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

Emma B. Dean, Josh Feng, Luca Maini†

July 29, 2024

Abstract

We document a novel channel through which commercial drug coverage affects Medicare Part B utilization. Exploiting plausibly exogenous differences in state-level commercial coverage of physician-administered drugs, we show that a 10pp increase in exclusion rates among commercial formularies leads to a 1.7pp reduction in market share among Part B beneficiaries. Prescribing behavior of physicians operating across multiple facilities suggests that facilities drive variation in product utilization, likely through the preferred stocking of products with better commercial coverage. Our results suggest that commercial market inefficiencies can affect government spending and support the idea that facilities can directly influence quality of care.

*We would like to thank 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, the International Industrial Organization Conference, the Southern Economic Association meetings, the Southeastern Health Economics Study Group, Utah Applied Economics Conference, Wayne State University, and Weill Cornell Medicine. 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, 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.

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 any decisions that may 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 to receive similar care regardless of payer—a concept called "commonality of care". The policy-relevant implication of this theory is that the incentives facing commercial payers when designing insurance plans can affect the care received by patients covered by a public payer. However, directly testing this implication has been difficult because of the limited availability of exogenous variation in commercial plan design.

In this article, we provide direct evidence that commercial insurance affects the care received by government-sponsored insurance patients 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 covered by different insurance types (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 effects are driven by facilities preferring to stock drug brands with broad coverage among commercial health insurance plans to unlock additional pricing concessions from manufacturers.

Our analysis focuses on the formulary coverage of physician-administered drugs. Formularies are lists of covered drugs and serve as the main tool for health plans and facilities to determine what drugs are available to patients. Two types of formularies are relevant for our study. *Prescription drug formularies* are tiered menus of drugs that determine what is covered by a health plan and at what cost to the patient. *Facility formularies* are lists of drugs stocked and available to physicians at a given healthcare facility. Commercial insurance health plans often employ prescription drug formularies that exclude certain drugs or impose additional access restrictions, such as requiring prior authorization from the insurer before dispensing a given medication (see, e.g., Brot-Goldberg et al., 2023). 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.

Although TM does not impose formulary restrictions, TM beneficiaries are still affected by facility formulary restrictions, which, in turn, may be influenced by commercial formularies. There are at least two reasons facilities may have a limited stock of drugs. First, limiting the number of drugs simplifies logistics such as inventory management and storage space (Dean et al., 2024). Second, even when multiple products are stocked, the facility has an incentive to favor a single brand over all others to unlock volume-based discounts from its manufacturer. 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 creates a spillover effect from commercial formularies to facility formularies. In turn, because facility formularies determine what is administered to commercially insured and Medicare patients, it introduces 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 these spillovers by examining coverage and utilization patterns in two samples of physician-administered drugs. Our "core" sample consists of all branded physician-administered drugs, which we divide into substitution classes using their Anatomical Therapeutic Chemical Level-4 Code (ATC-4), a commonly used standard in studies of drug markets (see, e.g., Dubois and Lasio, 2018; Kakani et al., 2020; Maini and Pammolli, 2023). Despite its widespread use, this classification system is imperfect because drugs can have multiple indications. To ensure that our results are not affected by errors in the market definition, we also study a second, smaller sample of substitution classes defined by an originator biologic product and its biosimilar competitors—clinically equivalent but non-identical alternatives that enter the market after a reference biologic loses patent protection (Food and Drug Administration, 2015). Biosimilars are considered "bioequivalent" to originator 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.[1]

[1]Before the IRA, Medicare paid 6 percent of the originator biologic's Average Sales Price (ASP) for administering either the originator or its cheaper biosimilars. The IRA increased the payment for biosimilars to 8 percent of the originator's ASP to provide an incremental incentive. TM patients generally face a coinsurance, which means that they pay less out of pocket for biosimilars versus the originator.

Our analysis proceeds in two steps. In the first step, we compare variation in the commercial formulary coverage of drugs to their market share among Part B patients at the state-year level and find a significant correlation. A key limitation of this approach is that the correlation could be driven by unobserved variation in state-level patient and physician preferences. To address this challenge, we employ an instrumental variable research design based on variation across states in exposure to *national* formularies. Commercial plans contract with intermediaries known as Pharmacy Benefit Managers (PBMs) to design their formulary and negotiate rebates. 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.[2] 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.

Our instrumental strategy is a shift-share design, where variation over time comes from inclusion and exclusion decisions of national formularies, and cross-sectional variation comes from different penetration of specific national formularies across states. This estimation strategy relies on one assumption and some stickiness in plan behavior. The key assumption for our approach to be valid is that inclusions and exclusions in national formularies cannot be motivated by patient or physician preferences at the state level. This assumption is plausible because PBMs offer the same national formularies to payers across all states, and we provide additional corroborating evidence. We also need a large enough fraction of payers to keep the same national formulary year after year, even if that formulary undergoes small changes, such as the inclusion or exclusion of a specific drug brand. Our data confirms that yearly health plan turnover rates in the MMIT data are relatively low. Intuitively, our empirical approach identifies the effect of formulary coverage by comparing changes in utilization among Part B beneficiaries across states where national formularies that include (or exclude) a given drug have different penetration rates. We deal with inference issues that arise in shift-share instrumental variables designs by following Borusyak et al. (2022).

Our results show that higher brand coverage rates in commercial formularies result in higher utilization rates among Part B beneficiaries. Among drugs in our core sample,

[2] 45.5 percent of individuals covered by commercial insurance are on a plan that adopts a national formulary in 2019.

with substitution classes defined at the ATC-4 level, a 10 percentage point increase in brand coverage by commercial plans leads to a 1.7 percentage 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 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 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 generally on the high end, with facility-level factors explaining between 80–90% of variation in prescribing behavior.

To link our two sets of results, we repeat step one but split the core sample between ATC-4s 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 ATC-4s for which facilities drive more of the variation.

Although the influence of commercial payers on Medicare Part B utilization is not problematic per se, our findings imply that commercial market failures or inefficiencies will also affect Medicare. Our biosimilar sample provides one such example. Consistent with evidence showing that biosimilars have struggled to gain formulary coverage despite competitive pricing (Maini et al., 2021; Frank et al., 2022), we find that Medicare would actually spend less money if commercial formularies stopped excluding drugs altogether.[3] If commercial formularies stopped covering the (more expen-

[3]This is driven by the fact that commercial insurance coverage of reference biologics is much more generous than that of cheaper biosimilars.

sive) originator products altogether, Medicare spending could fall by approximately $220 million a year (based on 2019 numbers), or approximately 40% of the total potential savings from switching all Medicare patients to the cheapest available biosimilar.[4]

The finding that facilities play an essential role in steering physician behavior also matters for policy more broadly. Physicians are not isolated in their practice; rather, they are influenced by their work environment. Our research suggests that changes in physician prescribing behavior could sometimes be more appropriately attributed to facility-level decisions. Recognizing this distinction is crucial because policies aimed at altering physician behavior may prove ineffective when they ignore the restrictions imposed by facilities. For example, recent interventions that attempt to change physician prescribing patterns through letter campaigns have obtained some success (see, e.g., Sacarny et al., 2018; Ahomaki et al., 2020; Schwartz et al., 2021).[5] Our results suggest that similar efforts may be less effective when targeting prescribing of Part B drugs.

Our paper contributes to three strands of the economics literature. The first is the literature on the role played by healthcare facilities in determining treatment. Our primary contribution is to provide 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).[6] Our contribution is to use plausibly exogenous geographic variation in commercial insurance—which is the driver of effects in the Glazer and McGuire (2002) framework—and trace its impacts on TM patients. Our results also speak to the role of a facility's financial incentives in determining drug choice, and the channel we identify may become more relevant with increasing consolidation and private equity involvement in healthcare facilities (see, e.g., Cooper et al., 2019; Bruch et al., 2021; Cerullo et al., 2021; Gupta et al., 2021; La Forgia, 2023). Our biosimilars results also speak to the literature on adoption of medical innovations (see e.g., Skinner and Staiger, 2015) by showing that share-based discounts could serve as a barrier to adoption.

[4] We do not find systematic evidence of excess spending in the core sample, suggesting that the commercial market may function relatively well in general.

[5] Not all letter campaigns are successful: see for example Sacarny et al. (2016).

[6] A related literature, mostly in medical and health policy journals, also studies the effect of changes in payment rates of a single payer type, on health outcomes (often mortality) of all patients, generally finding results that are consistent with the "commonality" hypothesis (see, e.g., Chen et al., 2010). Note that this does not mean, however, that all patients receive the exact same level of care (see, e.g., Langa and Sussman, 1993; Spencer et al., 2013).

We also contribute to the related literature on spillovers between public and commercial health insurance. Most existing work has focused on spillovers from Medicare and Medicaid insurance design to commercial equilibrium outcomes, partly because of the direct policy relevance and the availability of policy changes and natural experiments. This includes work by Finkelstein (2007) and Clemens and Gottlieb (2014) on spillovers from Medicare reimbursement and Duggan and Scott Morton (2006) and Feng et al. (2023) on spillovers from Medicaid drug reimbursement rules. More recent work has focused on capacity constraints (Garthwaite, 2012; Richards and Tello-Trillo, 2019; Gandhi, 2023; Hackmann et al., 2024). Under this channel, increased coverage in one segment can lead to decreased utilization in the other. A final strand of the literature has focused on physician practice styles as the driver of 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, has the opposite prediction to a capacity constraints mechanism, and operates at the facility level rather than the physician level.

