2018-2019), Fulphila (2018-2019), Udenyca (2018-2019)
epoetin alfa	treatment of low red blood cell count	Procrit (2018-2019), Retacrit (2018-2019)

B Validation of the Instrumental Variable

In this section, we present additional evidence supporting the validity of our instrumental variable approach.

B.1 Variation in national formulary uptake and coverage

Table A5 provides statistics related to our national formulary instrument. Panel A provides statistics at the formulary-by-year level, while Panel B provides statistics at the state-year level. We highlight three results. First, the data shows that national formularies operate across a large number of states, with no individual state having a dominant position. Most of the large national formularies operate in almost all states (a weighted average of 48 states).[50] The largest exposure to any state (typically California or Texas) averages about 0.16. Second, the average national formulary makes at least a handful of coverage changes from year to year. We find that, on average, 2.61 drugs that were previously excluded gain coverage, whereas 1.01 that were previously covered are excluded. Third, across all states and years, 42 percent of commercial patients are enrolled in plans that adopt a national formulary. Over 50 national formularies operate in the average state, but the distribution is very skewed. The most popular national formulary covers about 10 percent of commercial lives, meaning that a coverage change can generate large effects at the state level.

Figure A1 shows the overall penetration rate of national formularies across states in 2019. With two exceptions (North Dakota, with a penetration rate of 22 percent, and Hawaii, with a penetration rate of 27.1 percent), national formularies cover at least 30 percent of commercial insurance enrollees in all states. The average uptake is 45.5 percent.

[50] Several small national formularies operate across a small number of states. However, these have very low enrollment numbers, so they are given almost no weight in our instrumental variable approach.

Table A5: Characteristics of National Formularies

Panel A. Formulary-Year Level Statistics		
	Mean	St. Dev.
Number of States with Any Enrollees	48.01	8.41
Average Share (Operating States)	0.03	0.09
Share from Biggest State	0.16	0.12
Number of Drugs that Change Coverage	3.62	5.53
Number of Drugs that Gain Coverage	2.61	3.82
Number of Drugs that Lose Coverage	1.01	3.59
Number of observations	410	
Panel B. State-Year Level Statistics		
Share of commercial enrollees using a national formulary in previous year	0.42	0.12
Most popular National Formulary Share	0.10	0.06
Number of National Formularies	54.15	17.41
Number of observations	255	

Notes: This table provides statistics related to the national formulary instrument. In panel A, we provide statistics at the formulary-by-year level. We calculate the number of states in which each formulary operates (in a given year), the share of their business coming from all states in which they operate, and the share of their business coming from the largest state in which they operate. At the formulary-year level, we also provide statistics on the total number of drug coverage changes and how many are inclusions versus exclusions. All statistics are weighted by enrollment. In panel B, we provide statistics at the state-by-year level. We calculate the fraction of commercially insured patients on a plan that uses a national formulary ("Total Exposure"), the share for the biggest national formulary in the state ("Biggest National Formulary Share"), and the number of distinct national formularies used across all plans ("Number of National Formularies").

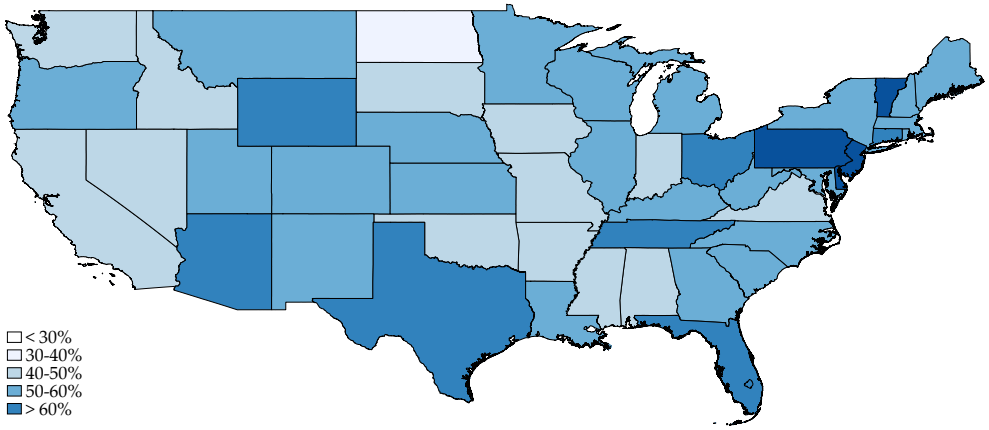


Figure A1: Uptake of national formularies among commercial insurance enrollees, 2019

Notes: Figure reports the fraction of commercial enrollees covered by a national formulary in each state, using 2019 data. Alaska (37.8 percent uptake) and Hawaii (27.1 percent uptake) are not pictured.

B.2 Inertia in Commercial Insurance Partnerships with PBMs and National Formularies

Next, we establish that plans stick with the same national formulary at high rates from year-to-year. This allows us to use previous year's exposure in constructing our instrument and still maintain relevance.

Table A6 reports the likelihood that a given plan that is using a national formulary remains in the data, keeps that same PBM, and keeps the same national formulary. From 2014 through 2019, the large majority of plans on a national formulary remained on the same national formulary the following year. The 2015 to 2016 rate is affected by the Optum-Catamaran merger, which consolidated two large PBMs and shuffled the identity of several large national formularies. Overall, the numbers show that the vast majority of plans who exist in both years stick with the same national formulary.

Table A6: Percentage of Commercial Insurers that Contract with the Same PBM and National Formulary as the Previous Year

Year Transition	Plan Remains	Same PBM	Same National Formulary
2014 to 2015	67.92%	65.78%	57.64%
2015 to 2016	59.94%	57.12%	27.80%
2016 to 2017	96.22%	95.29%	91.95%
2017 to 2018	87.38%	87.38%	78.79%
2018 to 2019	95.76%	94.82%	74.06%

Notes: For a given year, we keep all plans that use a national formulary. We then calculate the probability that the same plan remains in the MMIT data for the next year ("Plan Remains"). We then calculate the probability that the plan remains and uses the same PBM in the next year ("Same PBM"). Finally, we calculate the probability that the plan remains AND uses the same formulary ("Same National Formulary"). All probabilities are weighted by plan enrollment. The 2015 to 2016 rate is affected by the Optum-Catamaran merger, which consolidated two large PBMs and shuffled the identity of several large national formularies.

B.3 Correlation between coverage and state-level market shares

Finally, we check whether national formularies that operate in the same states make similar coverage choices, in order to assess whether national formularies cater to the tastes of their specific patient mix. To do so, we calculate two measures of similarity between two national formularies f and f': (i) a measure of state-share similarity, which we calculate as the Euclidian distance between the vector of each plan's shares of patients in each state; and (ii) a measure of coverage similarity, which we calculate as the number of drugs covered by both formularies divided by the number of drugs covered by f'.

Figure A2 shows that there is only a weak correlation between these two measures. In other

words, pairs of plans with very similar customers by geography make quite different coverage decisions across the 104 branded drugs in our core sample (about 14 different coverage decisions). Pairs of plans with very different customers have a similar number of differences (about 17.5). This provides further evidence that coverage decisions by national plans are not strongly determined by their geographic exposure.

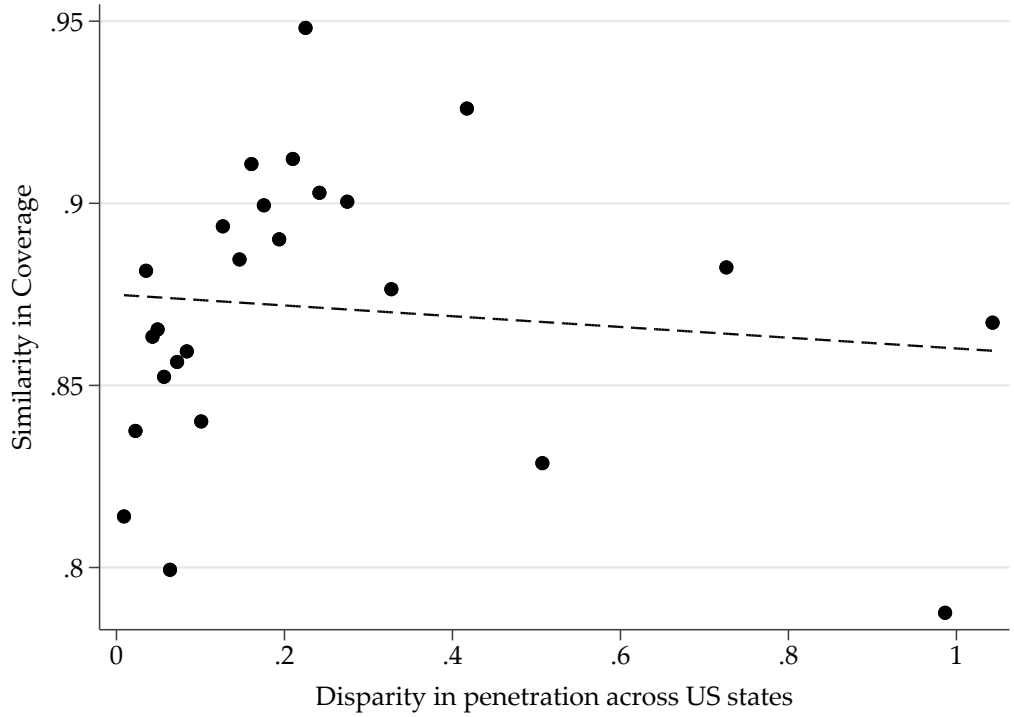


Figure A2: Correlation of plan design similarity and enrollee similarity across pairs of national formularies

Notes: For all pairs of national formularies (in the 2015 to 2019 period), we calculate the similarity of their exposure to different states in terms of enrollees (based on Euclidean distance) and the fraction of a given formulary's covered drugs that are also covered by the other formulary in the pair ("Coverage Overlap Fraction"). We then provide a binscatter of the correlation between the two quantities. Note that the Euclidean distance measure is bounded by the square root of two.

B.4 Placebo test

To validate our instrument further, we conduct a placebo test by examining whether changes in commercial insurance coverage affected Medicare market share in the year before the actual coverage change. We limit this test to the main sample because the biosimilar sample is too small to provide precise estimates with a one-year lookback.

