# Shadow Pricing Meets Antitrust: Heterogeneous PBM Compliance with the FTC Standard Offering in the First Post-Settlement Insulin Formulary


## Abstract

In February 2026 the Federal Trade Commission entered a consent order against Express Scripts, Inc. (ESI) in *In the Matter of Caremark Rx, Express Scripts, Inc., and OptumRx and affiliated GPOs* (Docket 9437). Section I of the order prohibits ESI from placing the Low-WAC version of a within-manufacturer authorized-generic insulin pair on a less favorable tier — or with stricter utilization management — than the same manufacturer's High-WAC version. CVS Caremark and OptumRx remain in active administrative litigation and are not legally bound by Section I. The substantive contribution of this paper is descriptive: the first post-settlement ESI standard offering (April 1, 2026) shows heterogeneous per-pair compliance. ESI's National Preferred Formulary (NPF) achieves Section I tier parity on the Lilly Humalog / Insulin Lispro pair and the Viatris Semglee dual-WAC pair, drops Novo Nordisk's NovoLog franchise entirely (rendering the rule moot for that manufacturer pair), and shows a Section-I-violation-direction configuration on the Sanofi Lantus / unbranded Insulin Glargine pair — Lantus is preferred at tier 2 and the unbranded counterpart is excluded. The Sanofi pair has no FTC-pleading-encoded pre-period Section-I-violation-direction allegation; the April 2026 configuration is therefore the first publicly observable Section-I-violation-direction configuration documented on that within-manufacturer pair, not a "preserved" or "reverted" one. The Flex variant of the standard offering applies stricter compliance, excluding the High-WAC branded versions of Humalog, Lantus, Semglee, and Toujeo in favor of their respective Low-WAC counterparts. To accompany the descriptive finding we report a structured pre/post contrast statistic on the Section-I subsample (n=502 pair-PBM-month cells). The coefficient on the within-pair tier gap (`tier_diff`) is −83.08; the magnitude is encoding-sentinel-dependent (excluded coded as tier 99; preferred coded as tier 2; FTC pleading cells pre-encoded at +97) and the operative precision statement is the permutation p of 1/3 across three PBM clusters — the lowest attainable with one treated cluster. Given that all 56 post-period control cells are LOCF-carried from pre-period snapshots (§7.8, Appendix F.6) and that the OptumRx low-WAC parser flags trace 91.4 percent to artifacts or LOCF carry-forward rather than genuine commercial Premium Formulary entries (§7.7, Appendix F.5), the coefficient is best read as a structured pre/post contrast on the treated side with a parser-conditioned control reference, not as a counterfactual estimate; the per-pair pattern in Table 2 is the substantive empirical contribution. The conventional asymptotic CRV1

PBM SE is reported for transparency (SE 4.08; asymptotic p = 0.0024, 95% CI [−100.64, −65.52]) but is mechanically biased toward over-rejection at G=3. We discuss four structural limitations — a two-cell pre-period treated baseline derived from FTC pleadings, a single treated cluster, CRV1 inference on three PBM clusters, and a parser-conditioned OptumRx low-WAC control side — and conclude that the substantive contribution is the per-pair compliance pattern, not the magnitude of the pre/post contrast. The April 2026 snapshot is one publication of one PBM; medium-run dynamics through Q3 2026 and Q1 2027 will determine whether the heterogeneous pattern is transient or structural. The settlement's rule-based architecture moves the formulary as advertised on three of four manufacturers, while leaving a violation-direction configuration visible on the fourth.

---

## 1. Introduction

The American pharmacy benefit manager (PBM) industry is the institutional product of a half-century of contracting between pharmaceutical manufacturers, plan sponsors, and prescription drug benefit administrators. Three firms — Caremark, Express Scripts, and OptumRx — process roughly 79 percent of U.S. prescription drug claims [@mattingly2023pbmhistory] and operate the formularies that determine which drugs appear on each commercial and Medicare Part D plan, on which tier, and subject to which utilization-management restrictions. Formulary placement is more than a clinical or administrative artifact. It is the contractual lever through which PBMs translate rebate negotiations with manufacturers into market share: tier-placement discretion gives the PBM a credible threat of exclusion or disfavor, which manufacturers buy off with rebates tied to the wholesale acquisition cost (WAC) of their products [@ho2024formulary]. When two molecularly identical products carry different WACs but the High-WAC version pays a larger absolute rebate dollar — even at a smaller percentage discount — the PBM's optimal contract steers volume toward the High-WAC version. The patient on coinsurance or in a high-deductible plan absorbs the inflated list price at the pharmacy counter, even as the net price to the plan sponsor falls [@dickson2023grosstonet; @vannuys2021insulin; @feldman2021devil].