Finally, our work 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 formulary coverage changes in the commercial segment. Existing work has relied on identification strategies based on local 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). These approaches, unlike ours, are unlikely to generate significant aggregate variation at the facility or market level. Second, we uncover a potential additional barrier to biosimilar adoption. Past work on biosimilar uptake has highlighted the aggressive response of incumbent biologics (Maini et al., 2021), the role of patent uncertainty (Van de Wiele et al., 2021; Hemphill and Sampat, 2022), the relative safety and efficacy of biosimilar drugs (Cohen et al., 2017; 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 and that most of the variation in U.S. physician-administered biosimilar adoption is due to facilities rather than physicians.

1 Institutional Background and Data

1.1 Commercial Insurance and Facility Drug Formularies

When prescribing drugs, physicians consider clinical concerns such as efficacy, safety, and drug interactions. In the case of drugs administered in an outpatient setting (e.g., hospital outpatient departments and physician offices) in the U.S., physicians also need to account for restrictions created by two types of formularies: the prescription drug formulary of the patient's health plan and the formulary of the facility where the doctor is prescribing.[7]

Prescription drug formularies are tiered menus listing all drugs covered by the health plan. Tiers determine the patient's out-of-pocket cost and whether the plan imposes any non-monetary restrictions on the drug's utilization, such as step therapy or requiring prior authorization before administering the drug. Drugs are assigned to tiers based on negotiations between drug manufacturers and intermediaries called Pharmacy Benefit Managers (PBMs). PBMs negotiate higher rebates in exchange for better tier placement. Often, rebates are contingent on the tier of direct competitors, with manufacturers granting additional concessions when direct competitors are placed in less generous tiers or excluded altogether (Senate Finance Committee, 2021). As a result, it is not uncommon for prescription drug formularies to exclude drugs with close substitutes. After negotiations, PBMs offer formularies to payers such as insurance companies and self-insured employers. Payers can create a custom formulary from the grids of contingent rebates negotiated by the PBM. Alternatively, most PBMs also offer some ready-made "national formularies." Because they require less expertise and effort, national formularies are usually preferred by self-insured employers. As a result, they have a significant influence on the commercial insurance segment, which we exploit for our identification strategy.

Facility formularies determine the drugs that facilities stock and make available to physicians. 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

[7] Due to differences in reimbursement structure, hospitals may have separate inpatient and outpatient formularies. Our analysis focuses on drugs prescribed primarily in outpatient settings and therefore on hospital outpatient formularies.

(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. These contracts offer significant price reductions to facilities that meet a pre-specified usage threshold for a given brand within a substitution class. However, facility formulary exclusions also restrict the set of available options for patients and physicians, which can be problematic for patients whose insurance company also uses a restricted prescription drug formulary.

The interaction between prescription drug formularies and facility formularies can potentially generate a spillover effect from commercial to government-sponsored health insurance, such as Medicare Part B. When facilities are choosing which drugs to stock, they will prioritize those with broad coverage on patients' formularies, 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 Medicare Part B does not have any formulary restrictions, commercial insurance formularies will end up having an outsized impact on the facility formulary choice.

Figure 1 summarizes the key agents involved in the choice of physician-administered drugs and their incentives, with blue arrows highlighting the channels that directly influence Medicare Part B utilization and green arrows highlighting the channels that directly influence commercial utilization. Commercial insurance formularies are separate from the blue part of the system but may affect Medicare Part B utilization through the hypothesized spillover channel, represented by the red arrow.

1.2 Data sources, sample selection, and definition of 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.[8] 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.

[8] 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.

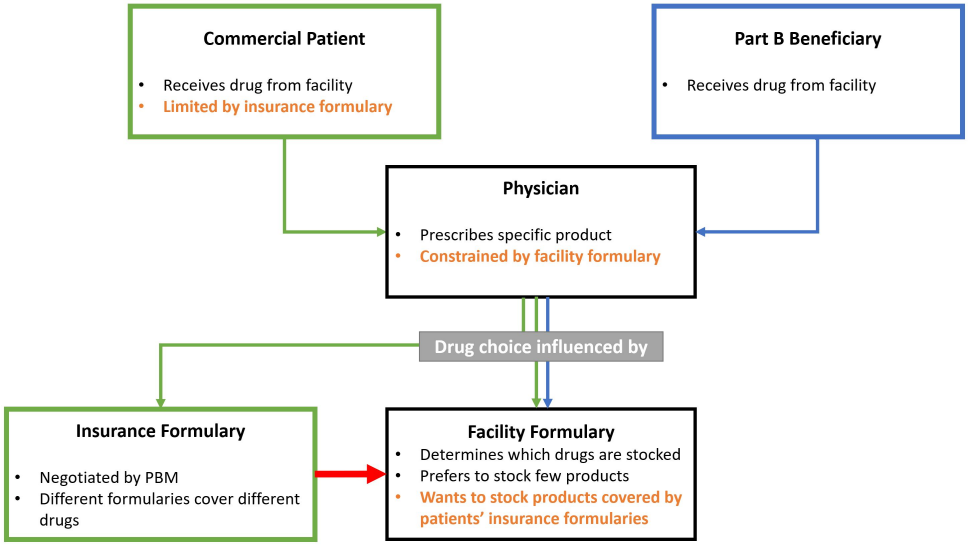


Figure 1: Diagram for impact of hospital formularies

To measure utilization, we use the 20 percent random sample of all Medicare Part B fee-for-service claims data from 2015 to 2019 (Centers for Medicare and Medicaid Services, 2019) from the Centers for Medicaid and Medicare Services (CMS). Physician-administered drugs are identified in claims data 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 to Medicare Part B claims, 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).[9] We also collect data on detailing payments to physicians from OpenPayments data to use as a control in one of our regressions.

Sample definition. Our primary analysis focuses on a broad group of physician-administered drugs. To identify drugs that are primarily administered in physician offices and hospital outpatient departments (i.e., in Medicare Part B) rather than through retail pharmacies (i.e., in Medicare Part D), we rely on publicly available spending data for Medicare Part B and D from the CMS Dashboard. Our initial sample includes branded drugs for which over 95% of Medicare spending comes from Medicare Part B.[10] From this list,

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

[10]We exclude generic drugs because we do not have formulary coverage data for generics. Even if we did, Medicare claims do not distinguish between different versions of the same generic (or branded)

we exclude instances where multiple drugs are assigned to the same Healthcare Common Procedure Coding System (HCPCS) code, which makes it impossible to calculate a drug-specific market share, and drugs 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 drug'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.[11] The ATC is a hierarchical classification system created by the WHO that is commonly used to define drug markets in economic research.[12] 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 not perfect. 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 result in a weakening of 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).

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

molecule, which precludes our measurement approach.

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

[12] 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)
Fraction uncovered lives	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.

these products are close substitutes and ideal candidates for percentage-based and exclusive contracts. Second, crucially for our analysis, biologics and biosimilars are identified through separate HCPCS codes in Medicare claims. Third, our analysis can speak to potential barriers to biosimilar adoption, which has been at the center of an important policy debate, because researchers and policymakers alike see biosimilar adoption as a crucial step towards moderating drug spending.[13]

[13]See for example, Marta Wosińska's interview on Tradeoff on January 26, 2023

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), pegfilgrastim (a bone marrow stimulant, brand name Neulasta), and epoetin alpha (a treatment for anemia, brand name Procrit).[14]

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

Table 1 reports some basic summary statistics on these products. Panel A shows that there are, on average, 4 drugs 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 drugs is about 11 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 the Medicare Part B market share of a given drug 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:

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

(https://tradeoffs.org/2023/01/26/humira-biosimilar-drug-prices/, retrieved June 8, 2023).

[14] 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.

where j denotes brand, k denotes state, and t denotes year. As controls, we include fixed effects for year, λ_t, and the interaction of drug and state, γ_{jk}. 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.

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 state-brand fixed effects γ_{jk} control for differences in the *average* preference for a given drug, 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 control for unobserved changes in state-level preferences over time, we propose an instrumental variable approach based on *national* formularies—stock products designed by PBMs, and available to all U.S. payers. Some payers design *custom* formularies based on the rebates negotiated by PBMs but tailored to enrollee preferences. However, national formularies are particularly popular among self-insured employers—who make up a significant fraction of the commercial insurance market—because designing a custom formulary requires specific medical expertise that most employers lack. Over the 2015–2019 period, 42 percent of commercial patient 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.[15]

The key assumption for our instrument to be valid is that changes in drug coverage on national formularies are independent of shocks to state-level preferences for a given drug, which plausibly holds because PBMs design them for the entire U.S. market. Most of the large national formularies operate in almost all states, and the largest exposure share to any single state averages about 0.16 across all formularies, with the largest typically being California or Texas. However, payers may still react to changes in national formularies by switching to a different national (or custom) formulary. Because this switching behavior would partly reflect shocks to state-level preferences, *current* penetration rate of a national formulary may still be correlated with state-level preference shocks. To ensure that our instrument is uncorrelated with these 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

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

relevance, because, in any given year, 70–90 percent of plans keep the same national formulary as the year before. Intuitively, any change in a national PBM formulary will mechanically affect exclusion rates at the state level in a way that is plausibly unrelated to changes in state-level preferences. Appendix B reports more details on the numbers included in this paragraph, and provides additional evidence in support of the validity of our instrument, including evidence showing that coverage differences across national formularies are only weakly correlated with the distribution of their enrollees across states.