When we include brand-year fixed effects, we find no effect on Medicare market share prior to the policy change and observe shifts in market share only at the time of the coverage change. In specifications with only state-brand fixed effects (columns 1 and 2), we find the placebo coef-

Table A7: Impact of Commercial Coverage on Brand Utilization in the Prior Year

	(1)	(2)	(3)	(4)
PANEL A. OLS ESTIMATES				
Fraction Excluded	-0.059 (0.012)	-0.068 (0.014)	-0.008 (0.009)	-0.008 (0.009)
PANEL B. SHIFT-SHARE IV RESULTS				
Fraction Excluded	-0.061 (0.027)	-0.062 (0.027)	0.005 (0.054)	0.005 (0.054)
Weak Identification *F*-stat. (Kleibergen–Paap)	110.1	109.4	111.8	111.8
PANEL C. IV FIRST STAGE				
Uncovered on National Formulary	0.150 (0.014)	0.150 (0.014)	0.024 (0.002)	0.024 (0.002)
R^2	0.219	0.219	0.127	0.127
Observations	12,370	11,746	12,364	12,364
Open Payments, ASP Controls	No	Yes	No	Yes
Fixed Effects	state-brand, Year	state-brand, Year	state-brand, brand-year	state-brand, brand-year

Notes: Estimates of Equation 2 where our dependent variable is the prior year's brand-level market share for each brand in a given state-year. The data is at the state-year-brand level from 2015 to 2019. Standard errors are clustered at the brand-national formulary level. All brands are included in the calculation of market share, but only brands without generic competitors are included in regressions. Columns 2 and 4 include controls for detailing payments as reported on OpenPayments, and Column 2 also includes controls for ASP. Observations indicates number of brand-year-state observations.

ficient is small (around -0.06) but significant, indicating some residual endogeneity when brand-year shocks are not absorbed.

B.5 Alternative specifications for regression 2

We present results from a set of alternative specifications for regression 2 (Tables A8 and A9). We present results using the traditional SSIV 2SLS approach, using population size and market size weights, and excluding large states from the analysis (out of a concern that national formularies may target their preferences). These states are California, Texas, Florida, and New York. The results largely confirm our main analysis.

B.6 Class-specific regressions

We report results from class-specific regressions (equation 2). Results at the class level are very noisy because several classes contain a small number of drugs. As a result, comparing these coefficients to the γ coefficients from the physician mover analysis is not a sharp test of our theory. Once we account for uncertainty in the CIs for both these coefficients and γ, we find small and

Table A8: Robustness Checks on Regression 2, core Sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. OLS Estimates								
Fraction Excluded	-0.035	-0.014	-0.055	0.012	-0.042	-0.004	-0.033	-0.019
	(0.011)	(0.010)	(0.011)	(0.008)	(0.010)	(0.008)	(0.012)	(0.009)
Panel B. Shift-Share IV results								
Fraction Excluded	-0.145	-0.134	-0.196	-0.086	-0.196	-0.081	-0.196	-0.092
	(0.023)	(0.112)	(0.034)	(0.044)	(0.033)	(0.026)	(0.032)	(0.033)
First-Stage *F*-stat.	998.9	60.02	381.5	107.2	367.8	133.5	405.6	123.8
Panel C. IV First Stage								
Uncovered on	1.181	0.267	0.221	0.010	0.221	0.017	0.225	0.020
National Formulary	(0.037)	(0.034)	(0.011)	(0.001)	(0.012)	(0.001)	(0.011)	(0.002)
R^2	0.682	0.868	0.310	0.0522	0.321	0.0955	0.325	0.0968
Observations	14,363	16,104	14,363	14,353	14,363	14,353	12,999	14,586
Controls	Detailing, ASP	Detailing	Detailing, ASP	Detailing	Detailing, ASP	Detailing	Detailing, ASP	Detailing
Fixed effects	state-brand, Year	state-brand, brand-year	state-brand, Year	state-brand, brand-year	state-brand, Year	state-brand, brand-year	state-brand, Year	state-brand, brand-year
Specification	Traditional SSIV	Traditional SSIV	State pop. weights	State pop. weights	Log mkt. size weights	Log mkt. size weights	Excluding large states	Excluding large states

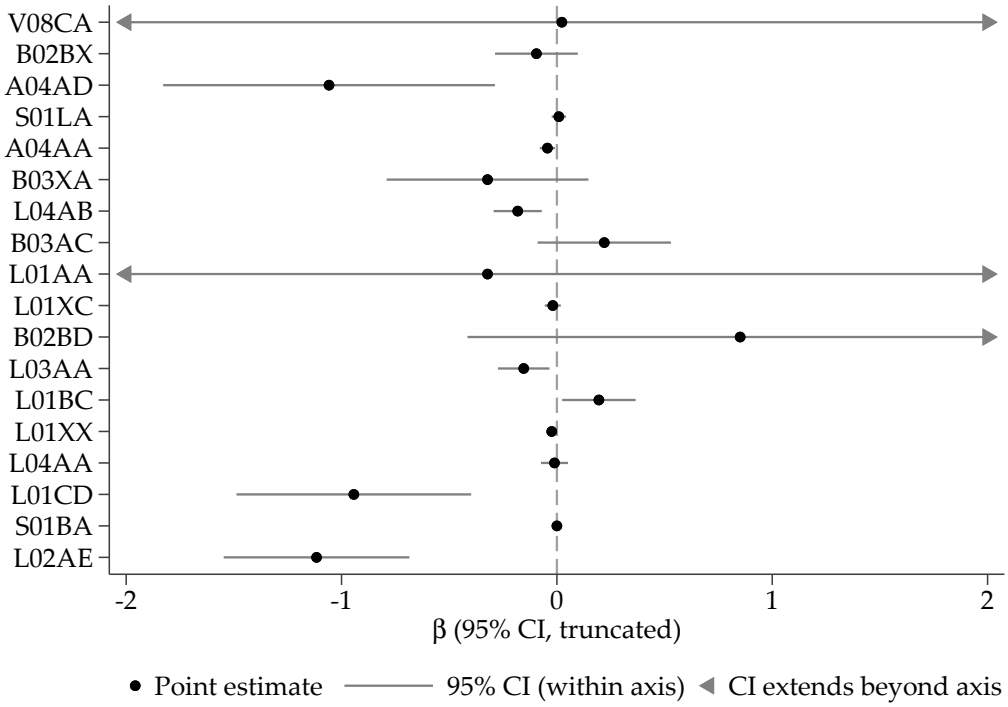
Notes: Estimates of Equation 2, using OLS and IV approaches. The data is at the state-year-brand level and include data from 2015 to 2019. The dependent variable is the brand-level market share for each brand in a given state-year. Standard errors are clustered at the state-brand level. All brands are included in the calculation of market share, but only brands without generic competitors are included in regressions. Columns 3 and 4 include regressions weighted by state population size while columns 5 and 6 includes regressions weighted by ATC-4 market size. Columns 7 and 8 exclude California, Texas, Florida, and New York.

Table A9: Robustness Checks on Regression 2, Biosimilar Sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PANEL A. OLS ESTIMATES								
Fraction Excluded	-0.522	0.005	-0.542	0.095	-0.541	0.053	-0.513	-0.004
	(0.034)	(0.034)	(0.032)	(0.056)	(0.033)	(0.048)	(0.036)	(0.035)
PANEL B. SHIFT-SHARE IV RESULTS								
Fraction Excluded	-0.653	-0.065	-0.286	0.016	-0.294	-0.082	-0.293	-0.114
	(0.048)	(0.174)	(0.036)	(0.082)	(0.035)	(0.074)	(0.034)	(0.084)
First-Stage *F*-stat.	1002.2	31.11	503.6	29.05	544.9	38.30	551.1	37.89
PANEL C. IV FIRST STAGE								
Uncovered on	1.569	0.407	0.257	0.010	0.258	0.014	0.261	0.016
National Formulary	(0.050)	(0.073)	(0.011)	(0.002)	(0.011)	(0.002)	(0.011)	(0.003)
R^2	0.876	0.925	0.405	0.118	0.413	0.153	0.417	0.146
Observations	1,809	2,142	1,809	2,142	1,809	2,142	1,657	1,962
Controls	Detailing, ASP	Detailing	Detailing, ASP	Detailing	Detailing, ASP	Detailing	Detailing, ASP	Detailing
Fixed effects	state-brand, Year	state-brand, brand-year	state-brand, Year	state-brand, brand-year	state-brand, Year	state-brand, brand-year	state-brand, Year	state-brand, brand-year
Specification	Traditional SSIV	Traditional SSIV	State pop. weights	State pop. weights	Log mkt. size weights	Log mkt. size weights	Excluding large states	Excluding large states

Notes: Estimates of Equation 2, using OLS and IV approaches. The data is at the state-year-brand level and include data from 2015 to 2019. The dependent variable is the brand-level market share for each brand in a given state-year. Standard errors are clustered at the state-brand level. All brands are included in the calculation of market share, but only brands without generic competitors are included in regressions. Columns 3 and 4 include regressions weighted by state population size while columns 5 and 6 includes regressions weighted by ATC-4 market size. Columns 7 and 8 exclude California, Texas, Florida, and New York.

Figure A3: Coefficient plot of class-specific regressions



Notes: *This Figure plots coefficients of formulary coverage from equation 2 estimated at the ATC-4 level, and their confidence intervals. We report the coefficient from a specification including both state-brand and brand-year fixed effects, as well as controls for detailing payments (equivalent to column 4 in Table 2). ATC-4s are ranked in descending order of γ, matching Figure 1. Select confidence intervals are truncated to improve visualization of the estimated coefficients.*

insignificant correlation between the two (the correlation is either positive or negative depending on how the weighting is applied).

B.7 Estimates using an event study framework

As an alternative to our SSIV methodology, we also present results using an event study approach. Traditional event studies are challenging in our setting because our design relies on differential state exposure to numerous national-level formulary changes, rather than discrete, state-specific policy events. To address this, we first construct measures of commercial coverage for each state-brand-year, capturing the average inclusion or exclusion of a drug from commercial formularies within the state. We then focus on large coverage changes, defining positive coverage shocks as the largest 5% of year-over-year inclusion changes and negative coverage shocks as the largest 5% of exclusion changes. This threshold-based approach isolates meaningful deviations in insurance coverage relative to typical year-to-year variation; however, unlike a traditional event study, state-brand coverage continues to evolve year-to-year in both treated and control groups.