Insulin is the textbook case. Discovered in 1921 and never genericized — molecular reformulations and a regulatory pathway that treated insulin as a biologic rather than a small-molecule drug sustained branded monopoly pricing for a century [@greene2015generic] — insulin list prices grew 15 to 17 percent annually over 2012 to 2016 [@cefalu2018insulinaccess], while net prices (the prices plan sponsors actually paid after rebates) were flat or declining. The gross-to-net gap absorbed nearly the entire price wedge. By 2019, voluntarily negotiated commercial discounts — the rebate channel PBMs control — captured 60.5 percent of gross insulin sales, up from 27.0 percent in 2012 [@dickson2023grosstonet]. Patient-facing cost-sharing pegged to list rather than net, combined with cost-related underuse documented at 25 percent in clinic sam-

ples [@herkert2019insulinunderuse], turned the rebate channel from an arcane contracting mechanism into a public-health emergency.

Three regulatory responses ensued. The Inflation Reduction Act (IRA) capped Medicare beneficiary cost-sharing for insulin at $35 per month [@myerson2023ira], a policy that mechanically lowered out-of-pocket costs without addressing the underlying list-net distortion. The Federal Trade Commission opened a Section 6(b) investigation in 2022, issuing its first interim staff report in July 2024 [@ftc2024interim1] and a second on specialty generics in January 2025 [@ftc2025interim2]. In September 2024, the FTC filed an administrative complaint under Section 5 of the FTC Act against the three largest PBMs and their affiliated rebate-aggregating group purchasing organizations [@ftc2024complaint]. The Part 3 administrative complaint named specific instances of conduct the agency characterized as anticompetitive: in 2014, ESI's National Preferred Formulary excluded all rapid-acting insulins except branded Humalog (¶117); in 2019, after Lilly launched a 50-percent-lower-WAC authorized generic of Humalog, ESI kept branded Humalog as the sole preferred rapid-acting insulin and excluded the authorized generic (¶145); in 2021, after Viatris launched Semglee as a dual-WAC biosimilar — selling both a High-WAC branded "Semglee" and a Low-WAC unbranded "insulin glargine-yfgn" under the same Biologics License Application — ESI's NPF included High-WAC Semglee while excluding the Low-WAC counterpart (¶154); and in 2024, ESI's NPF was alleged to prefer High-WAC Tresiba and High-WAC Semglee, excluding the Low-WAC versions of each (¶246).

In February 2026, the FTC announced an asymmetric settlement: a consent order binding ESI only, leaving CVS Caremark and OptumRx in active administrative litigation [@ftc2026consent]. The order's centerpiece is Section I, a categorical prohibition framed not as a list of drugs but as a rule on formulary architecture: whenever a Drug Manufacturer markets both a High-WAC Version and a Low-WAC Version of a Drug Product — defined as products sharing the same active ingredient, dosage form, route of administration, and strength, and bearing the same FDA labeler segment — ESI's Standard Formulary may not place the Low-WAC Version on a less favorable tier or with stricter utilization management than the High-WAC Version. The Implementation Date was set at the first standard-offering publication after February 4, 2026; ESI's flagship standard-offering documents are typically refreshed each April. On April 1, 2026, ESI published four documents: the National Preferred Formulary (NPF, 23 pages), the NPF Exclusions list (20 pages), the National Preferred Flex Formulary Exclusions list (20 pages), and the High Performance Formulary (HPF, 3 pages). These are the first direct observation of how ESI complied with the Section I obligation.

This paper documents the first publicly observable post-Implementation-Date ESI Standard Offering and reports the per-pair Section-I configuration across the four manufacturer pairs the rule reaches. We do not claim a causal estimate. We construct a pair-level panel of seven Section-I-qualifying within-

manufacturer authorized-generic insulin pairs — Lantus / Sanofi unbranded insulin glargine (vial and SoloStar), Humalog / Lilly Insulin Lispro (vial and KwikPen), NovoLog / Novo Insulin Aspart (vial and FlexPen), and the Viatris Semglee dual-WAC pair — across three PBMs (ESI treated, CVS Caremark and OptumRx control), over 47 months from July 2022 through May 2026. The within-pair tier gap (`tier_diff = low_wac_tier − high_wac_tier`) and a strict-low-WAC-preference indicator (`low_wac_preferred = 1[tier_diff < 0]`) are the two outcomes. Treatment is the February 2026 consent order. Pair, PBM, and month fixed effects absorb time-invariant pair characteristics, common time shocks, and PBM-level differences in formulary architecture. Standard errors are clustered at the PBM level.