We implement this shift-share style design by following the procedure laid out in Borusyak et al. (2022) (henceforth, BHJ).[16] BHJ show that, under certain conditions, the shift-share instrumental variable (SSIV) regression is equivalent to an instrumental variables regression that includes each shock as a separate instrument. Our setting likely satisfies the theoretical conditions for the BHJ equivalence results because: i) exposure shares to national formularies could be non-random but the changes to national formularies are likely independent of state-level preferences; ii) there are a large number (117) of national formularies (law of large numbers condition); and iii) each state is exposed to a small number of national formulary changes (relevance condition). In addition, we also need to address several additional issues raised in BHJ. First, we have a panel data setting, so we use the recommended first-difference approach described in BHJ in order to allow for changing exposure shares over time. Second, not all plans in each state use a national formulary, so the total exposure shares do not sum to one. We account for this by controlling for the total exposure shares.

Formally, our modified specification is

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

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

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

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

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), $\Delta\mathbb{I}^{excluded}{}_{jft}$ reflects whether national formulary f changed how it covered drug 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 drugs across national formularies.[17] Overall, we identify 381 exclusion events and 727 inclusion events across all products in our core sample, and 30 exclusion events and 212 inclusion 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 basic panel into an equivalent national formulary-brand-year level dataset.[18] 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.[19]

2.2 Results for the core sample

We report our results 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

[17] 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, whereas a third PBM (Optum Rx) did not. New Jersey had significantly higher exposure to Optum Rx's national plans than Connecticut in 2018. Both states had the same coverage rate of Udenyca in 2018 (44 percent), but, partly due to the difference in exposures, ended up with a 6pp difference in total commercial coverage rates in 2019 (90 percent for Connecticut and 84 percent for New Jersey).

[18] We implement this transformation using the *ssaggregate* package associated with BHJ.

[19] 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.

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

	(1)	(2)	(3)	(4)
PANEL A. OLS ESTIMATES				
Fraction Excluded	-0.091	-0.091	-0.115	-0.094
	(0.011)	(0.011)	(0.009)	(0.010)
Specification	OLS	OLS	OLS	OLS
PANEL B. SHIFT-SHARE IV RESULTS				
Fraction Excluded	-0.174	-0.174	-0.173	-0.170
	(0.030)	(0.030)	(0.031)	(0.030)
Specification	2SLS	2SLS	2SLS	2SLS
First-Stage *F*-stat.	418.4	418.3	388.9	369.1
PANEL C. IV FIRST STAGE				
Uncovered on National Formulary	0.235	0.235	0.234	0.235
	(0.012)	(0.012)	(0.012)	(0.012)
Specification	OLS	OLS	OLS	OLS
R^2	0.330	0.330	0.321	0.330
Observations	16,114	16,114	16,114	16,114
Open Payments Control	No	Yes	Yes	Yes
State Population Weights	No	No	Yes	No
Log Market Size Weights	No	No	No	Yes

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. Column 3 includes regressions weighted by state population size, whereas column 4 includes regressions weighted by log ATC-4 market size, measured as the sum of annual beneficiary service days per ATC-4. Observations indicates number of drug-year-state observations.

sufficient power. The reduced-form regression 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 (Column 2) and weight results by state population (Column 3) or log market size (measured as the sum of annual beneficiary service days at the ATC-4 level, Column 4) return almost identical coefficients.

The instrumental variables estimates are economically significant. As an example, assume that there are three competing drugs with equal market share and each earns 50 percent of its revenue from Medicare. If a drug were excluded by a large PBM—leading to a 25pp decrease in commercial coverage—our estimates imply that it would lose about $-0.174 \times 0.25 = 0.0435$ fraction of TM market share, equivalent to a 6.5 percent loss in revenue *just from the spillover to Medicare Part B*. The IV estimates are larger in magnitude than the OLS estimates, which could reflect commercial payers unobservably 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.

Appendix C reports results from additional robustness checks, including: (i) traditional shift-share IV methodology; (ii) different functional forms of market size weights; (iii) adding controls for product age; (iv) including drug-year fixed effects; and (v) excluding the four most populous states, whose preferences could be targeted in the design of national formularies (California, Texas, Florida, and New York). All of these checks confirm our results, although the magnitude of the effects varies slightly across specifications.[20]

2.3 Results for the biosimilar sample

The analysis of the core sample shows that commercial insurance formularies impact the brands Part B recipients receive, with this spillover applying broadly across physician-administered drugs. However, our market definitions may be misspecified, because ATC-4s may be broader than substitution classes used in actual contracts. To provide a more precise estimate of the magnitude of this spillover within a well-defined substitution class, we run a secondary analysis focusing on biologic drugs with biosimilar competitors, with markets defined by the molecule of the originator biologic. As noted earlier, results from the biosimilars sample also speak to potential barriers to adoption and government policies aimed at encouraging biosimilars adoption in Medicare.

Results are shown in Table 3. Estimates are qualitatively consistent with our main results in 2, but the point estimates are larger in magnitude, suggesting that our main

[20] Including drug-year fixed effects and age controls leads to a meaningful reduction in the magnitude of our coefficient. We believe this is to be expected, as variation in coverage across years and over the lifecycle of drugs helps us identify our coefficient of interest, especially in the biosimilar sample, because newly launched biosimilars tend to obtain better coverage over time.

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

	(1)	(2)	(3)	(4)
PANEL A. OLS ESTIMATES				
Fraction Excluded	-0.377	-0.375	-0.393	-0.386
	(0.023)	(0.023)	(0.022)	(0.022)
Specification	OLS	OLS	OLS	OLS
PANEL B. SHIFT-SHARE IV RESULTS				
Fraction Excluded	-0.301	-0.300	-0.295	-0.303
	(0.033)	(0.033)	(0.035)	(0.034)
Specification	2SLS	2SLS	2SLS	2SLS
First-Stage *F*-stat.	526.1	539.5	490.3	526.4
PANEL C. IV FIRST STAGE				
Uncovered on National Formulary	0.262	0.262	0.259	0.260
	(0.011)	(0.011)	(0.012)	(0.011)
Specification	OLS	OLS	OLS	OLS
R^2	0.417	0.418	0.408	0.415
Observations	2,142	2,142	2,142	2,142
Open Payments Control	No	Yes	Yes	Yes
Population Weights	No	No	Yes	No
Log Market Size Weights	No	No	No	Yes

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. Column 3 includes regressions weighted by state population size, whereas column 4 includes regressions weighted by log molecule market size, measured as the sum of annual beneficiary service days per molecule. Observations indicates number of brand-year-state observations.

estimates may be biased downward by misspecifications in the definition of substitution classes. As with the broader Part B results, our first stage estimates in Panel C show a very strong association between our instrument and brand coverage rates. Our reduced-form regression suggests that a 10pp increase in the exclusion rate of a given brand is associated with a 3.7pp decline in utilization among Medicare Part B beneficiaries (Column 1), whereas our IV estimates find a smaller effect (3.0pp). Robustness checks controlling for marketing payments to physicians (Column 2) and weighting results by state population (Column 3) or log market size (Column 4) return nearly iden-

tical coefficients. We report additional robustness checks in Appendix C.

3 The Role Played by Facilities and its Implications

Section 2 establishes a causal link between commercial formulary coverage and market share in Medicare Part B. In this section, we investigate potential mechanisms behind this link and discuss its implications for policy.

3.1 Mechanism: the role of physicians versus facilities

Our conjecture is that facilities are the main drivers of this spillover effect, through their choice of facility formulary. 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, purchase of machinery, and other 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.

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).[21] 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

[21] 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.

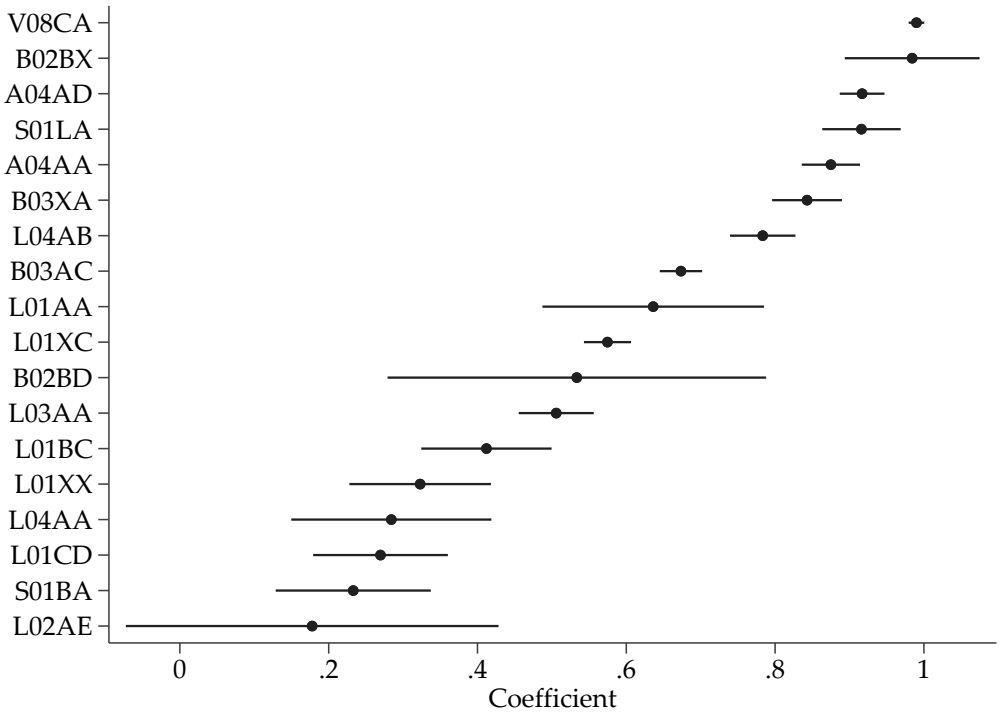


Figure 2: Change in Physician Brand Prescribing Across Facilities

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.