In total, we include 839 negative coverage shocks involving 89 brands, and 838 positive cov-

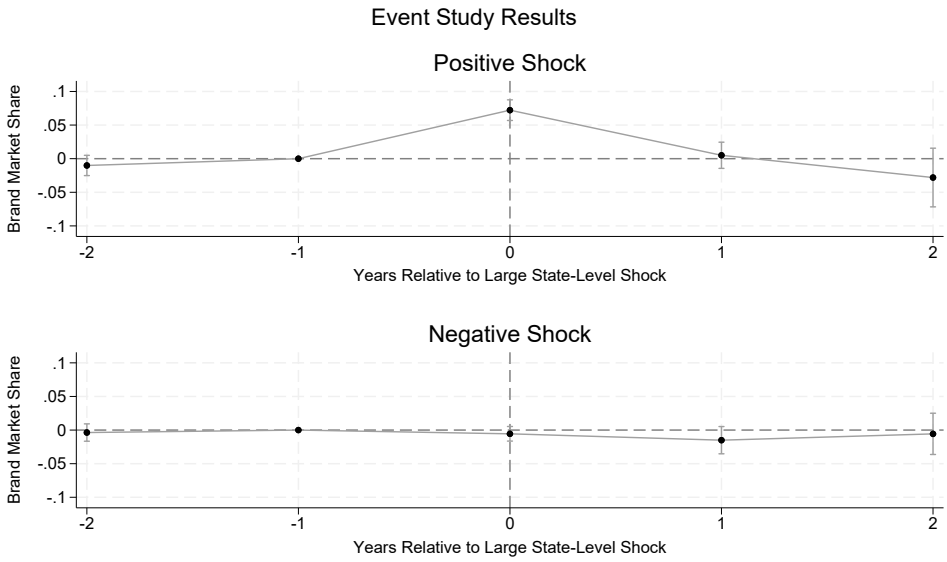


Figure A4: Event study

Notes: Event-study estimates of the effect of large state-level commercial coverage changes on a brand's Medicare market share, using the staggered difference-in-differences estimator of Callaway and Sant'Anna (2021). Positive (negative) coverage shocks are defined as the largest 5% of state-brand year-over-year inclusion (exclusion) changes between 2015 and 2019. Specification includes brand-year fixed effects, with never-treated and not-yet-treated drug-states as the control group.

erage shocks involving 69 brands, with at least one event observed in all 50 states and DC.[51] The average negative coverage shock corresponds to a 9.4pp reduction in state-level coverage, while a positive coverage shock corresponds to a 17.5pp increase.

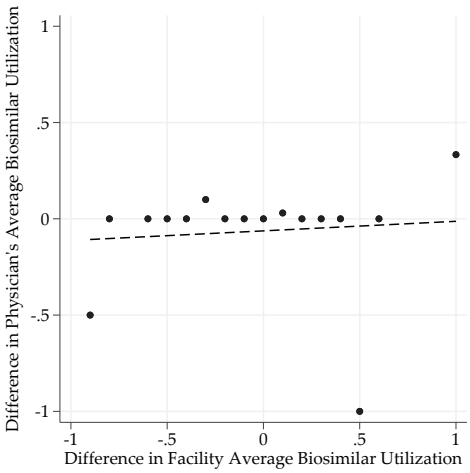
To estimate how Medicare utilization changes around these events, we follow Callaway and Sant'Anna (2021), whose methodology can accommodate the staggered timing of events across states. The specification includes brand-year fixed effects to account for secular trends in drug utilization and uses as a control group the never-treated and not-yet treated observations (drug-state data points without large coverage shocks during the period).

The results show no evidence of pre-existing trends in Medicare utilization prior to large coverage shocks. For positive shocks, we find an immediate increase in utilization in the year of the coverage change, which attenuates within one year. For negative shocks, we observe a non-significant decline in Medicare market share. The absence of pre-trends supports the validity of our identification strategy, and the event study results are broadly consistent with our primary estimates.

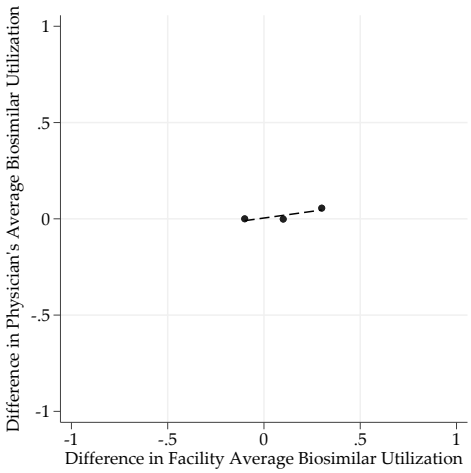
B.8 Dual-practicing physician plots for all ATC-4 classes

We report "mover" graphs for each class in our core sample (18 graphs in total, Figures A5–A7), and for the biosimilar sample (4 graphs in total, Figure A8).

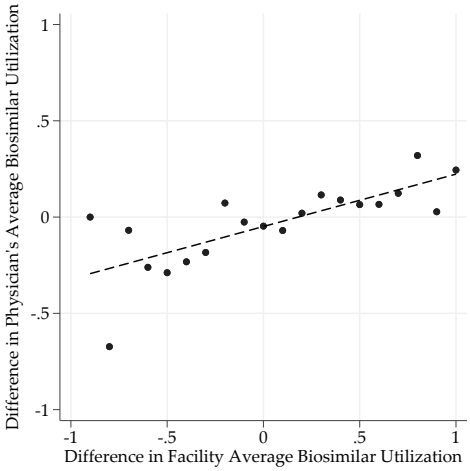
[51] We exclude drug-states with multiple shocks between 2015 and 2019 from the analysis.