We organize the empirical findings around the per-pair pattern rather than the pre/post average. **The substantive contribution is the observed per-pair compliance pattern across the four manufacturer pairs the rule reaches.** ESI's April 2026 NPF achieves Section I tier parity on the Lilly Humalog / Insulin Lispro pair (both sides preferred) and the Viatris Semglee dual-WAC pair (both sides preferred). The Flex variant is stricter: it excludes branded Humalog, branded Lantus, branded Semglee, and Toujeo while preserving their Low-WAC counterparts on the formulary. Novo Nordisk's entire rapid-acting insulin franchise (NovoLog and the unbranded aspart authorized generic) is excluded from the NPF — the within-manufacturer rule is moot when the manufacturer's line is not preferred on either side. But on the Sanofi Lantus / unbranded insulin glargine pair on the standard NPF, ESI preserves a Section-I-violation-direction configuration. Lantus is preferred (tier 2); the Sanofi unbranded insulin glargine is excluded. This is the same configuration the FTC's Part 3 complaint alleged for ESI's Humalog pair in 2019 (¶145) and Semglee pair in 2021 (¶154) — and it appears on the first standard offering published after the consent order's Implementation Date.

To accompany the descriptive finding, we report a pair-level pre/post contrast. On the Section-I subsample (n=502 pair-PBM-month cells with both sides observed), the coefficient on `tier_diff` is −83.08 with conventional asymptotic CRV1 SE of 4.08. We do not present this as a credible causal estimate. The magnitude reflects the encoding sentinel (excluded coded as tier 99; preferred coded as tier 2; pre-period FTC pleading cells encode the violation pattern at +97), so it should be read as a direction-of-compliance summary rather than as a calibrated tier-distance effect. The operative precision statement is the permutation p-value of 1/3, the lowest attainable with three PBM clusters and one treated cluster; the asymptotic CRV1 p of 0.0024 is mechanically biased toward over-rejection (Section 5.2, Section 9). The `low_wac_preferred` strict-inequality outcome is statistically indistinguishable from zero on the Section-I subsample ($\beta$=−0.414, SE 0.164, asymptotic p=0.128), because Section I compliance is typically achieved by tier parity (low = high) rather than by strict preference, and tier parity makes the strict-inequality indicator false. The order requires non-inferiority, not strict preference, and ESI complied at the floor.

The contribution is fourfold. Substantively, to our knowledge this is the first publicly available pair-level evaluation of FTC anti-PBM enforcement at the level of the legal mechanism the consent order defines. Methodologically, the paper introduces a rule-based pair-level descriptive design that maps the order's textual obligation directly to an empirical object — an approach generalizable to other categorical regulations of formulary architecture. Empirically, the paper assembles the first publicly available enumeration of within-manufacturer authorized-generic insulin pairs in the United States, with FDA Purple Book BLA crosswalks and primary-source launch documentation. Policy-wise, the observed compliance pattern identifies a specific configuration — the Sanofi Lantus pair on the standard NPF — that warrants regulatory attention as the first publicly observable post-settlement Section-I-direction violation.

The paper proceeds as follows. Section 2 describes the institutional setting and the FTC Docket 9437 timeline. Section 3 lays out the conceptual framework. Section 4 details the data. Section 5 specifies the empirical strategy. Section 6 reports the headline estimates and the per-pair compliance pattern. Section 7 reports robustness checks. Section 8 interprets the heterogeneous-compliance finding mechanistically. Section 9 acknowledges limitations. Section 10 concludes.

---

## 2. Institutional Background

### 2.1 PBM market structure and the rebate-driven formulary

The PBM industry as a vertically integrated intermediary emerged in the 1990s and 2000s as plan sponsors outsourced the administration of drug benefits to firms that could negotiate prices, design formularies, and operate mail-order and specialty pharmacies. Consolidation accelerated through the 2010s: by 2023, three vertically integrated insurer-PBM-pharmacy combines — CVS Health (Aetna / CVS Caremark / CVS Pharmacy), Cigna (Evernorth / Express Scripts / Accredo), and UnitedHealth Group (UnitedHealthcare / OptumRx / Optum Specialty) — processed about 79 percent of all U.S. prescription drug claims [@mattingly2023pbmhistory]. Each of the three PBMs operates one or more "standard" commercial formularies that are sold as a default to small and mid-size plan sponsors, along with custom formularies for large employers. The standard offering is the unit of policy interest: it is the formulary that governs the median commercially insured American.