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

Core Sample Results Figure 2 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 releas-

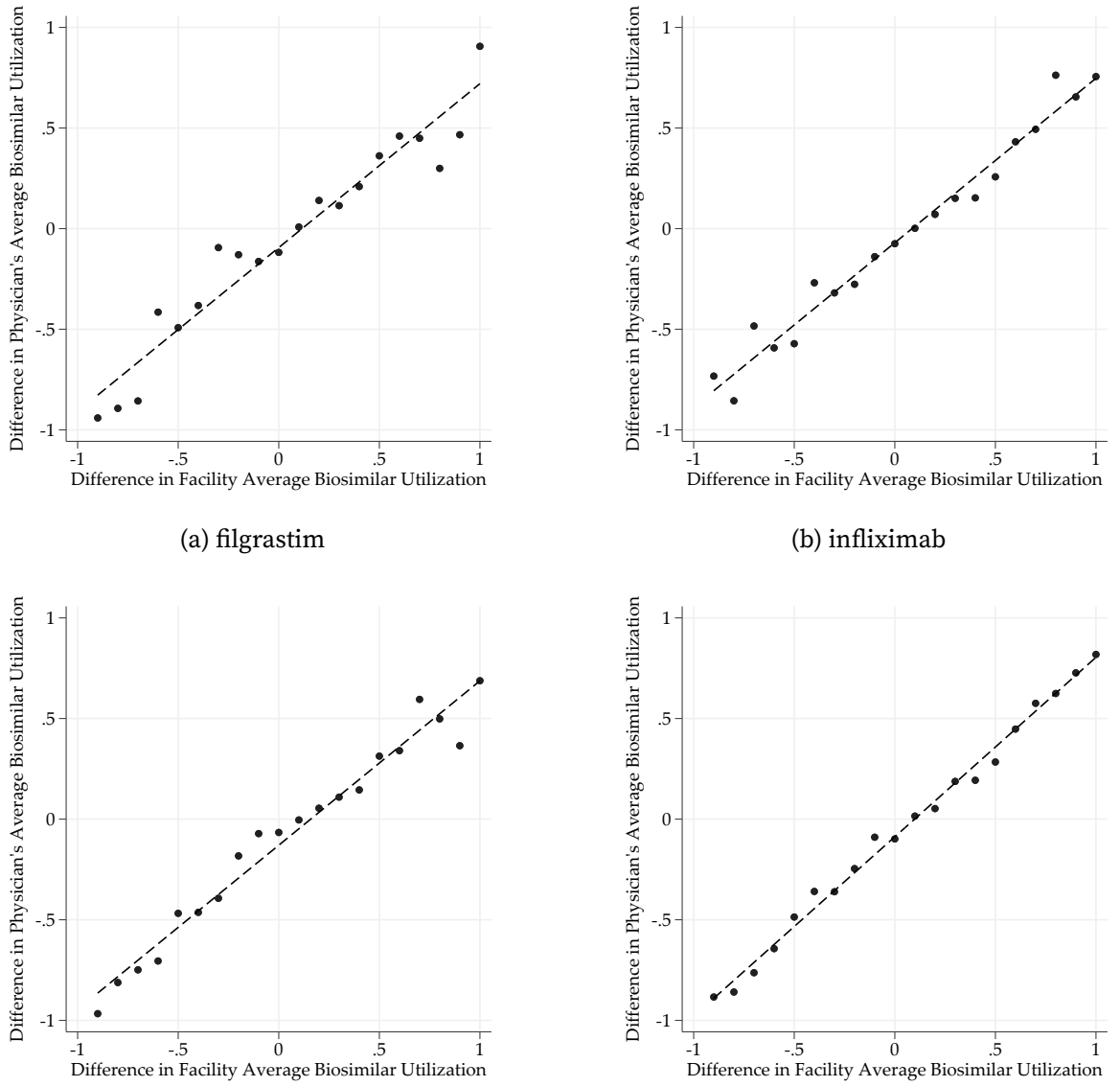
ing hormone analogues," which includes drugs used to treat prostate cancer, but that also have on- and off-label indications in breast cancer, endometriosis, infertility, and precocious puberty). Conversely, higher slopes are associated with more narrowly defined categories, such as A04AD and A04AA (two 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 category in Appendix C.2, and report additional statistics about facility product-mix in Appendix C.3.

Biosimilar sample results Figure 3 plots our results separately by molecule for our biosimilar sample. Our variable of interest is the market share of the reference biologic brands within our four molecules of interest. 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% confidence interval: 0.739–0.827) for infliximab, 0.768 (95% confidence interval: 0.719–0.817) for pegfilgrastim, and 0.843 (95% confidence interval: 0.796–0.890) for epoetin alpha. Hence, across all molecules, physicians adapt their prescription patterns to adhere very closely to the choices of other physicians at the same facility.

3.2 Linking facility effects to coverage and utilization changes

Our original hypothesis is that the link between commercial coverage and Medicare Part B utilization is driven by facility-level decisions. The results in the previous section shows that facilities *do* influence prescription choices, but do not tie this evidence directly to the main analysis in Section 2. In this section, we show that the magnitude of the facility effect we just estimated (that is, γ_c) is correlated with the magnitude of the effect of 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.

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 60 percent smaller than our main estimate. Conversely, among drugs in substitution classes where facilities exert a stronger influ-



(a) filgrastim

(b) infliximab

(c) pegfilgrastim

(d) epoetin alfa

Figure 3: 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.

ence on prescription drug choices, the effect is approximately double our main estimate.

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	-0.068	-0.059	-0.062
	(0.011)	(0.011)	(0.010)	(0.011)
Specification	2SLS	2SLS	2SLS	2SLS
First-Stage *F*-stat.	223.8	223.8	210.5	197.7
Observations	10,208	10,208	10,208	10,208
Panel B. Shift-Share IV results (Above Median)				
Fraction Excluded	-0.359	-0.359	-0.377	-0.374
	(0.067)	(0.067)	(0.071)	(0.068)
Specification	2SLS	2SLS	2SLS	2SLS
First-Stage *F*-stat.	213.0	213.0	197.4	201.6
Observations	4,352	4,352	4,352	4,352
Open Payments Control	No	Yes	Yes	Yes
Population Weights	No	No	Yes	No
Log Market Size Weights	No	No	No	Yes

Notes: Estimates of Equation 2 run on two different groups of core sample drugs separated based on whether the coefficient from Figure 2 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. Column 3 includes regressions weighted by state population size, whereas column 4 includes regressions weighted by log molecule market size, measured as the sum of annual beneficiary service days per ATC-4. Observations indicates number of brand-year-state observations.

3.3 Implications for the Medicare program

In principle, linking Medicare Part B utilization to commercial coverage is not obviously problematic. In a well-functioning market, more expensive products would be more likely to be excluded from commercial formularies, which would help reduce Medicare spending—relative to, for example, a counterfactual world with open formularies. However, the presence of a spillover effect also implies that any failures or inefficiencies in the commercial market will affect the Medicare program.

Evidence suggests that coverage and utilization of biosimilar products may represent one such failure. Evidence suggests that biosimilar uptake and coverage has lagged in commercial markets, despite the fact that virtually all biosimilars are significantly cheaper than originator biologics, even after accounting for rebates (Frank et al., 2022; Maini et al., 2021). Our data confirms these results (see, for example, the numbers reported in Table 1): we find that originator biologics are less likely to be excluded and have higher market share than their biosimilar competitors.

Although calculating the exact impact of the spillover effect would require more detailed data and several additional assumptions, we can obtain an approximate value of the money at stake using a simple back of the envelope calculation simulating what would happen if commercial formularies made different coverage choices. To perform the calculation we take the difference in the average cost of a claim between each reference biologic and biosimilar brand. Then, using the estimated coefficient from our preferred specification in Column (1) of Table 2, we estimate the change in utilization as $\Delta Q_{jk} = \hat{\beta} \times \Delta Coverage_{jk} \times Q_{jk}$, where Q_{jk} is the number of Part B claims for biosimilar j in state k and $\Delta Coverage_{jk}$ is change in coverage under our counterfactual assumptions. We also assume that the additional claims for biosimilar j come from the other brands proportionally to their observed market share.[22]

Using this method, we perform two simulations. First, we calculate spending if all commercial formularies covered both the originator biologic and all biosimilars. Under this scenario, we find that spending would actually *decrease* (by approximately $8 million USD), thus confirming that, in the current equilibrium, more expensive drugs have better coverage. Second, we calculate spending if commercial plans were to completely exclude the reference biologic and instead fully cover all the biosimilar alternatives. We find potential excess Medicare Part B spending of up to $220 million in 2019—approximately 3 percent of overall spending on our sample of molecules and 40

[22] See Appendix D for more details.

percent of the total potential savings from switching all patients to the cheapest available biosimilar.