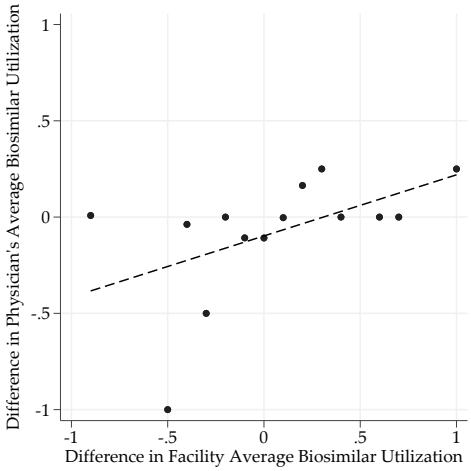
(a) L02AE



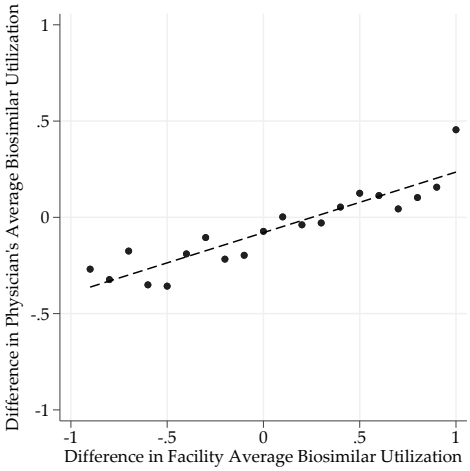
(b) S01BA



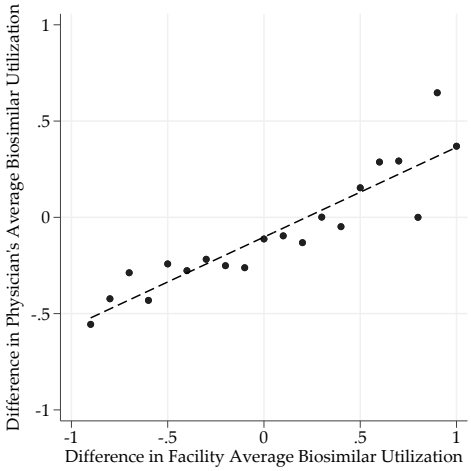
(c) L01CD



(d) L04AA

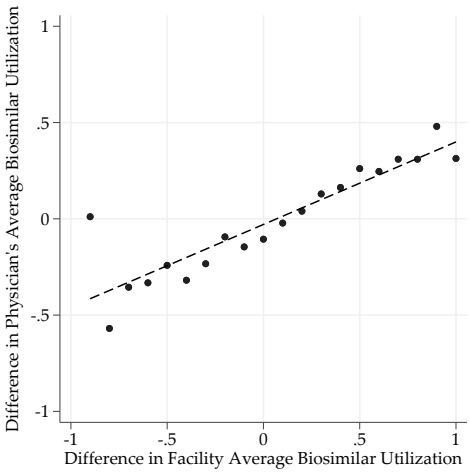


(e) L01XX

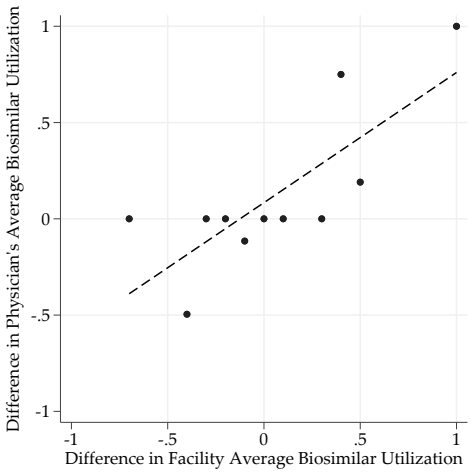


(f) L01BC

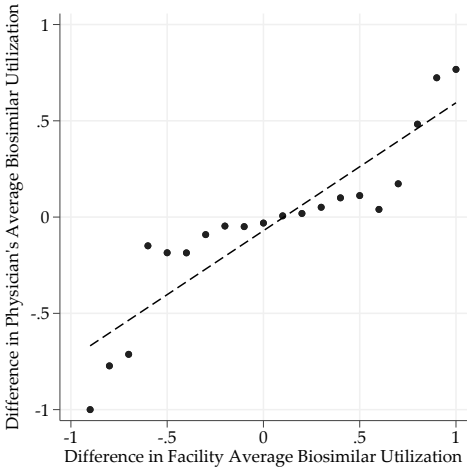
Figure A5: ATC-4 Plots, Part 1



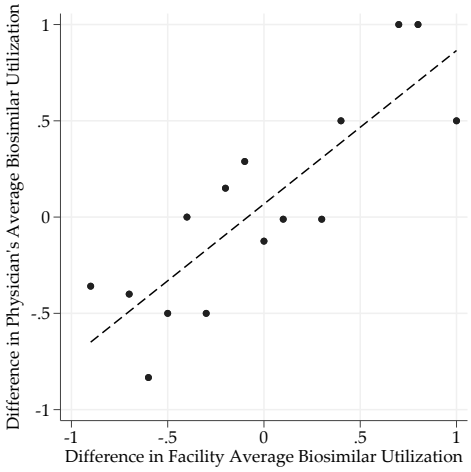
(a) L03AA



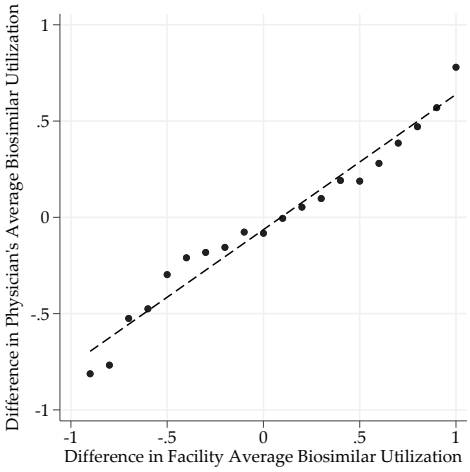
(b) B02BD



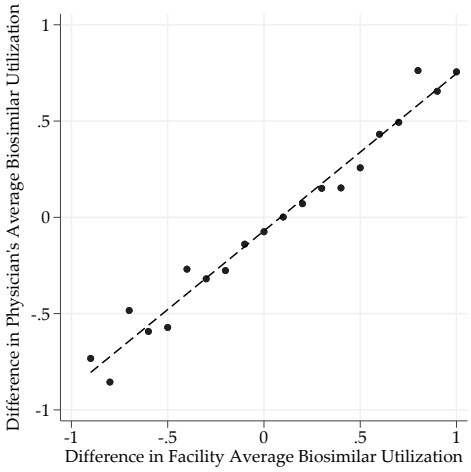
(c) L01XC



(d) L01AA

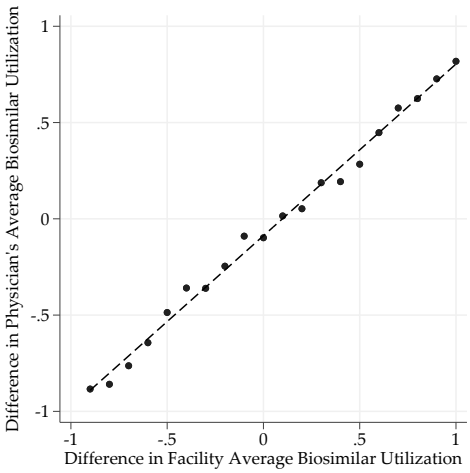


(e) B03AC

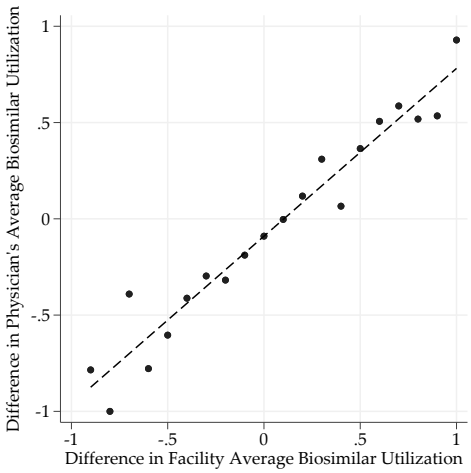


(f) L04AB

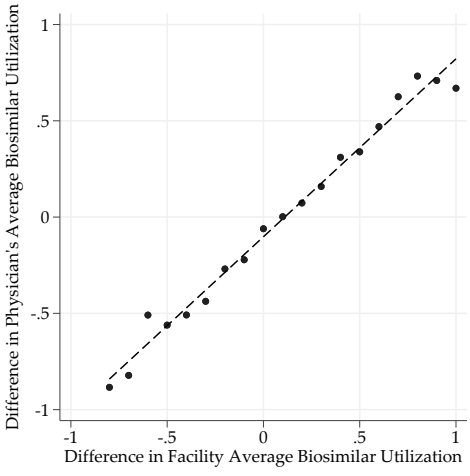
Figure A6: ATC-4 Plots, Part 2



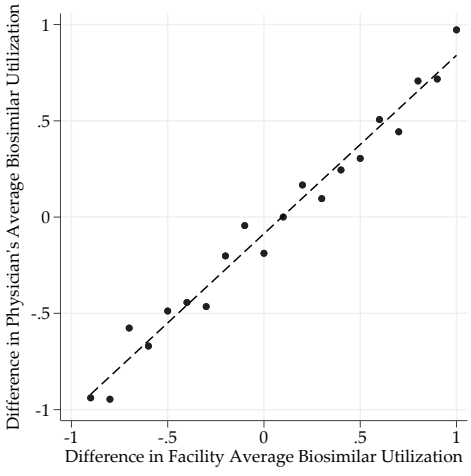
(a) B03XA



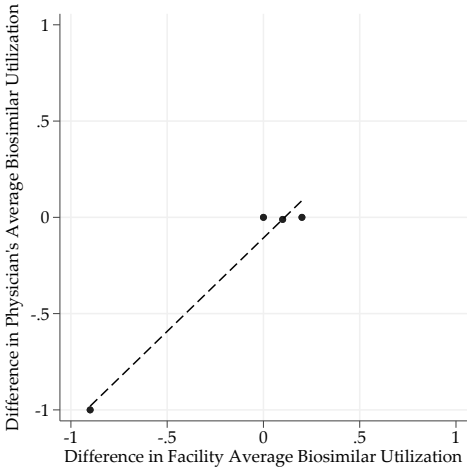
(b) A04AA



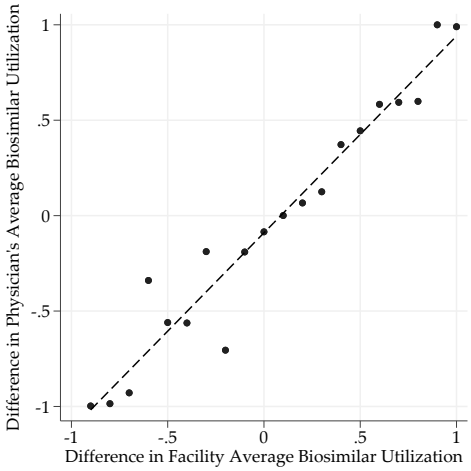
(c) S01LA



(d) A04AD

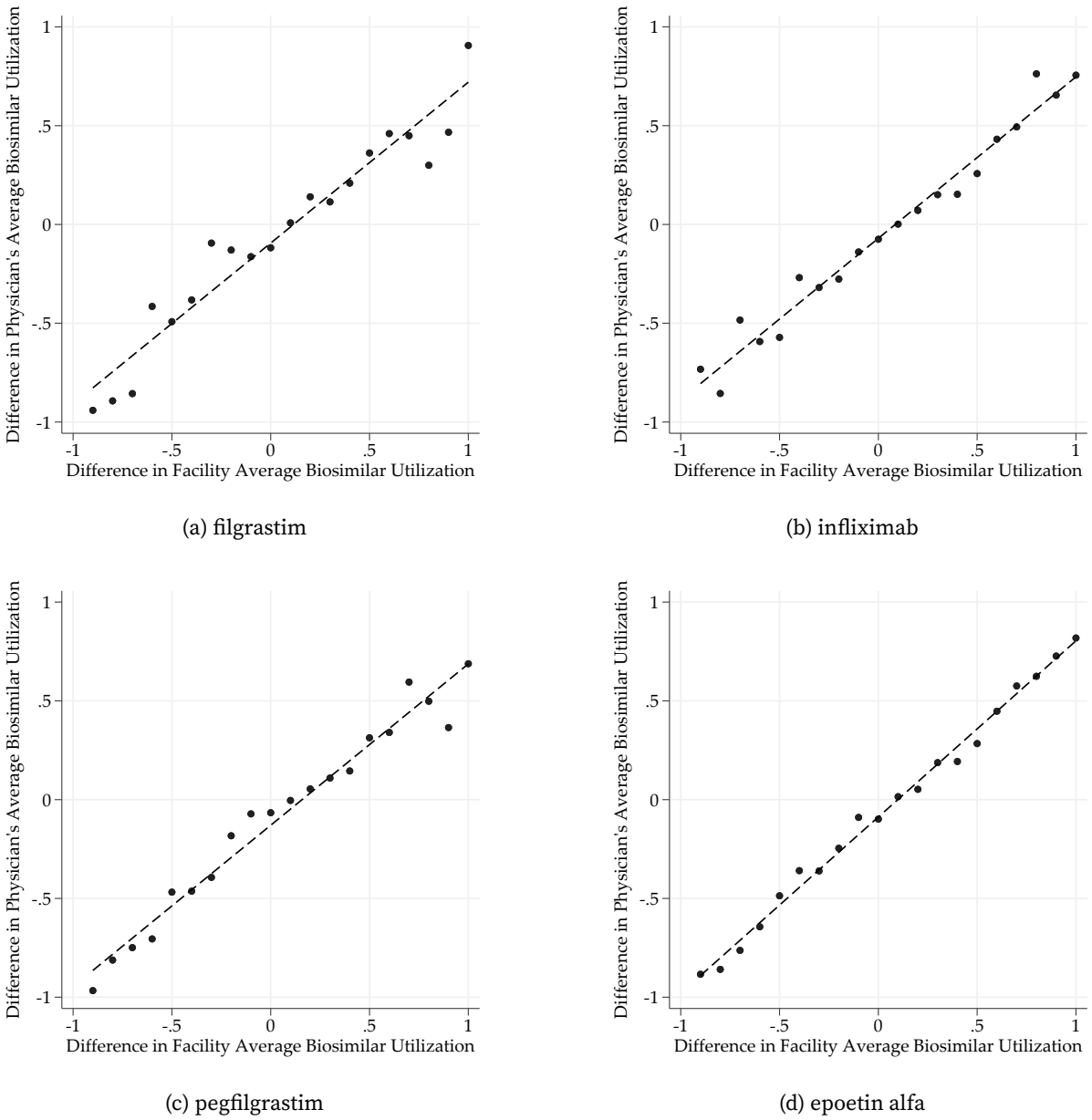


(e) B02BX



(f) V08CA

Figure A7: ATC-4 Plots, Part 3



(a) filgrastim

(b) infliximab

(c) pegfilgrastim

(d) epoetin alfa

Figure A8: Change in Physician Biosimilar Prescribing Across Facilities

Notes: Figure reports estimates based on Equation 3 in the biosimilars sample. The sample includes all physicians prescribing a given drug in two facilities, using data from the 20% Medicare Part B data between 2015 and 2019.

B.9 Variation analysis using the Cooper et al. (2019) methodology

As noted in the main text, only 10% of physicians in our sample practice at exactly two facilities. To provide broader complementary evidence, we adapt the approach in Cooper et al. (2019) to our setting, in order to assess the importance of facilities versus physicians.

The regression equation we estimate projects the physician-facility-level market share of the leading product in an ATC-4 (or reference biologic in a class defined by a molecule) onto the fol-

Table A10: R^2 values from different specifications for predicting the market share of the leading drug in each ATC-4

ATC-4	R^2				Observations	Number of drugs
	(1)	(2)	(3)	(4)		
V08CA	0.14	0.95	0.69	0.98	91,786	2
B02BX	0.93	0.99	0.92	1.00	7,252	2
A04AD	0.32	0.89	0.81	0.97	63,114	3
S01LA	0.47	0.67	0.75	0.85	104,288	4
A04AA	0.74	0.89	0.92	0.98	95,032	4
B03XA	0.27	0.77	0.73	0.94	41,074	2
L04AB	0.59	0.86	0.79	0.94	61,293	3
B03AC	0.21	0.72	0.77	0.93	252,978	6
L01AA	0.70	0.95	0.88	0.99	15,111	3
L01XC	0.20	0.48	0.79	0.87	1,456,575	25
B02BD	0.10	0.69	0.67	0.86	5,819	11
L03AA	0.24	0.55	0.68	0.87	205,198	7
L01BC	0.24	0.63	0.69	0.90	56,745	5
L01XX	0.38	0.56	0.74	0.90	137,930	10
L04AA	0.58	0.86	0.97	1.00	17,760	5
L01CD	0.23	0.39	0.69	0.83	43,155	3
S01BA	0.99	0.99	0.99	1.00	7,983	3
L02AE	0.29	0.70	0.90	1.00	7,818	3
Combined	0.43	0.74	0.80	0.92	2,670,911	101
Drug FE	X					
Drug-Facility FE		X		X		
Drug-Physician FE			X	X		

Notes: we report the R^2 values from four regression specifications that use the market share of the leading product at the ATC 4-facility-physician level: (i) drug FEs only, (ii) Drug-Facility FEs, (iii) Drug-Physician FEs, and (iv) Drug-Facility and Drug-Physician FEs.

lowing four sets of variables: (i) drug FEs, (ii) Drug-Facility FEs, (iii) Drug-Physician FEs, and (iv) Drug-Facility and Drug-Physician FEs. We run a separate regression for each substitution class. Tables A10 and A11 report the R^2 coefficients at the ATC-4 and molecule level, respectively. Classes are ordered in descending value of γ (i.e., the estimated facility coefficient in the dual-homing physician analysis). In other words, classes in the first few rows are the ones where we'd expect the relative explanatory power of Drug-Facility FEs to be strongest compared to Drug-Physician FEs.