The economic mechanism by which PBM rebate negotiation distorts formulary placement away from list-price minimization is well documented [@ho2024formulary; @feldman2021devil]. Manufacturers compete for preferred placement on the PBM's formulary by offering rebates tied to the list price (WAC) of their products. Because the absolute rebate dollar grows with the WAC, a manufacturer with a High-WAC product can outbid a Low-WAC competitor even if the Low-WAC competitor offers a larger percentage discount.

The PBM, retaining a portion of the rebate as administrative revenue or returning it to plan sponsors as a rebate guarantee, internalizes the High-WAC product's larger rebate dollar; the plan sponsor receives a lower net price. The patient on coinsurance or in a high-deductible plan, however, faces a copay calculated against list rather than net — and absorbs the inflated WAC at the pharmacy counter [@meiri2020insulinoop; @tseng2020insulinpartd]. Vertical integration of PBMs with insurers compounds the dynamic: a vertically integrated firm internalizes both the rebate revenue and the premium revenue, while a non-integrated rival faces input costs set by a competitor [@gray2023verticalintegration].

The empirical magnitude of the distortion is substantial. Feldman's [-@feldman2021devil] analysis of approximately one million Medicare Part D claims (2010–2017) found that the share of generics placed on the most preferred tier dropped from 73 to 28 percent, while the share of drugs on tiers inappropriate to their generic-or-brand status rose from 47 to 74 percent. Klebanoff et al. [-@klebanoff2024adalimumab] found that 98.9 percent of 5,609 Medicare Part D plans covered branded Humira, while only 53.4 percent covered any adalimumab biosimilar. Axelsen et al. [-@axelsen2025biosimilars] reported that 99 percent of adalimumab Part D spending in 2023 still flowed to the branded originator despite multiple lower-priced biosimilars. Kakani et al. [-@kakani2025semglee] estimated that improved insurance coverage — not pharmacist substitution or prescriber behavior — was the dominant mechanism behind Semglee's market-share gain in employer-sponsored insurance. The binding constraint on Low-WAC product adoption is the formulary, not the regulatory pathway.

### 2.2 Insulin as the canonical case

Insulin's pricing pathology is well documented. Dickson et al. [-@dickson2023insulinprices; -@dickson2023grosstonet] show that net prices of long-acting insulin fell about 8 percent per year after biosimilar entry while list prices continued to grow; total discounts on the four leading insulins grew from $4.9 billion (37.6 percent of gross) in 2012 to $22.0 billion (81.4 percent) in 2019. Van Nuys et al. [-@vannuys2021insulin] found that of $100 in U.S. insulin expenditure, the manufacturer share fell from $70 (2014) to $47 (2018) while the PBM share rose 154.6 percent. Meiri et al. [-@meiri2020insulinoop] documented that commercial insurers absorbed rising list prices through plan design — average out-of-pocket was flat over 2006–2017 — except for high-deductible plan enrollees, whose exposure to list rose sharply. Tseng et al. [-@tseng2020insulinpartd] documented analogous Part D dynamics. The combined picture: list-net divergence is large, growing, and concentrated in the rebate channel; the channel runs through formulary architecture; the patients exposed to the inflated list are those on coinsurance or in HDHPs.

In response, the major insulin manufacturers launched authorized-generic or unbranded counterparts of their flagship products at substantially lower

list prices. Eli Lilly introduced "Insulin Lispro Injection" — a Lilly-labeled authorized generic of Humalog, at 50 percent lower WAC — in May 2019 [@ftc2024complaint, ¶145]. Sanofi launched an unbranded "Insulin Glargine" at approximately 50 percent lower WAC than Lantus in June 2019, marketed through Sanofi's Winthrop U.S. generics arm. Novo Nordisk launched an unbranded "Insulin Aspart" at 50 percent lower WAC than NovoLog in January 2020. Viatris and Biocon launched the FDA-interchangeable Semglee biosimilar in November 2021 as a dual-WAC product — selling both a High-WAC branded "Semglee" and a Low-WAC unbranded "insulin glargine-yfgn" under the same Biologics License Application. By 2022, every major U.S. insulin franchise that was Section-I-eligible had a within-manufacturer Low-WAC counterpart on the market. Whether that counterpart appeared on PBM formularies — and on which tier — became an empirical question with first-order welfare implications.