Performing similar calculations for the broader set of Part B products is more challenging because substitution patterns within a substitution class defined by an ATC-4 code are likely more complex than substitution patterns of near-identical biosimilar products who all share the same molecule. Moreover, any assumption we make will be more consequential in the core sample, where it is not uncommon for products in the same ATC-4 to have very different prices, than in the biosimilar sample, where within-molecule prices are relatively comparable. That being said, replicating the procedure outlined above shows that Medicare Part B spending would increase by 7.9 percent under fully open formularies ($4 billion off of a base of $50 billion), although most of that effect is concentrated in two broad classes of oncology drugs. Conversely, fully excluding the most expensive product in each class (while leaving coverage for all other products the same) would lower spending by approximately 5 percent.

4 Conclusion

This article contributes causal evidence for one of key implications of the care commonality hypothesis first theorized by Glazer and McGuire (2002). We demonstrate the existence of a quantitatively significant spillover from commercial insurance coverage restrictions to Medicare Part B utilization, and provide further evidence that facility stocking decisions drive this effect.

Our results raise three important issues. First, we provide concrete evidence that commercial health insurers can influence the care received by publicly-sponsored insurance enrollees. Unlike the reverse effects, the impact of private markets on government-sponsored markets is harder to pin down because equilibrium outcomes in the former are potentially endogenous to many variables that researchers do not generally observe, such as patient and provider preferences. Although evidence of the impact of government on private markets is more abundant, this should not be interpreted as a sign that these effects dominate. Rather, it is likely a function of the fact that such evidence is easier to cleanly identify.

Second, we highlight the role of facilities as a significant player in determining care. Our finding that physician behavior changes drastically across facilities shows that choice of drugs can be influenced not only by patient and provider preferences, but also by the environment where a physician practices. This in turn has implications for

models of drug choice, which traditionally focus solely on patient and physician incentives. More broadly, the mechanism identified in this article could apply in a number of healthcare delivery settings. For example, Clemens and Gottlieb (2014) find that Medicare payment reform affects technology adoption (specifically, the acquisition of MRI machines), and Hackmann (2019) finds that an increase in Medicaid reimbursement leads to staffing increases in nursing homes. Although government insurance may have driven the facility-level changes found in previous work, our analysis suggests that private insurance changes could also have similar effects, which would change the technology and staffing facing patients covered by government insurance.

Third, we provide proof of concept for a broadly applicable instrumental variable strategy that can be used to analyze the effect of formulary coverage decisions at the market level. Our method addresses the inherent endogeneity of formulary design by leveraging changes in national-level products that are popular across the commercial segment. Such an instrument could be potentially used to address endogeneity concerns to study how demand, access, and health outcomes respond to insurance formulary changes.

References

Ahomaki, Iiro, Visa Pitkanen, Aarni Soppi, and Leena Saastamoinen, "Impact of a physician-targeted letter on opioid prescribing," *Journal of Health Economics*, July 2020, *72*, 102344.

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, Peter Hull, 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.

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

Bruch, Joseph, Dan Zeltzer, and Zirui Song, "Characteristics of Private Equity-Owned Hospitals in 2018," *Annals of Internal Medicine*, February 2021, *174* (2), 277–279.

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

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

Cerullo, Marcelo, Kelly Kaili Yang, James Roberts, Ryan C. McDevitt, and Anaeze C. Offodile, "Private Equity Acquisition And Responsiveness To Service-Line Profitability At Short-Term Acute Care Hospitals," *Health Affairs*, November 2021, *40* (11), 1697–1705.

Chandra, Amitabh, Evan Flack, and Ziad Obermeyer, "The Health Costs of Cost Sharing," *The Quarterly Journal of Economics*, May 2024, *forthcoming*.

Chen, Hsueh-Fen, Gloria J. Bazzoli, David W. Harless, and Jan P. Clement, "Is Quality of Cardiac Hospital Care a Public or Private Good?," *Medical Care*, November 2010, *48* (11), 999.

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.

Cohen, Hillel, Donna Beydoun, David Chien, Tracy Lessor, Dorothy McCabe, Michael Muenzberg, Robert Popovian, and Jonathan Uy, "Awareness, knowledge, and perceptions of biosimilars among specialty physicians," *Advances in Therapy.*, January 2017, *33* (12), 2160–2172.

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.

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.

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.

Feng, Josh, 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.

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, "Scientific Considerations in Demonstrating Biosimilarity to a Reference Product: Guidance for Industry," Technical Report 2015.

_ , "Biosimilar and Interchangeable Biologics: More Treatment Choices," Technical Report 2021.

Forgia, Ambar La, "The Impact of Management on Clinical Performance: Evidence from Physician Practice Management Companies," *Management Science*, August 2023, *69* (8), 4646–4667.

Frank, Richard G., Mahnum Shahzad, Aaron S. Kesselheim, and William Feldman, "Biosimilar competition: Early learning," *Health Economics*, 2022, *31* (4), 647–663.

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

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.

Gupta, Atul, Sabrina T. Howell, Constantine Yannelis, and Abhinav Gupta, "Owner Incentives and Performance in Healthcare: Private Equity Investment in Nursing Homes," February 2021.

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.

Kakani, Pragya, 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.

Langa, K. M. and E. J. Sussman, "The effect of cost-containment policies on rates of coronary revascularization in California," *The New England Journal of Medicine*, December 1993, *329* (24), 1784–1789.

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.

_ , Josh Feng, Thomas Hwang, and Jacob Klimek, "Biosimilar Entry and the Pricing of Biologic Drugs," Technical Report 2021.

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.

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.

Sacarny, Adam, David Yokum, Amy Finkelstein, and Shantanu Agrawal, "Medicare Letters To Curb Overprescribing Of Controlled Substances Had No Detectable Effect On Providers," *Health Affairs*, March 2016, *35* (3), 471–479.

_ , Michael L. Barnett, Jackson Le, Frank Tetkoski, David Yokum, and Shantanu Agrawal, "Effect of Peer Comparison Letters for High-Volume Primary Care Prescribers of Quetiapine in Older and Disabled Adults: A Randomized Clinical Trial," *JAMA Psychiatry*, October 2018, *75* (10), 1003–1011.

Schwartz, Kevin L., Noah Ivers, Bradley J. Langford, Monica Taljaard, Drew Neish, Kevin A. Brown, Valerie Leung, Nick Daneman, Javed Alloo, Michael Silverman, Emily Shing, Jeremy M. Grimshaw, Jerome A. Leis, Julie H. C. Wu, and Gary Garber, "Effect of Antibiotic-Prescribing Feedback to High-Volume Primary Care Physicians on Number of Antibiotic Prescriptions: A Randomized Clinical Trial," *JAMA Internal Medicine*, September 2021, *181* (9), 1165–1173.

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

Skinner, Jonathan and Douglas Staiger, "Technology Diffusion and Productivity Growth in Health Care," *The Review of Economics and Statistics*, December 2015, *97* (5), 951–964.

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.

Spencer, Christine S., Darrell J. Gaskin, and Eric T. Roberts, "The Quality Of Care Delivered To Patients Within The Same Hospital Varies By Insurance Type," *Health Affairs*, October 2013, *32* (10), 1731–1739.

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), E668–678.

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 without 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).[23] 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, 2.60 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

[23]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.

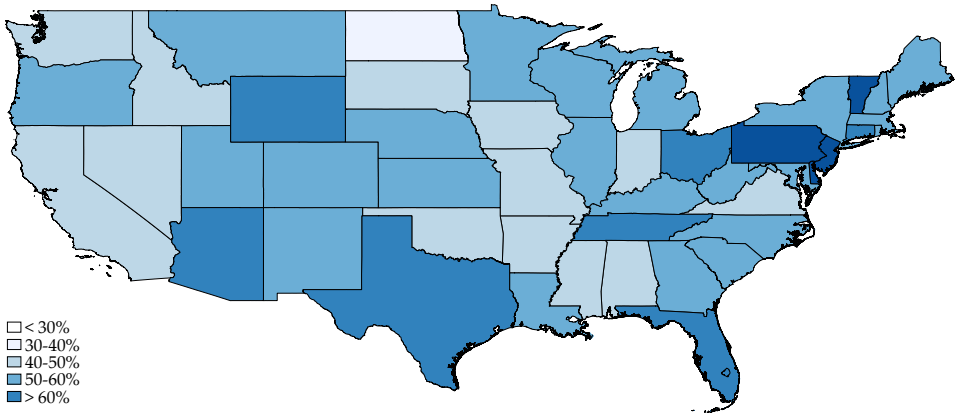


Figure A1: Uptake of national formularies among commercial insurance enrollees, 2019

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.

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 45.6 percent.

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 maintaining 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. The numbers show that the vast majority of plans who exist in both years stick with the same national formulary.

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	3.61	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").

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 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 national formularies: (i) a measure of state-share similarity, which we calculate as the Euclidian distance between the vector of market shares; and (ii) a measure of coverage similarity, which we calculate as the total number of drugs that are excluded by a formulary and covered by the other.