As one might expect, there are far more physicians (152,250) than facilities (9,636) in the data. In spite of this, for 7 out of 18 ATC-4 substitution classes and 3 out of 4 molecule substitution classes the (unadjusted) R^2 coefficient is higher in column (2) than in column (3).

Table A11: R^2 values from different specifications for predicting the market share of the leading drug in each biologic-biosimilar market

Molecule	R^2				Observations	Number of drugs
	(1)	(2)	(3)	(4)		
Epoetin Alfa	0.27	0.77	0.73	0.94	41,074	2
Filgrastim	0.14	0.89	0.68	0.98	62,644	4
Infliximab	0.59	0.86	0.79	0.94	61,293	3
Pegfilgrastim	0.62	0.84	0.85	0.96	63,786	3
Combined	0.44	0.85	0.78	0.95	228,797	12
Drug FE	X					
Drug-Facility FE		X		X		
Drug-Physician FE			X	X		

Notes: we report the R^2 values from four regression specifications that use the market share of the reference biologic at the molecule-facility-physician level: (i) drug FEs only, (ii) Drug-Facility FEs, (iii) Drug-Physician FEs, and (iv) Drug-Facility and Drug-Physician FEs.

In addition, Figure A9 plots γ (i.e., the parameter estimated from variation using dual-homing physicians) vs. a new parameter ζ calculated as

$$\zeta = \frac{R^2_{(2)} - R^2_{(1)}}{R^2_{(4)} - R^2_{(1)}}$$

where $R^2_{(n)}$ is the R^2 of the regression specification from column (n) in the table. Like γ, ζ is intended to capture the *relative contribution of facilities to drug choice relative to the combined contribution of facilities and physicians.* Column 2 reflects the overall contribution of facilities, while column (4) reflects the combined contribution. We then subtract the R^2 from column (1) because we may otherwise overstate the contribution of facilities.[52] We find that the two metrics are strongly positively correlated (correlation coefficient is around 0.63).

B.10 Summary stats on facility-mix

Next, we provide additional evidence on the potential role played by hospitals versus other types of facilities. For example, some classes of drugs may be more likely to be administered at hospitals versus other outpatient clinics, and hospitals may be better positioned to enforce stocking decisions. To test this story, we compare, at the ATC-4 level, the estimated γ coefficient and the

[52] To see why, consider, e.g., the ATC-4 class B02BX. Since the leading product has 93% market share, there is very little identifying variation in the data to begin with, and we want to avoid erroneously attributing it to facilities.

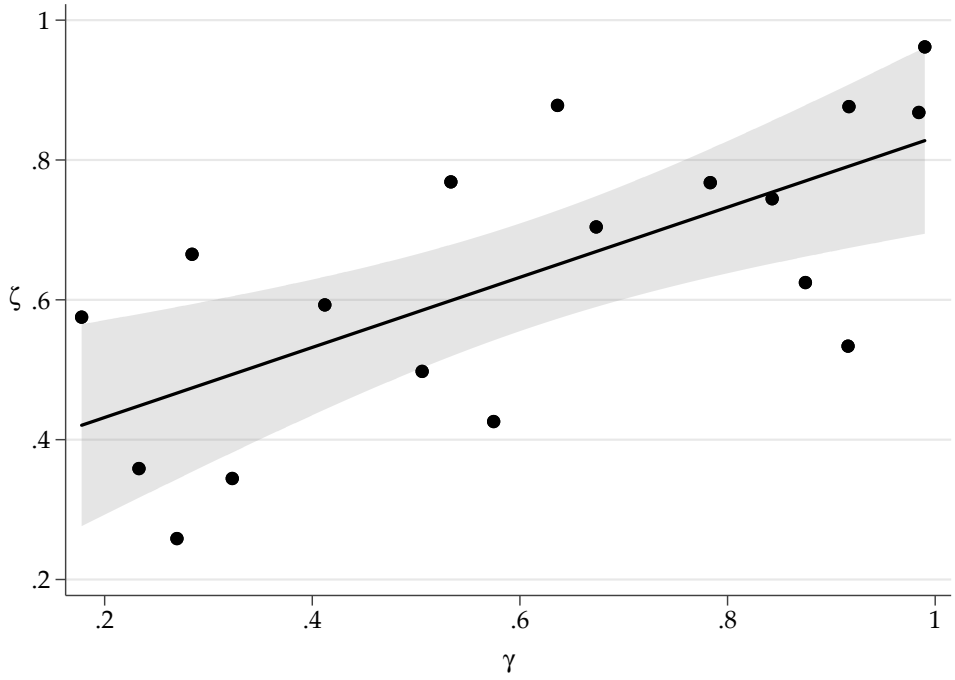


Figure A9: Correlation between γ and ζ

share of claims that are hospital-based. Figure A10 presents the results. There is variation in the share of drug administrations at hospitals versus other facilities across ATC-4 classes. Four classes have hospital shares in the 0.15 to 0.3 range, whereas the remaining classes have hospital shares around 0.8. However, we find high and low values of γ in both buckets. The correlation between the two metrics is slightly negative, but insignificant.

B.11 Facility-level utilization rates

From qualitative interviews, we learned that facilities set internal formularies and earn financial considerations from manufacturers if they prefer one brand within a given molecule. We provide additional descriptive evidence consistent with this behavior by looking at facility-level prescribing patterns using the 20% random sample of fee-for-service Medicare beneficiaries.

In the majority of substitution classes we consider, facilities overwhelmingly prescribe a single brand over all possible alternatives (Figure A11). Within our core sample (Panel A), in three-fourths of substitution classes, we find the majority of facilities concentrate their administrations on a single brand within the substitution class.[53] Classes where most facilities administer a mix of brands are also ones where our mechanism analysis suggests physicians play a larger role than facilities (see Figure 1 in the main text). Within our biosimilar sample (Panel B), about three quarters

[53] We use 80 percent as a threshold for dominance because discussions with hospital pharmacists and administrators responsible for facility purchasing indicated this threshold as a common requirement for percentage-of-sales contracts. We exclude facilities with fewer than 5 claims.

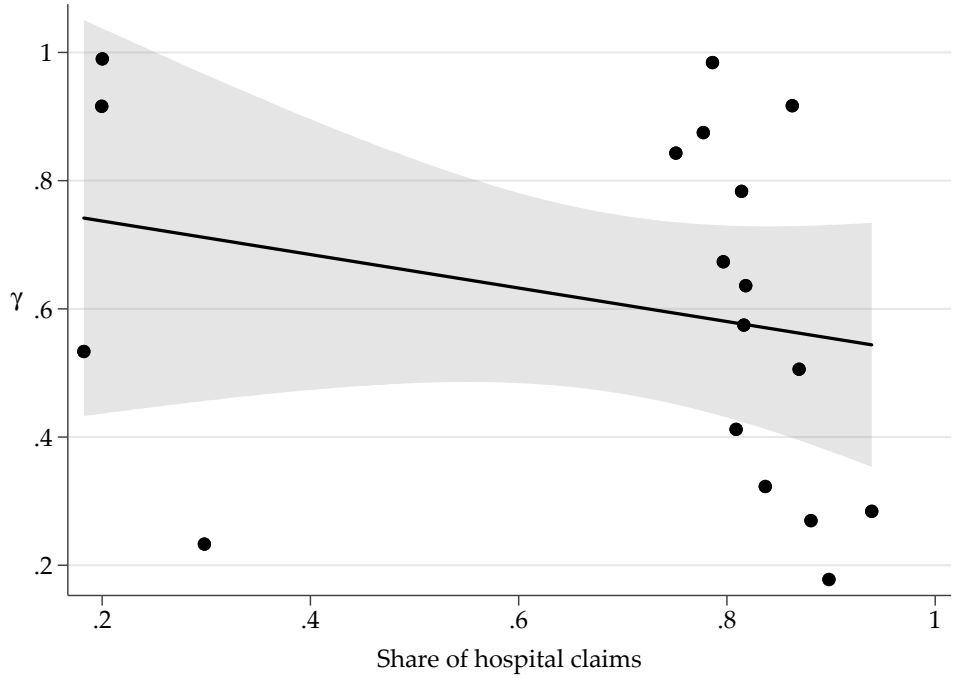


Figure A10: Class-level comparison between γ and share of hospital claims

of facilities had a dominant brand (either a biosimilar or the reference biologic). In both samples, drugs other than the market leader have greater than 80 percent market share in a significant fraction of facilities.