### 2.3 FTC Docket 9437

The Federal Trade Commission's investigation of PBM rebate practices proceeded along two parallel tracks. The Section 6(b) investigation, opened in 2022, produced two interim staff reports: a July 2024 report on PBM market structure and rebate-driven formulary placement [@ftc2024interim1], and a January 2025 report quantifying specialty-generic markups of thousands of percent and $7.3 billion in excess revenue 2017–2022 [@ftc2025interim2]. In parallel, in September 2024, the FTC filed an administrative complaint under Section 5 of the FTC Act — *In the Matter of Caremark Rx, LLC, Zinc Health Services, LLC, Express Scripts, Inc., Evernorth Health, Inc., Ascent Health Services, LLC, OptumRx, Inc., and Emisar Pharma Services, LLC* (Docket 9437) — alleging that the three largest PBMs and their affiliated rebate-aggregating GPOs engineered formulary exclusions and tier placements to favor High-WAC insulin products, inflating list prices and harming patients on coinsurance and high-deductible plans [@ftc2024complaint]. The complaint contained 274 numbered paragraphs; 64 named at least one specific insulin SKU. None named a GLP-1 receptor agonist.

Four paragraphs are load-bearing for the present paper. Paragraph 117 alleges that in 2014, ESI's National Preferred Formulary excluded all rapid-acting insulins other than Humalog. Paragraph 145 alleges that in 2019, after Lilly launched the 50-percent-lower-WAC Insulin Lispro authorized generic, ESI kept branded Humalog as the sole preferred rapid-acting insulin and excluded the authorized generic from the NPF. Paragraph 154 alleges that in 2021, ESI's NPF preferred High-WAC Semglee while excluding the Low-WAC unbranded insulin glargine-yfgn. Paragraph 246 alleges that ESI's 2024 NPF preferred High-WAC Tresiba and High-WAC Semglee while excluding the Low-WAC versions of each. These are sworn FTC pleadings, not adjudicated facts. They describe specific ESI conduct at specific moments — exactly the conduct Section I now prohibits — and provide the pre-period evidentiary baseline on which the consent order's

prohibition was negotiated.

On February 4, 2026, the FTC and Express Scripts announced an Agreement Containing Consent Order to resolve the administrative complaint as to ESI alone; CVS Caremark and OptumRx remain in active administrative litigation, with no announced settlement timeline as of this writing [@ftc2026consent]. The consent order is structured in six operative sections. Section I (the prohibition on placing a Low-WAC Version on a less favorable tier than the same-manufacturer High-WAC Version, with no stricter utilization management) is the textual core of the order and the empirical focus of this paper. Section II requires ESI's Standard Formulary to be available without charge to any plan sponsor that requests it. Section III mandates pass-through of manufacturer rebates to plan sponsors based on net rather than list price for the standard offering. Section IV protects manufacturer patient-assistance program (PAP) revenues for the four Section IV insulins (Toujeo, Tresiba, Lyumjev, Fiasp). Section V institutes monitoring and compliance reporting. Section VI sets the Implementation Date as the first standard-offering publication after February 4, 2026.

ESI's flagship standard offering — the National Preferred Formulary — is typically refreshed each April. On April 1, 2026, ESI published the four documents listed in Section 4 of this paper: the NPF, the NPF Exclusions, the Flex variant Exclusions, and the High Performance Formulary. These are the first publication after the Implementation Date and the first direct observation of how ESI complied with the Section I rule.

In addition to the four load-bearing pre-period paragraphs above, the Part 3 complaint documents the rebate-contracting context of the alleged conduct in extensive (largely redacted) detail. Paragraph 110 alleges that in 2023 ESI preferred Lilly's rapid-acting insulins (Humalog and Lyumjev) on the NPF while excluding Novo's rapid-acting line (NovoLog and Fiasp). Paragraph 111 alleges that in 2023 ESI excluded Sanofi's Lantus from the NPF in favor of Novo's Levemir and Tresiba — a configuration in which the originator brand (Lantus) was *excluded* rather than preferred. The 2023 long-acting configuration is therefore *not* Section-I-violation-direction on the Sanofi pair (Sanofi has no preferred branded long-acting on the 2023 NPF to pair against an excluded Low-WAC version). The 2024 NPF configuration alleged in ¶246 — preferred High-WAC Tresiba and Semglee, excluded Low-WAC versions — re-introduces the Section-I-violative within-manufacturer pattern. The historical record from the FTC's own pleadings is therefore a sequence of formulary configurations, some of which fall in Section-I-violation-direction and some of which do not.

### 2.4 Why this is a clean test of a rule-based remedy

The asymmetric application of the consent order — binding ESI but not CVS Caremark or OptumRx — combined with the rule-based (rather than list-based) prohibition in Section I, creates an unusually clean test of behavioral antitrust enforcement. The treatment is sharp: February 2026, applied to one of three

competitors in a concentrated market. The empirical object is sharply defined: the within-manufacturer authorized-generic pair, with seven Section-I-qualifying pairs textually grounded in the FDA Purple Book and verified by primary-source manufacturer press releases. The outcome is directly observed: the formulary tier and utilization-management flag on the PBM's own published standard offering each quarter. The pre-period is documented by the agency's own pleadings at paragraph-level granularity. There is no claims-lag, no measurement opacity, and no restricted data — all of the load-bearing observations are public.