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.

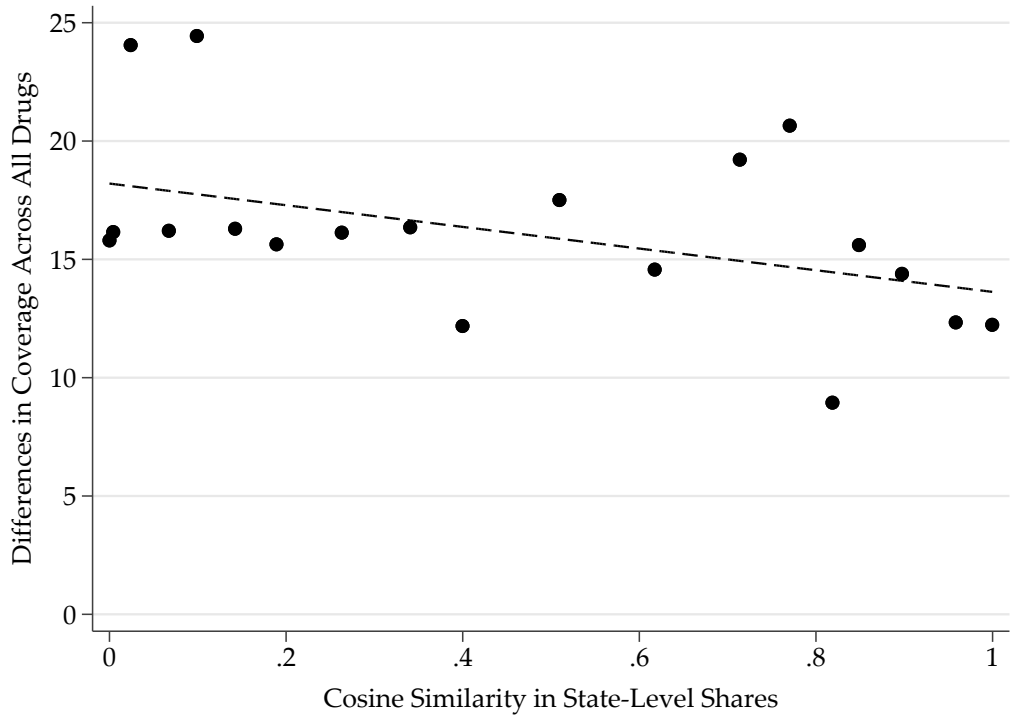


Figure A2: Correlation of plan design similarity and enrollee similarity across pairs of national formularies

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 ("Cosine Similarity") and differences in their coverage of the 104 drugs in our core sample. We then provide a binscatter of the correlation between the two quantities.

C Additional Results and Robustness Checks

In this section, we present a few additional results and robustness checks for our main specifications.

C.1 Alternative specifications for regression 2

We present results from a set of alternative specifications for regression 2 (Tables A7 and A8). We present results using the traditional SSIV 2SLS approach, using different functional forms for market size weights, adding controlling for life-cycle effects using either a second-order polynomial function or drug-year fixed effects, 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 confirm our main analysis, with two minor caveats: using less concave functional forms for market size as weights leads to smaller estimated coefficients, likely because the effects are driven by smaller, more precisely measured substitution classes. Adding controls for life-cycle effects also returns smaller coefficients, because these controls capture some of the variation that helps us identify the coefficient of interest: as newer products enter the market, their coverage usually improves over time, and this variation is helpful in pinning down the spillover effect on Medicare utilization. This is especially useful in the context of our biosimilar sample as biosimilars tend to have low coverage rates early on.

C.2 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 A3–A5).

C.3 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 A6). 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.[24] 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 2 in the main text). Within our biosimilar sample (Panel B), about three quarters 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.

[24] 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.

Table A7: Robustness Checks on Regression 2, core Sample

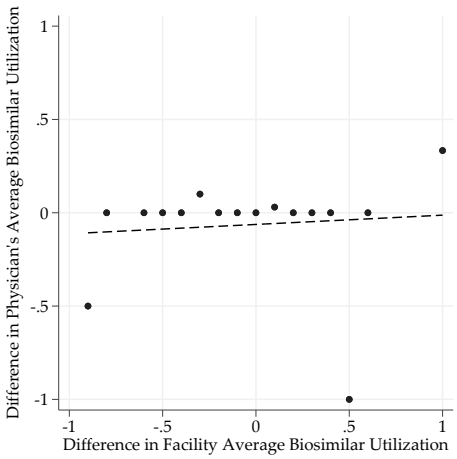
	Traditional SSIV	Sq. Root Mkt. Size Weights	Linear Mkt. Size Weights	Age controls	Drug-Year FE	Exclusion of large states
PANEL A. OLS ESTIMATES						
Fraction Excluded	-0.091	-0.082	-0.056	-0.018	-0.014	-0.087
	(0.011)	(0.008)	(0.008)	(0.010)	(0.010)	(0.012)
Specification	OLS	OLS	OLS	OLS	OLS	OLS
PANEL B. SHIFT-SHARE IV RESULTS						
Fraction Excluded	-0.226	-0.135	-0.087	-0.091	-0.093	-0.174
	(0.022)	(0.026)	(0.019)	(0.034)	(0.033)	(0.029)
Specification	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
First-Stage *F*-stat.	1804.1	226.3	133.9	398.3	129.9	417.5
PANEL C. IV FIRST STAGE						
Uncovered on	1.262	0.233	0.230	0.202	0.019	0.237
National Formulary	(0.0297)	(0.015)	(0.020)	(0.010)	(0.002)	(0.012)
Specification	OLS	OLS	OLS	OLS	OLS	OLS
R^2	0.760	0.326	0.319	0.289	0.0973	0.332
Observations	16,114	16,114	16,114	16,114	16,104	14,594
Detailing Control	No	Yes	Yes	Yes	Yes	No
Population Weights	No	No	No	No	No	No
Mkt. Size Weights	No	Sq. root	Linear	No	No	No

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. Column 3 includes regressions weighted by state population size while column 4 includes regressions weighted by ATC-4 market size. The number of observations in column 5 (specification with Drug-Year FEs) is lower than in columns 1–4 because the Stata command *reghdfe* excludes singleton observations from the regression, leading to the exclusion of drugs who are in the data for multiple states, but only a single year.

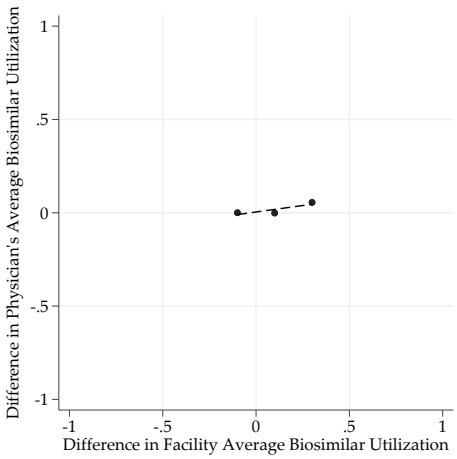
Table A8: Robustness Checks on Regression 2, Biosimilar Sample

	Traditional SSIV	Sq. Root Mkt. Size Weights	Linear Mkt. Size Weights	Age controls	Drug-Year FE	Exclusion of large states
PANEL A. OLS ESTIMATES						
Fraction Excluded	-0.377	-0.421	-0.486	-0.147	0.005	-0.373
	(0.023)	(0.024)	(0.036)	(0.018)	(0.034)	(0.025)
Specification	OLS	OLS	OLS	OLS	OLS	OLS
PANEL B. SHIFT-SHARE IV RESULTS						
Fraction Excluded	-0.343	-0.311	-0.339	-0.041	-0.093	-0.303
	(0.030)	(0.037)	(0.040)	(0.035)	(0.078)	(0.033)
Specification	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
First-Stage *F*-stat.	837.7	518.3	535.8	316.0	39.28	519.9
PANEL C. IV FIRST STAGE						
Uncovered on	1.418	0.255	0.248	0.193	0.015	0.263
National Formulary	(0.049)	(0.011)	(0.011)	(0.011)	(0.002)	(0.012)
Specification	OLS	OLS	OLS	OLS	OLS	OLS
R^2	0.884	0.408	0.398	0.332	0.149	0.419
Observations	2,142	2,142	2,142	2,142	2,142	1,962
Detailing Control	No	Yes	Yes	Yes	Yes	No
Population Weights	No	No	No	No	No	No
Mkt. Size Weights	No	Sq. root	Linear	No	No	No

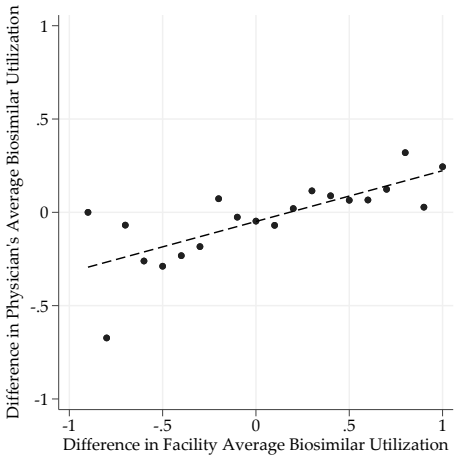
Notes: Estimates of Equation 2, using OLS and IV approaches. The data is at the state-year-brand level and include data from 2015 to 2019. The dependent variable is the brand-level market share for each brand in a given state-year. Standard errors are clustered at the state-brand level. All brands are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