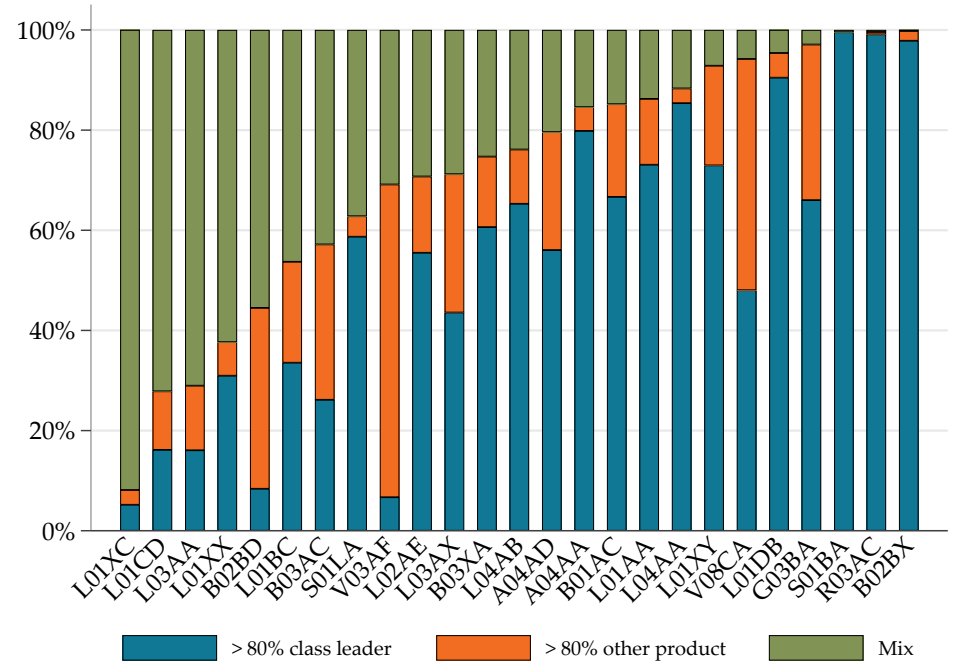
C Additional details on the model calibration

C.1 Simplifying assumptions and avenues to relax them

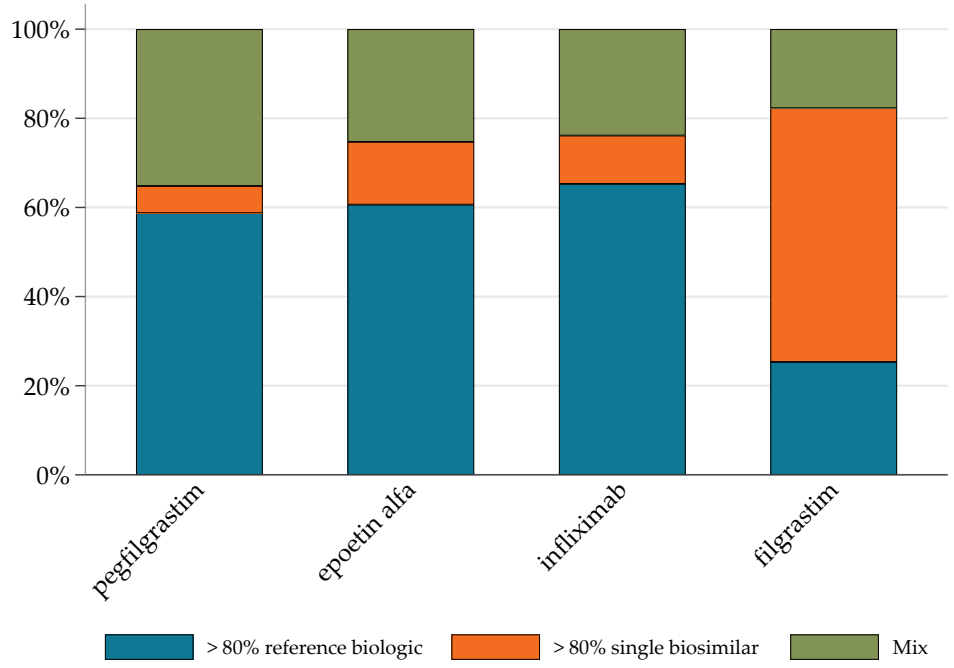
Our model is best understood as a calibration exercise rather than a full structural estimation, and it imposes a number of simplifying assumptions that trade realism for tractability. This appendix catalogs the most consequential of these assumptions, explains the rationale behind each one, and discusses what relaxing them would require. We group them into three categories: features of the market structure that our framework abstracts from; simplifications to the behavior of individual agents; and dynamic or cross-class considerations that fall outside our static, single-class setup.

Market-structure abstractions

- **Absence of 340B facilities.** The single most important simplification is that we do not model the 340B Drug Pricing Program. As discussed in the main text, the program requires manufacturers to offer deeply discounted ceiling prices to eligible facilities, and the resulting rebates accrue as spread revenue on facility margins rather than as savings to payers. Be-



(a) Core Sample



(b) Biosimilar sample

Figure A11: Fraction of Facilities with over 80% Utilization of a Single Brand, by ATC4 and Molecule

Notes: Figure is calculated using the 20% Medicare Part B data from 2019. We exclude facilities with fewer than 5 administrations in a given class.

cause these rebates are larger in dollar terms on the originator than on the biosimilar, 340B facilities have a strong incentive to stock the originator exclusively even when its ASP is substantially higher than the biosimilar's (Bond et al., 2023). This mechanism is a plausible explanation for the central puzzle left by our calibration — namely, why the originator retains a dominant stocking share in spite of high prices — and it is one that our model, which does not distinguish 340B from non-340B facilities, cannot reproduce. Relaxing this assumption would require facility-level 340B status, reliable data on 340B ceiling prices (which are not publicly reported), and a modified facility profit function with a separate margin channel for 340B eligible units.

- **Geographic concentration of formulary coverage.** In our model, marginal gains in formulary coverage are diffused across the entire payer-mix distribution of facilities, including facilities that are far from the stocking margin. In practice, biosimilar manufacturers can concentrate their marketing efforts on insurance companies whose enrollees are geographically clustered, which concentrates the effect of a coverage change on a smaller but potentially more marginal slice of the facility distribution.[54] Relaxing this assumption would require a spatial version of the model in which facilities, insurers, and patients are linked through a geographic structure rather than aggregated into a single national market. The required data—facility location, insurer market areas, and plan enrollment by geography—is in principle available (e.g., through all-payer claims datasets), but acquiring it and solving this more complicated model goes beyond the scope of this paper.

- **Exogenous list prices.** We take manufacturer list prices p_j^{list} as given. This is a simplification of convenience: solving a model in which list prices, exclusive prices, and rebates are set jointly is considerably more complex, and list prices are known to respond to many forces external to the pricing game we model. The Medicaid best-price and inflation-penalty rules lead manufacturers of drugs with high Medicaid exposure to raise list prices more slowly over time (Feng et al., 2023). Political-economy concerns, such as the threat of harsher price regulation, also shape list-price growth (Che, 2025). High list prices raise the indirect value of insurance by making uninsurance less attractive, benefiting manufacturers (Conti et al., 2022). And list prices interact with patient cost-sharing: insurers can use higher list prices as a tool to circumvent benefit-design regulations that mandate minimum coverage generosity (Feng and Maini, 2025). Incorporating these forces would require embedding our game inside a richer pricing environment that is well beyond the scope of this paper.

- **Representative insurer and posted rebates.** We model a single representative insurer that

[54] An extreme example of this phenomenon is Kaiser Permanente, which has been among the most aggressive adopters of biosimilars among US insurers, systematically favoring them over the reference biologics across multiple drug classes. Kaiser is highly vertically integrated: its health plans operate an exclusive network of Kaiser-owned facilities, served by physician groups that contract exclusively with Kaiser. This close alignment between insurer and provider means that a product that gains preferred formulary status with Kaiser health plans is virtually guaranteed to obtain exclusive stocking across all Kaiser facilities.

chooses a formulary arrangement and receives a single per-unit rebate from each manufacturer. This is a tractability assumption: allowing multiple plans with heterogeneous bargaining power over rebates would introduce a standard bilateral negotiation problem on top of the already non-trivial equilibrium we solve. As noted in the main text, our formulation is equivalent to one where manufacturers post a single rebate and atomistic plans each pick a formulary in response, which we view as a reasonable approximation.

- **Exogenous insurer-to-facility markup.** Commercial insurers pay facilities a fixed multiple β of the drug's ASP, which we set to 1.75 using Transparency in Coverage (TiC) data. We do not model the insurer–facility negotiation that determines this multiplier, both because we lack the commercial claims data needed to identify a negotiation model at the drug level and because the prevailing view in the literature is that insurer–hospital negotiations occur over a single hospital-wide multiplier applied to a menu of services, rather than drug-by-drug (Lewis and Pflum, 2015; Ho and Lee, 2017; Prager, 2020). If this view is approximately correct, the multiplier depends on broad market conditions and is unlikely to respond to within-class dynamics in a drug class representing a small share of overall hospital revenue. Relaxing this assumption would require explicit data on commercial transaction prices.

Behavioral abstractions

- **Exogenous patient-to-facility assignment.** We treat patients as exogenously assigned to facilities, ruling out the possibility that a patient responds to a facility's stocking choice by seeking care elsewhere. Allowing an extensive margin of facility choice would introduce recapture effects into both the manufacturer's pricing problem and the facility's stocking problem, and identifying the relevant substitution patterns would require commercial claims data that we do not have. This assumption has implications for welfare. Taken literally, it implies that a patient whose assigned facility exclusively stocks a product not covered by her plan cannot receive the drug at all — which would generate large welfare losses, even though in practice patients can seek care at other facilities. Two features of our setting mitigate this concern. First, exclusive pricing arrangements do not actually require exclusive stocking; informal interviews suggest that an 80% market-share threshold is common, so most patients can still receive a covered drug at their assigned facility. Second, our insurer utility function includes a baseline preference ζ_{12} for open formularies that absorbs some of the welfare benefit of facility choice. The model is therefore slightly internally inconsistent: the absence of an extensive margin implies that all patients access the covered drug, while ζ_{12} implicitly prices in the value of choice. We do not believe this inconsistency meaningfully affects our estimates.

- **Homogeneous facility size and parametric payer-mix distribution.** Facilities differ only in their payer mix θ_h and not in the volume of patients they serve; in addition, we take $G(\theta)$ to be a scaled Dirichlet distribution with a fixed concentration parameter. Relaxing the size

assumption is conceptually straightforward (facility size would enter every profit expression as a multiplicative weight) but adds another moment (size distribution) that would need to be calibrated. Relaxing the Dirichlet assumption—for instance, by using the empirical distribution of payer mixes from claims data—is also feasible, but requires access to commercial claims data.

No outside option. The normalizations $\delta_2 = \zeta_2 = \kappa_2 = 0$ and the absence of an outside option in the consumer, facility, and insurer logits imply that every patient in the model that has the option ultimately receives one of the two products we study (the only exception is what happens when the intersection of the stocking and formulary coverage choice is the empty set). This is a modeling choice, not a clinical claim: in practice, patients with the indications we study (primarily chemotherapy-induced anemia in non-ESRD settings) can also be treated with darbepoetin alfa — a longer-acting erythropoiesis-stimulating agent in the same class — or with alternative strategies such as blood transfusion or watchful waiting, and the latter two have become more common since the 2007–2008 FDA black-box warnings on ESA safety. We focus on epoetin alfa because it has a clean two-product (originator + biosimilar) structure that makes the pricing game tractable; the outside option is absorbed into the sample-selection margin (we condition on observed ESA administrations) rather than the within-sample choice margin, and the zero normalizations simply fix the scale of the remaining logit. Extending the model to admit molecular substitution or non-treatment would require data on darbepoetin use and on the extensive margin of ESA treatment, and would materially change the interpretation of the consumer and insurer preference parameters.

Dynamic and cross-class considerations

- **Static, steady-state equilibrium.** Our model is static, whereas the real Medicare Part B reimbursement rule uses a six-month lagged ASP. The lag generates dynamic incentives to set a high price at the beginning of a drug's life cycle and lower it slowly over time, as shown in Acquatella et al. (2026). Our equilibrium is best interpreted as the steady state that such a dynamic game converges to once the transitory incentives associated with the lag have been fully priced in; it is not intended to describe the transition path immediately following biosimilar entry. Embedding our stage game inside the dynamic environment of Acquatella et al. (2026) is a natural extension but is computationally demanding given the already expensive inner-loop equilibrium solve.