The design has four structural limitations we make no attempt to paper over. First, the FTC's Part 3 pleadings are allegations, not admissions; their use as the pre-period ESI baseline depends on the proposition that an FTC administrative complaint accurately describes the conduct it alleges. Second, only one of three PBMs is treated; cluster-robust inference with three clusters is mechanically biased. Third, the post-period observation window — April through May 2026 — is short. Fourth, the control-side post-period cells (CVS Caremark and OptumRx) are entirely last-observation-carried-forward (LOCF) from pre-period snapshots; the Wayback historical control corpus is heterogeneous (it includes Medicare Part D, FEP, state/union/employer-specific, specialty, and criteria-PDL documents that are not directly comparable to ESI's Standard Offering); and the OptumRx low-WAC parser flags are 91.4 percent artifact or carry-forward rather than verified commercial Premium Formulary entries. The control side is therefore a parser-conditioned LOCF reference, not a clean commercial-mail-order counterfactual. Section 9 returns to each.

---

### 2.5 Why insulin and not GLP-1s

A natural question is whether the rule-based analysis extends to the GLP-1 receptor agonist class — Ozempic, Wegovy, Mounjaro, Zepbound, Trulicity — which dominates the diabetes and obesity drug-spending landscape and is the subject of substantial public attention. The answer is no, and the reason is structural: Section I is a rule about within-manufacturer authorized-generic pairs. The GLP-1 class has no within-manufacturer Low-WAC counterparts. Each GLP-1 product is marketed by a single manufacturer as a single SKU at a single WAC. There is no Lilly-unbranded Mounjaro, no Novo-unbranded Ozempic. Section I therefore cannot bind on the GLP-1 class. The FTC Part 3 administrative complaint reflects this empirical reality: of 274 numbered paragraphs, 64 name at least one specific insulin SKU; zero name a GLP-1 receptor agonist. The agency's evidence base, the consent order's textual obligation, and the empirical scope of this paper all converge on the insulin class.

This is a feature, not a bug, of the rule-based design. A list-based remedy would name specific drugs and would face implementation challenges as new products enter and exit the market. The rule-based design is forward-compatible: if a

manufacturer launches a within-manufacturer Low-WAC counterpart to a GLP-1 brand in the future, that pair will fall under Section I automatically. The rule's domain is the set of pairs the manufacturer chooses to create; the regulator does not need to update the list to keep the remedy current.

### 2.6 Adjacent biosimilar evidence

The empirical findings on biosimilar uptake outside the insulin space — Klebanoff et al. [-@klebanoff2024adalimumab] on adalimumab, Axelsen et al. [-@axelsen2025biosimilars] on Medicare Part D biosimilar spending — establish that the binding constraint on Low-WAC product adoption is formulary coverage rather than regulatory approval. The Kakani et al. [-@kakani2025semglee] interrupted-time-series analysis of Semglee market share post-interchangeability designation, and the Kwon et al. [-@kwon2025substitution] state-law DiD, both identify formulary placement and insurance coverage as the dominant mechanism behind Low-WAC uptake. The present paper is upstream of these utilization-level papers: it estimates whether the formulary placement itself changes when the regulator mandates that the within-manufacturer Low-WAC tier be no worse than the High-WAC tier. The downstream question — whether Section I compliance translates into Low-WAC utilization gains for patients on coinsurance — is a natural follow-up that requires claims data this paper does not access.

---

## 3. Conceptual Framework: Partial and Over-Compliance under Categorical Regulation

The Section I obligation is categorical: it specifies a comparison (within-manufacturer pair) and a constraint (the Low-WAC tier must be no worse than the High-WAC tier, with no stricter utilization management). It does not require the PBM to prefer the Low-WAC version. It does not require the PBM to cover either version. It does not specify a tier number, only a relative ordering. The minimum-compliance behavior, given this structure, is a corner solution: drop both sides of any inconvenient pair off the formulary entirely (no pair, no constraint), or place them at parity (low = high), which clears the rule at the lowest restructuring cost.