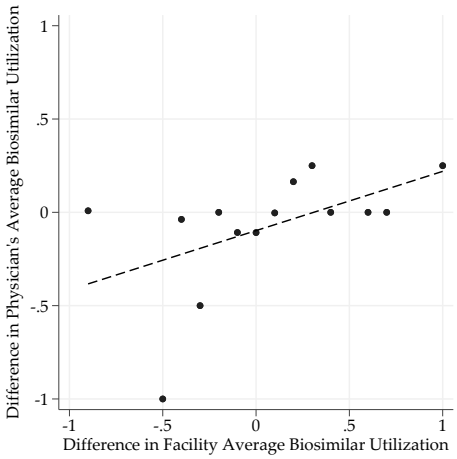
(a) L02AE



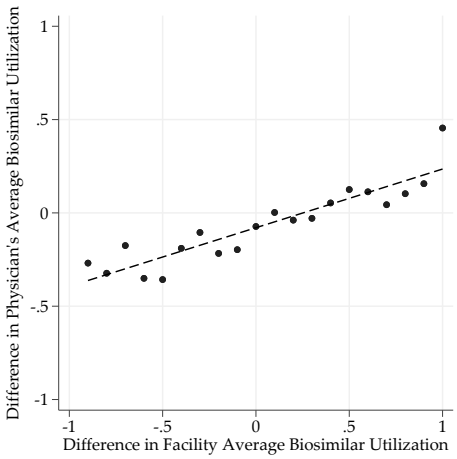
(b) S01BA



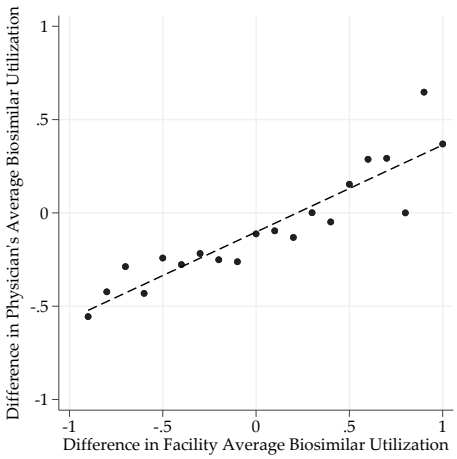
(c) L01CD



(d) L04AA

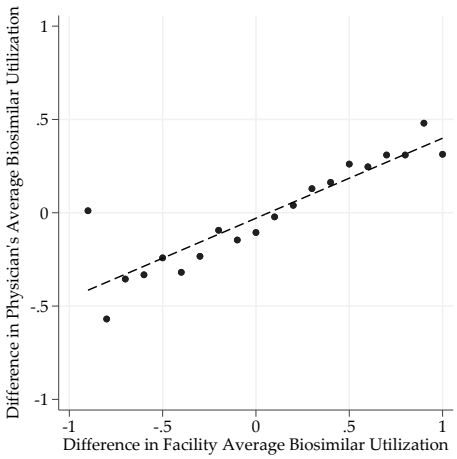


(e) L01XX

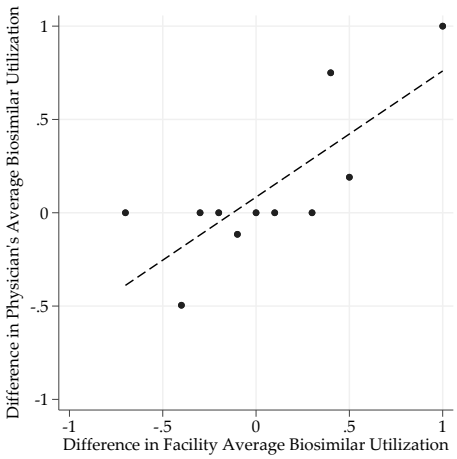


(f) L01BC

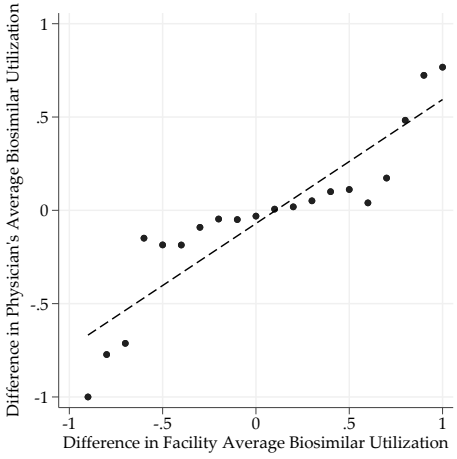
Figure A3: ATC-4 Plots, Part 1



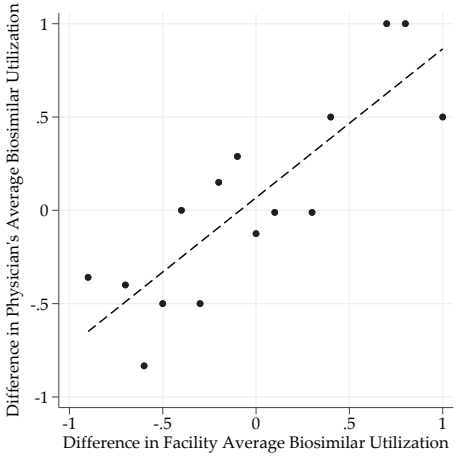
(a) L03AA



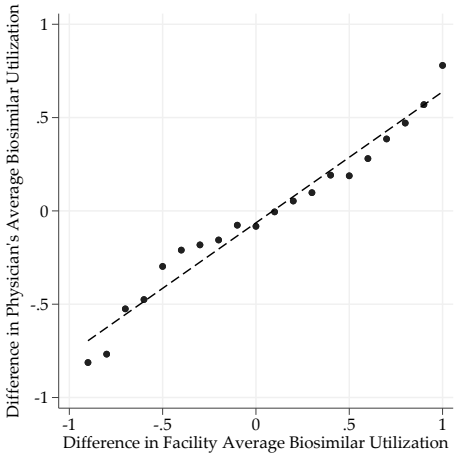
(b) B02BD



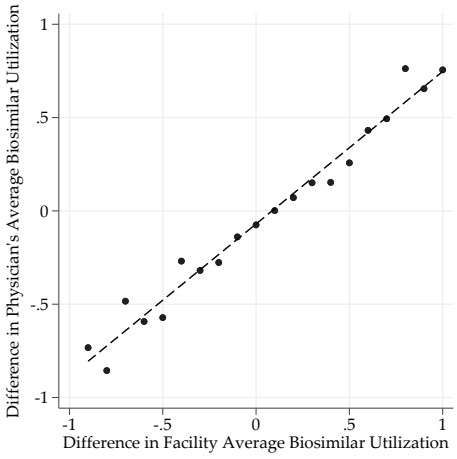
(c) L01XC



(d) L01AA

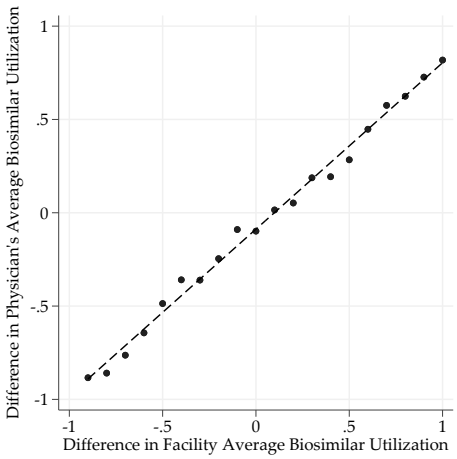


(e) B03AC

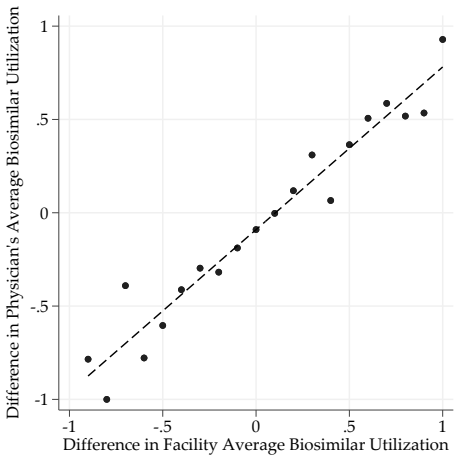


(f) L04AB

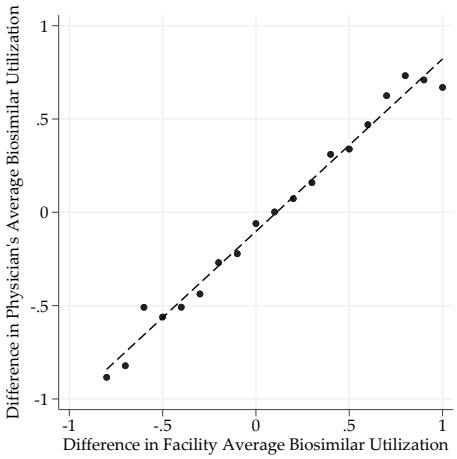
Figure A4: ATC-4 Plots, Part 2



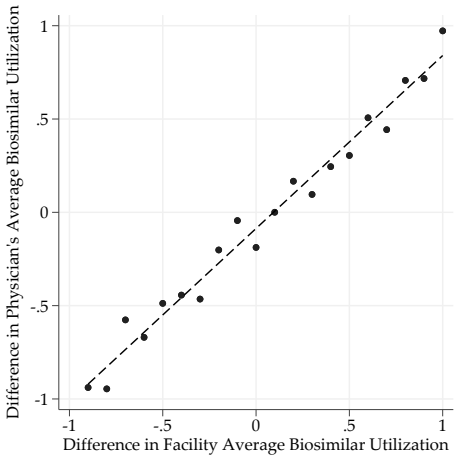
(a) B03XA



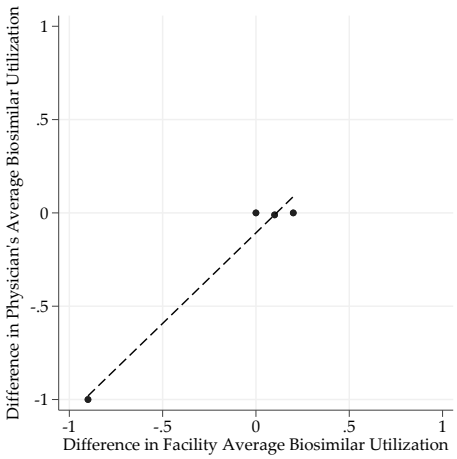
(b) A04AA



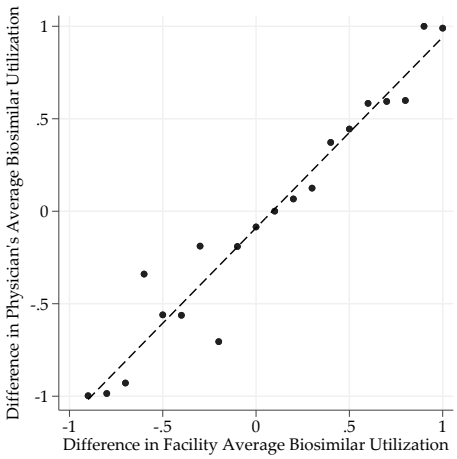
(c) S01LA



(d) A04AD

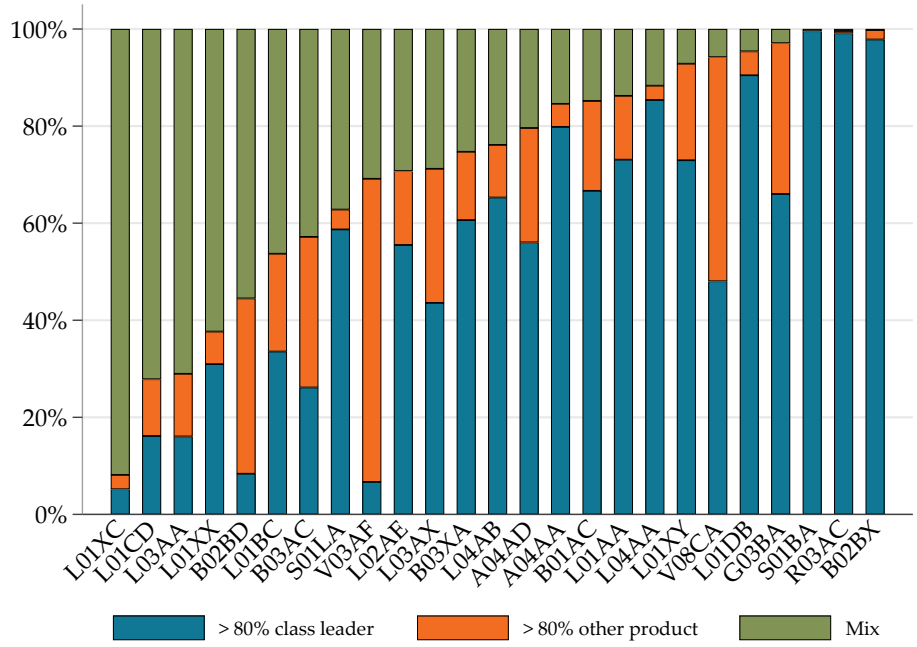


(e) B02BX

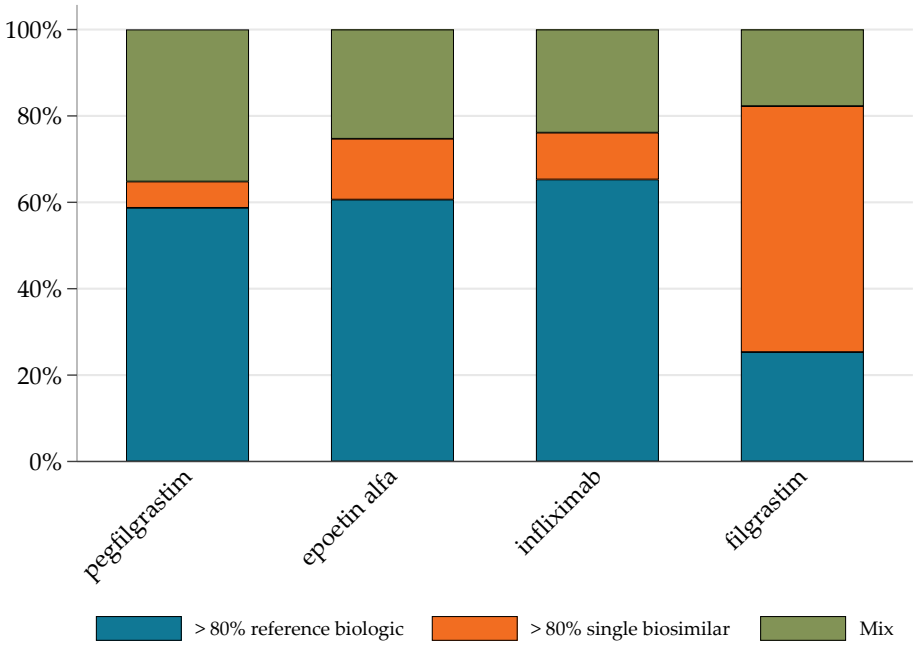


(f) V08CA

Figure A5: ATC-4 Plots, Part 3



(a) Core Sample



(b) Biosimilar sample

Figure A6: 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.

D Calculation of savings from unrestricted formularies

In the main paper, we estimate potential Medicare Part B savings if all commercial formularies covered all biosimilar brands. We perform this calculation in three steps.

1. First, we compute the change in coverage relative to the counterfactual of interest for all drugs and state, which we denote as $\Delta\text{Frac. Uncovered}_{jk}$.

2. Then, we use the estimated coefficient in our preferred specification, (e.g., $\hat{\beta} = -0.173$ from Table 2, column 1) and calculate the predicted change in utilization rate (e.g. market share) for each brand j and state k as $\hat{\beta} \times \Delta\text{Frac. Uncovered}_{jk}$ (note that all data refers to the year 2019)

3. Because first-order changes in market share will generally not sum up to 1, we need to adjust the number from step 1 to account for the fact that some of the increased or decreased utilization of a given brand j may affect other brands. To make this adjustment, we assume that substitution patterns are proportional to observed market shares (this is roughly what would happen with logit demand). The table below illustrates the adjustment with a numerical example