- **Zero marginal cost.** We treat both manufacturers as having zero marginal cost, so that the profit expressions in equation (15) coincide with revenue. In practice, while biologic manufacturing costs are non-trivial, we expect marginal costs of the two manufacturers to be similar (minimizing the distortionary effect of our assumption), though differences may arise if, e.g., the originator manufacturer has managed to lower production costs over time through

learning by doing, or if the biosimilar manufacturer is better able to incorporate newer manufacturing techniques. Introducing heterogeneous marginal costs would scale the prices our model predicts and may have implications for the competitive equilibrium. Unfortunately we do not have sufficient moments to estimate these costs.

- **No bundling.** We model purchases of each drug in isolation. In practice, facilities often purchase drugs from manufacturers that supply them with other products, either through formal bundled contracts or through informal cross-product discounting. Evidence of bundling exists in vaccine procurement (Gardenghi, 2024), and Hermosilla et al. (2025) documents bundling-like patterns specifically in the biosimilar market. Relaxing this assumption would require tracking manufacturer product portfolios and modeling cross-product discounting, both of which take us beyond the two-drug scope of the current paper.
- **Single drug class.** Our calibration is based on a single drug class (epoetin alfa) with a single biosimilar entrant, which greatly simplifies the computational burden of solving the equilibrium and identifying the parameters. Extending the model to classes with three or more biosimilars would not alter the economic logic, but it would enlarge the strategy space, complicate the stocking game (the number of exclusive and partial-stocking configurations grows combinatorially), and increase the sensitivity of the solver to the choice of starting values. Joint estimation across classes would additionally require parameter restrictions to exploit the cross-class information — for example, common consumer preferences within a therapeutic area—which we leave to future work.

C.2 Medicare Part B reimbursement and how it enters the model

This subsection documents the Medicare Part B reimbursement rule that applies to physician-administered biologics and biosimilars in our sample, and the way the rule enters the model. It also records the 2022 Inflation Reduction Act (IRA) adjustment to the rule, which is the policy behind the ASP-multiplier counterfactual in Section 4.

The 2019 reimbursement rule

Medicare Part B reimburses physician-administered drugs based on a *Medicare-allowed amount* A_j that is tied to a drug's Average Sales Price (ASP). Three statutory components determine A_j and the way it is shared across payers.

First, the *statutory add-on*. For originator biologics, the allowed amount is the drug's own ASP plus a 6% mark-up: $A_1 = (1 + \psi_1)\, p_1^{ASP}$ with $\psi_1 = 0.06$. For biosimilars, the 2018 reference-biologic rule ties the add-on to the *reference product's* ASP rather than the biosimilar's own: $A_2 = p_2^{ASP} + \psi_2\, p_1^{ASP}$ with $\psi_2 = 0.06$ in 2019. The reference-biologic rule was introduced in 2018 to make the biosimilar add-on — and therefore the facility margin on biosimilar administrations — less sensitive to the biosimilar's own price, addressing a concern that the pre-2018 formula made entry at a low biosimilar ASP mechanically unprofitable for facilities.

Second, the *Medicare sequester*. Since 2013, the Bipartisan Budget Control Act of 2011 has reduced the federal share of the allowed amount by 2%. Because Medicare pays only 80% of the allowed amount (the remainder is patient coinsurance; see below), the sequester applies to that 80% only and translates into a uniform reduction factor on the provider receipt of $\sigma = 1 - 0.80 \times 0.02 = 0.984$.[55]

Third, *patient coinsurance*. Medicare beneficiaries pay a 20% coinsurance share of the allowed amount, which is unaffected by the sequester. Writing $c = 0.20$ for the coinsurance share and collecting terms, the per-unit amount the facility actually collects on a Medicare unit is

$$P_j = \sigma A_j = \underbrace{(\sigma - c) A_j}_{\text{Medicare program payment } M_j} + \underbrace{c A_j}_{\text{patient coinsurance } C_j} = 0.784 A_j + 0.20 A_j. \tag{A1}$$

The facility books P_j per Medicare unit, of which $M_j = 0.784 A_j$ comes from the Medicare program and $C_j = 0.20 A_j$ comes from the patient. Combining the statutory add-on and the sequester, the net mark-up over ASP that the facility ultimately collects on an originator Medicare unit is $(1 + 0.06) \cdot 0.984 - 1 \approx 4.3\%$, which is the value of ψ used in the hospital profit equations of the main text.

The 2022 Inflation Reduction Act adjustment

The Inflation Reduction Act of 2022 increased the biosimilar statutory add-on from 6% to 8% for qualifying products, leaving the originator's 6% add-on unchanged.[56] In the notation above, the IRA bump replaces $\psi_2 = 0.06$ with $\psi_2 = 0.08$, so that the biosimilar allowed amount becomes

$$A_2^{\text{IRA}} = p_2^{\text{ASP}} + 0.08\, p_1^{\text{ASP}},$$

and the post-sequester provider receipt on a biosimilar Medicare unit rises in lockstep. The originator reimbursement is untouched. The Medicare sequester and the 20% coinsurance split continue to apply as before.

How the rule enters the model

The Medicare rule enters the model in three places and does not enter in two others. Separating the "where" from the "where not" is useful because it explains why the counterfactual ASP-multiplier change has a muted equilibrium effect even though it looks, at first glance, like a direct revenue boost to biosimilar administrations.

1. **Hospital profit (provider receipt P_j).** Each Medicare administration generates a per-unit margin of $P_j - p_j^{\text{list}}$ at dual-stocking facilities and $P_j - p_j^{\text{excl}}$ at exclusive-stocking facilities,

[55] The sequester was briefly suspended during the COVID-19 pandemic but was in force throughout 2019, which is the sample year for our calibration.

[56] The bump is conditional on the biosimilar satisfying certain criteria; since all of the biosimilars we study satisfy those criteria, we take the bump as universally applied in the counterfactual.

where $P_j = \sigma A_j$. The provider receipt therefore enters all three branches of the hospital profit expression in equation (10) (exclusive stocking of drug 1, exclusive stocking of drug 2, and stocking both). This is the most direct channel through which a change in the statutory add-on affects the equilibrium: a higher ψ_2 raises the margin on biosimilar Medicare units and, all else equal, makes exclusive biosimilar stocking more attractive to facilities with a larger Medicare share.

2. **Medicare consumer logit (allowed amount A_j).** At dual-stocking facilities, the patient–physician choice between the two drugs for a Medicare patient is evaluated at the full allowed amounts (A_1, A_2), which are the Medicare-side analogue of the commercial-side marked-up prices $(\beta p_1^{ASP}, \beta p_2^{ASP})$. Allowed amounts — rather than the coinsurance share $c\,A_j$ — are the appropriate price in the logit because the price sensitivity parameter η is estimated on total prices and therefore already absorbs any population-average scaling between what the payer sees and what the patient pays out of pocket (as well as physician agency). Under the reference-biologic rule, changes in the originator's ASP shift *both* A_1 and A_2, which matters for the feedback between consumer choices and the manufacturer pricing game.

3. **Payer spending accounting.** At the equilibrium stocking and formulary choices, total Medicare-side spending is computed as the sum of the Medicare program payment and patient coinsurance over all Medicare units:

$$\text{Spend}^{\text{med}} = \sum_{j \in \{1,2\}} P_j\, V_j^{\text{med}} = \sigma\big[A_1 V_1^{\text{med}} + A_2 V_2^{\text{med}}\big],$$

   where V_j^{med} is the total Medicare volume of drug j implied by the equilibrium choices. The Medicare program component $M_j V_j^{\text{med}}$ and the patient coinsurance component $C_j V_j^{\text{med}}$ are reported separately in the counterfactual tables.

Two places where the Medicare rule does *not* enter directly are worth highlighting because they shape the magnitude of the counterfactual effects reported in Section 4.

- **Manufacturer revenue.** Manufacturers sell units to facilities at either p_j^{list} or p_j^{excl}, and separately pay a per-unit rebate r_j to commercial insurers. The Medicare statutory add-on, sequester, and coinsurance are all downstream of the manufacturer–facility transaction; none of them appears in equation (15). A change in the admin fee therefore affects manufacturer revenue only indirectly, through its influence on facility stocking decisions and on the consumer logit at dual-stocking facilities.

- **ASP identity (equation 6).** The statutory ASP is the volume-weighted manufacturer-side transaction price net of commercial rebates. Neither the statutory add-on nor the sequester appears in the numerator or denominator of equation (6). The rule still feeds into the ASP indirectly through the consumer logit at dual-stocking facilities — which determines how

volume at those facilities splits between the two drugs — but it does not add a term to the identity itself.

These two features together explain the central finding of the ASP-multiplier counterfactual in Section 4: because the admin fee enters the facility's stocking margin but not the manufacturer's direct revenue or the ASP identity, a change from 6% to 8% alters facility incentives only at the margin, and only for the Medicare share of volume. Whether this is sufficient to flip a facility's stocking decision depends on the counterweight from commercial-side margins, which are governed by βp_j^{ASP} and are unaffected by the admin-fee change. The counterfactual finds that the commercial counterweight dominates at most facilities, so the IRA bump has a muted equilibrium effect.

C.3 Estimation routine

Sample moments

- **Fraction of enrolled lives in commercial plans whose formulary covers** F **(**λ_F**):** calculated from MMIT data.
- **Fraction of patients served by facilities stocking** S **(**γ_S**):** calculated from CMS 20% sample carrier and outpatient files, under the assumption that facilities with > 95% of administration going to a single brand are exclusively stocking that brand.
- **Fraction of reference product administrations among Medicare patients visiting open stocking facilities (**$s_1^{\textbf{open}}$**):** calculated from CMS 20% sample carrier and outpatient files, restricted to open-stocking facilities according to the definition above.
- **ASP (**p_j^{ASP}**):** volume-weighted average of quarterly ASP, with volume coming from SSR Health and ASP coming from publicly reported CMS data.
- **Facility exclusive acquisition cost (**$p_j^{\textbf{excl}}$**):** Dosage-adjusted average NDC-level price for brand j (converted to HCPCS-level price to align with ASP) from the Federal Supply Schedule (FSS). FSS prices are guaranteed to all federal purchasers of drugs and are based on the manufacturer-reported "most-favored" pricing for comparable purchasers in the commercial market, including hospitals and physician practices that purchase drugs for use in facilities. Importantly, FSS prices exclude rebates given to 340B facilities, which apply only to covered entities and include legally mandated discounts that would not be reflective of exclusive prices. Using FSS prices as a proxy for exclusive pricing is reasonable, because they represent the lowest prices offered to facilities, analogous to the discounted "exclusive" price a facility would receive for committing to use a single product, above and beyond standard prompt-pay or other routine discounts.