A PBM facing a rule of this form makes three decisions per manufacturer pair. First, whether to keep the pair on the formulary at all. Excluding both sides removes the pair from the rule's domain — the constraint binds only on pairs the PBM chooses to cover. Second, conditional on keeping the pair, what tier configuration to adopt. Three configurations are Section-I-compliant in tier terms: both sides at the same tier (parity), or the Low-WAC version on a strictly more preferred tier than the High-WAC version (strict preference). A fourth configuration — the Low-WAC version on a less favorable tier — is Section-I-violation-direction. Third, conditional on tier, what utilization management to

apply. Stricter prior-authorization, step-therapy, or quantity-limit restrictions on the Low-WAC version are also prohibited.

The PBM's optimization weighs three forces. (i) The rebate gradient: keeping the High-WAC version preferred maximizes the absolute rebate dollar per script. (ii) The compliance cost: violating Section I exposes the PBM to enforcement and the formula's reputational damage. (iii) The substitution cost: dropping a pair entirely sacrifices any rebate revenue from either side and may reduce plan sponsor satisfaction if the molecule is clinically important.

The categorical structure of Section I produces several testable predictions. First, pairs where the rebate gradient is steep — where the High-WAC dollar is large relative to the Low-WAC dollar — will be the most likely sites of partial or non-compliance. Second, pairs whose manufacturers offer a single product line (the manufacturer cannot easily substitute volume from a different molecule) will be more likely to receive compliance, because the PBM cannot drop the pair without losing all manufacturer revenue. Third, variants of the PBM's standard offering with different audiences (a flagship NPF for cost-sensitive small employers; a stricter Flex variant for cost-aggressive plans; a narrow High Performance Formulary for low-cost-share plans) may apply Section I with different intensity, because the rebate-share economics differ. Fourth, the rule's minimum-compliance behavior is tier parity, not strict Low-WAC preference; observed strict preference would constitute over-compliance and is informative about the PBM's negotiation position with the relevant manufacturer.

These predictions structure the empirical analysis. The heterogeneous-compliance finding reported in Section 6 — Sanofi pair non-compliant on the NPF; Lilly pair at parity on the NPF and strictly compliant on the Flex variant; Novo Nordisk pair dropped entirely; Viatris pair at parity on the NPF and strictly compliant on the Flex variant — maps onto these forces in a way Section 8 develops in detail.

---

## 4. Data

### 4.1 Sample

The empirical object is the (within-manufacturer authorized-generic insulin pair × PBM × month) cell. The pair table contains 15 rows: 7 Section-I-qualifying within-manufacturer authorized-generic insulin pairs, 1 UNVERIFIED pair (Tresiba's alleged authorized generic, asserted only by FTC complaint ¶246 with no public Novo Nordisk press-release confirmation as of 2026-05-12), and 7 economist comparator pairs (Toujeo, Lyumjev, Fiasp, Levemir, Basaglar, Admelog, Rezvoglar). The seven Section-I-qualifying pairs are: Lantus / Sanofi unbranded insulin glargine vial (BLA 021081); Lantus SoloStar / Sanofi unbranded insulin glargine SoloStar (BLA 021081, pen formulation); Humalog / Lilly Insulin Lispro authorized generic vial (BLA 020563); Humalog

KwikPen / Lilly Insulin Lispro KwikPen authorized generic (BLA 020563, pen formulation); NovoLog / Novo Nordisk unbranded Insulin Aspart vial (BLA 020986); NovoLog FlexPen / Novo Nordisk unbranded Insulin Aspart FlexPen (BLA 020986, pen formulation); Semglee / Viatris unbranded insulin glargine-yfgn (BLA 761201, the dual-WAC pair). The full pair table — with manufacturer NDC labeler codes, launch dates, FDA Purple Book BLA crosswalks, primary-source press release URLs, and FTC complaint paragraph citations where applicable — appears in Appendix B and is available as `data/raw/insulin_pairs/within_manufacturer_authorized_generic_pairs.csv`.

A pair qualifies for the Section-I-qualifying set when (a) the High-WAC and Low-WAC versions share an FDA Purple Book BLA, (b) the manufacturer of record on the NDC labeler segment is the same on both sides, (c) the dosage form, route, and strength match across the two versions, and (d) the Low-WAC version's launch is independently documented by a primary-source press release that specifies the launch date and WAC differential. Pen-formulation pairs (Lantus SoloStar, Humalog KwikPen, NovoLog FlexPen) share the BLA of their vial counterparts; we keep them as separate pair rows for completeness of the formulary panel.

The pair-level panel contains 2,115 rows (15 pairs × 3 PBMs × 47 months from 2022-07 through 2026-05). The Section-I-only subsample contains 987 rows (7 pairs × 3 PBMs × 47 months). After the April 2026 ESI PDF ingest, 502 of the Section-I subsample rows have both sides of the pair observed (a non-NaN `tier_diff` cell), and 1,053 of the full-pair-panel rows do.