	Drug A	Drug B	Drug C
Initial market share	70%	20%	10%
$\Delta\text{Frac. Uncovered}_{jk}$	−10%	20%	30%
Raw change in market share	−1.7%	3.4%	5.1%
Drug A adjustment		1.13%	0.57%
Drug B adjustment	−2.98%		−0.42%
Drug C adjustment	−3.97%	−1.13%	
New adjusted market share	61.4%	23.4%	15.2%

In the example, Drug A, would experience a decrease in market share of ~1.7%, Drug B an increase of about 3.4%, and Drug C an increase by 5.1%. We assume that those changes come from adjustments in the market shares of other drugs that are proportional to their market shares. Hence, the 1.7% decline in the market share of Drug A is split between Drug B and Drug C with Drug B absorbing

$^2/_3$ of the increases (1.13%) and Drug C absorbing $^1/_3$ (0.57%). Similarly, the 3.4% increase in the market share of Drug B is split between Drug A and Drug C with Drug A absorbing $^7/_8$ of the decrease (−2.98%) and Drug C absorbing $^1/_8$ (−0.42%), and the 5.1% increase in the market share of Drug C is split between Drug A and Drug B with Drug A absorbing $^7/_9$ of the decrease (−4.54%) and Drug B absorbing $^2/_9$ (−1.13%).[25]

4. Finally, we multiply the difference between the initial and new adjusted market share by the total number of claims for a given molecule (inclusive of both biosimilars and biologics) and by the difference in average spending between claims for the reference biologic and for each biosimilar.[26]

Note that this procedure does not take into account two potential effects:

1. changes in demand that may arise as a result of a change in formularies; and

2. changes in equilibrium price (which could arise if the government imposed restrictions to formulary coverage that affected the negotiation between manufacturers and payers).

[25] Occasionally, since our initial projection is linear, the counterfactual $\Delta\text{Frac. Uncovered}_{jk}$ implies negative shares or shares above 1. When that happens we cap shares at their natural bounds.

[26] An alternative way to compute this number would have been to use the publicly available Average Sales Price (ASP) series. However, ASP often refers to a different number of administered units for different drugs, which would have required additional adjustments.