- **Rebate (r_j):** We do not observe rebates directly. We back them out by inverting the ASP identity in equation 6. Define the volume-weighted manufacturer-side transaction price as

$$\text{WAP}_j \;=\; f_j^{\text{excl}}\, p_j^{\text{excl}} + \left(1 - f_j^{\text{excl}}\right) p_j^{\text{list}},$$

where $f_j^{\text{excl}} = Q_j^{\text{excl}}/Q_j$. Since rebates are paid only on commercial volume, equation 6 can be rewritten as $p_j^{\text{ASP}} = \text{WAP}_j - r_j \cdot Q_j^{\text{comm}}/Q_j$, which yields

$$r_j \;=\; \frac{\text{WAP}_j - p_j^{\text{ASP}}}{Q_j^{\text{comm}}/Q_j}.$$

Computing the right-hand side requires three auxiliary assumptions: (i) $Q_j^{\text{comm}}/Q_j = 1 - \alpha = 0.67$, so the Medicare share is uniform across facility types; (ii) the stocking shares $(\gamma_1, \gamma_2, \gamma_{12}) = (0.6526, 0.1007, 0.2467)$ and conditional commercial formulary shares $(\lambda_1, \lambda_2, \lambda_{12}) = (0.179, 0.030, 0.791)$ are taken from the data; (iii) the drug-1 share at both-stocking facilities is $s_1^{\text{open}} = 0.559$, applied to both Medicare and open-formulary commercial patients. Then

$$f_j^{\text{excl}} \;=\; \frac{\gamma_j\left[\alpha + (1-\alpha)\left(\lambda_j + \lambda_{12}\right)\right]}{\gamma_j\left[\alpha + (1-\alpha)\left(\lambda_j + \lambda_{12}\right)\right] + \gamma_{12}\, s_j^{\text{open}}},$$

which gives $f_1^{\text{excl}} \approx 0.823$ and $f_2^{\text{excl}} \approx 0.449$, hence $\text{WAP}_1 \approx \$127.72$ and $\text{WAP}_2 \approx \$100.66$. Substituting into the inverted ASP identity yields

$$r_1^* \;\approx\; \frac{127.72 - 110}{0.67} \;\approx\; \$26.45, \qquad r_2^* \;\approx\; \frac{100.66 - 96}{0.67} \;\approx\; \$6.96.$$

Calibration of fixed parameters

- **Insurer hospital payment multiplier (β):** Our estimate of commercial markups comes from the Transparency in Coverage (TiC) legislation, provided to us by Serif Health. TiC requires health plans to publish machine-readable files reporting in-network and out-of-network payment rates for commercial insurers. Serif Health aggregates these large (~1 TB) files and substantially enhances the raw TiC data, which often contain "zombie" rates, which are prices reported for providers that do not actually offer the service. To address this, Serif Health restricts the data to practice–drug pairs where the practice either (1) is associated with a relevant specialty or (2) has an observable claim for the drug in Komodo Health's multi-payer claims data. This approach aligns with prior work using other private vendors such as Clarify Health (Whaley et al., 2025). We calculate β =1.75 by restricting to HCPCS codes for drugs in our biosimilar sample and weighting each by the number of facilities (identified by EIN) prescribing them.

- **Medicare admin fees (ψ_j):** see Section C.2.

- **Dirichlet concentration:** The Dirichlet distribution generates a set of shares that sum to one by drawing a series of independent Gamma random variables and dividing each by their sum. The scaled variant additionally multiplies each draw by a scaling vector prior to normalizing, which centers the distribution around a desired set of proportions. The variance of the distribution affects how aggressively companies compete with one another. A very low variance means that facilities are very homogeneous, generally leading to a winner-take-all equilibrium where the reference product takes all of the market share thanks to its built-in advantage in formulary coverage. However, a variance that is too high implies that only a few facilities are on the margin between the reference product and the biosimilar, eliminating price competition incentives. The value we chose strikes a balance between these two extremes and generates stable equilibria where both firms gain meaningful market share.

Simulated method of moments

The estimation routine simulates the model at a candidate parameter vector, computes the moments described in Table 5 on the resulting equilibrium, and minimizes a weighted sum of squared deviations from the corresponding sample values. Let $\Theta = (\delta_1, \eta, \kappa_1, \kappa_{12}, \sigma_\omega, \zeta_1, \zeta_{12}, \sigma_\nu, \rho)$ denote the nine estimated parameters and let $m_j(\Theta)$ denote the gap between the j-th simulated moment and its sample counterpart. Choice-share moments enter as raw differences; price and rebate moments enter as percentage deviations, so that all moment errors are dimensionless and comparable in magnitude. The estimator solves

$$\hat{\Theta} = \arg\min_{\Theta} \sum_{j=1}^{11} w_j \, m_j(\Theta)^2,$$

where the weights w_j are listed in Table 5.

Simulation. We approximate the integrals over the facility payer-mix distribution $G(\theta)$ by Monte Carlo. At setup, we pre-draw the underlying uniform variates with a fixed seed and convert them into the per-facility Medicare share α_h and the three commercial-mix variates that combine into $(\lambda_{1h}, \lambda_{2h}, \lambda_{12,h})$ once the formulary shares are known. The same draws are reused at every parameter guess, so any change in the objective function reflects a change in the parameters rather than simulation noise. We use 1,000 facility draws.

Inner equilibrium solver. For each candidate Θ, we numerically solve for the equilibrium of the manufacturer pricing game. The unknowns are the four manufacturer instruments $(p_1^{excl}, p_2^{excl}, r_1, r_2)$ together with the two ASPs that close the system in equation 6. We stack these into a six-dimensional vector, expressed as fractions of list prices to keep them on a common numerical scale, and solve for the fixed point of the best-response map by damped iteration. Each step updates the candidate vector toward the best response with a damping factor of 0.5, recomputing the inner formulary

and stocking shares along the way. If the iteration fails to converge to a tight tolerance within several hundred steps, we fall back on a derivative-free optimizer (Nelder-Mead simplex) applied to the squared norm of the best-response gap, warm-started from the last fixed-point iterate. Best responses for each manufacturer are themselves obtained by direct maximization of the profit expression in equation 15, anticipating how the insurer's formulary choices respond to the candidate rebate.

Because we don't know if the game has multiple equilibria, the choice of starting point may matter. We always initialize the inner solver at the observed data moments, which anchors the routine to the equilibrium basin consistent with the calibration target. The same starting point is used in counterfactual simulations to keep the comparisons across scenarios meaningful.

Outer optimization. The objective surface is non-smooth (the inner equilibrium solver is itself iterative) and has multiple local minima, so a single local search from a single starting point would not be reliable. We address this with a multistart strategy.

In a first stage, we evaluate the objective at 3,000 Latin Hypercube Sampling draws of Θ within the parameter bounds. The resulting map of the loss landscape identifies regions of the parameter space where the model fits the data best. In a second stage, we run the same Nelder-Mead optimizer from the eight LHS points with the lowest loss, in parallel. Each worker writes its converged parameter vector to a checkpoint file as soon as it finishes, so partial results survive interruptions. A stall-stopping rule terminates a run when the objective fails to improve by more than 10^{-4} over 50 consecutive iterations, which prevents weak starting points from consuming wall time when a better fit has already been found by another worker. The resulting parameter vector is reported as $\hat{\Theta}$.

C.4 Counterfactual implementation

Each counterfactual takes the calibrated parameter vector $\hat{\Theta}$ as given, modifies a primitive of the model or restricts an agent's choice set, and re-solves the equilibrium of the modified game using the same numerical machinery described in Appendix C.3. Three implementation choices are common to all three scenarios and worth flagging up front.

First, every counterfactual reuses the 1,000 pre-drawn facility variates that were used for estimation, rather than drawing a new sample. Holding the simulation draws fixed eliminates Monte Carlo noise as a source of differences between the baseline and counterfactual equilibria; any variation in the reported quantities reflects only the change in the underlying policy or constraint.

Second, every Nash solve is initialized at the observed data moments. The manufacturer game may admit multiple equilibria so anchoring the solver at the empirical configuration helps keep each counterfactual in the same equilibrium basin as the baseline.

Third, payer spending and the other aggregate quantities reported in Table 7 are computed from the equilibrium objects—prices, rebates, stocking shares, and formulary shares—using the spending decomposition described in Appendix C.2. When a counterfactual shuts down the re-

bate game or restricts the insurer's choice set, the spending formulas adapt mechanically to the modified equilibrium.

We describe the three counterfactuals in the order in which they appear in Section 4.

Increase in the biosimilar admin fee. The first counterfactual mimics the 2022 Inflation Reduction Act adjustment, which raises the biosimilar Medicare admin-fee rate from 6% to 8% of the reference biologic's ASP. We implement this by setting ψ_2 to its post-IRA value and re-solving the standard Nash equilibrium without modifying any other primitive. The admin fee enters both the Medicare-allowed amount on the demand side and the provider receipt on the hospital margin side (see Appendix C.2), so the counterfactual propagates through the consumer logit, the hospital stocking decision, the manufacturer pricing problem, and the ASP identity. It does not enter the manufacturer revenue equation directly, which is the channel that mutes the equilibrium effect discussed in Section 2.

Open insurance formulary mandate. The second counterfactual forces all commercial insurers to adopt an open formulary, so $\Lambda = (0, 0, 1)$ exogenously. Because insurers no longer make an active formulary choice, manufacturers cannot influence formulary placement through rebates, and we set $r_1 = r_2 = 0$. We then re-solve the manufacturer game over only $\left(p_1^{\text{excl}}, p_2^{\text{excl}}\right)$. A specialized version of the inner solver removes the rebate state variables and the formulary fixed point from the system, leaving a two-dimensional manufacturer game closed by the ASP fixed point. The hospital stocking game, the implied ASPs, and the consumer logit are unchanged. Because all facilities now face the same commercial-side formulary distribution, the heterogeneity that drives some facilities toward exclusive stocking flows entirely through the random logit shock (and, to a lesser degree, the draws of α_h), while the commercial formulary-driven component vanishes.

Biosimilar coverage mandate. The third counterfactual removes the exclusive reference-product formulary $F = \{1\}$ from the insurer's choice set. Insurers therefore choose between exclusive biosimilar coverage ($F = \{2\}$) and an open formulary ($F = \{1, 2\}$), with $\lambda_1 = 0$ enforced throughout. Implementation requires re-normalizing the insurer logit in equation 7 over a two-element choice set rather than a three-element one; manufacturers continue to set exclusive prices and rebates, and the hospital stocking game, the consumer logit, and the ASP fixed point are unchanged. We re-solve the modified game using the same damped fixed-point iteration with optimizer fallback as in the baseline. Of the three scenarios, this counterfactual changes the model in the narrowest way—only the insurer side—and is the most direct comparison to the baseline equilibrium.