### 4.2 Outcomes

For each (pair $p$, PBM $b$, month $t$) cell we observe $\texttt{high\_wac\_tier}_{p,b,t}$ (the tier on which the PBM places the High-WAC version) and $\texttt{low\_wac\_tier}_{p,b,t}$ (the corresponding tier for the Low-WAC version). The two derived outcomes are

$$\Delta_{p,b,t} = \text{low\_wac\_tier}_{p,b,t} - \text{high\_wac\_tier}_{p,b,t}$$

and

$$Y_{p,b,t} = \mathbb{1}\{\Delta_{p,b,t} < 0\}.$$

$Y_{p,b,t}$ is one when the PBM places the Low-WAC version on a strictly more preferred tier than the same-manufacturer High-WAC version. $\Delta_{p,b,t} > 0$ corresponds to Section-I-violation-direction conduct; $\Delta_{p,b,t} = 0$ corresponds to tier-parity compliance (the legal floor); $\Delta_{p,b,t} < 0$ corresponds to active Low-WAC preference (over-compliance). Excluded products are coded as a sentinel tier (99) for the purposes of the `tier_diff` calculation. The choice of 99 is arbitrary; any monotone transformation of "excluded > any preferred tier" would yield the same sign on the pre/post contrast but a different magnitude. The encoding is dropped easily for robustness and pipelines for alternative sentinels (NaN; rescaled exclusion penalty) are wired but not used in the headline.

### 4.3 Sources for the (pair × PBM × month) tier cells

Four classes of evidence fill each cell.

**FTC Docket 9437 Part 3 administrative complaint.** Of the complaint's 274 numbered paragraphs, 19 describe specific ESI formulary behaviors on insulin pairs. Four paragraphs are load-bearing for the pre-period treated baseline. **¶117** describes ESI's 2014 NPF as preferring branded Humalog as the sole rapid-acting insulin. **¶145** alleges that in 2019, ESI kept High-WAC Humalog as the only preferred rapid-acting insulin and excluded the Low-WAC Lilly Insulin Lispro authorized generic after its May 2019 launch. **¶154** alleges that ESI's 2021 NPF included High-WAC Semglee while excluding Low-WAC unbranded insulin glargine-yfgn. **¶246** alleges that ESI's 2024 NPF preferred High-WAC Tresiba and High-WAC Semglee while excluding the Low-WAC versions of each. Cells filled from these paragraphs carry `source = 'ftc_part3_complaint'` and are encoded as tier sentinels (preferred = 2; excluded = 99) because the complaint does not state a specific numeric tier. We label these as allegations, not adjudicated facts, throughout. These cells provide the pre-period treated observations: $n = 2$ Section-I-subsample ESI pre-period cells with both sides observed, encoding the Tresiba 2024 ¶246 episode at tier_diff = +97.

**Direct ingestion of the April 1, 2026 ESI Standard Offering PDFs.** Four standard-offering documents dated April 1, 2026 were retrieved from the Express Scripts public portal on 2026-05-12: the 2026 National Preferred Formulary (23 pages); the 2026 NPF Exclusions list (20 pages); the 2026 National Preferred Flex Formulary Exclusions list (20 pages); and the 2026 Express Scripts High Performance Formulary (3 pages). All four carry the April 1, 2026 effective date — the first standard-offering publication after the Section I Implementation Date. The PDFs do not print numeric tier numbers per SKU. They are categorical: a SKU is either listed as covered on the formulary (encoded `tier_num = 2`) or listed as excluded (encoded `excluded = TRUE`, `tier_num = NaN`). The class-level matrix on page 10 of the NPF and page 7 of the two Exclusions PDFs renders a three-column structure — Drug Class | Excluded Medications | Preferred Alternatives — and is the load-bearing source for per-pair Section I adjudication. The HPF prints no exclusion matrix and is parsed as a flat covered-list only. The parser `data/scripts/07_parse_esi_april_2026_formularies.py` extracts each PDF into a per-PDF CSV under `data/raw/esi_post_feb_2026/parsed/`; the cross-referenced pair-level table is written to `data/raw/ftc-docket/esi_postsettlement_observed_april_2` (120 rows = 15 pairs × 2 sides × 4 variants). Cells filled from these PDFs carry `source = 'esi_april_2026_pdf'`.

**Internet Archive Wayback Machine PBM-PDF parses.** For each PBM we retrieve every available snapshot of the standard commercial formulary PDF between January 2023 and the present, parse it for (drug, tier, UM flag) rows, and join the parsed rows to the pair table via FDA-NDC-aware brand-name matching. The Wayback corpus refreshed on 2026-05-12 successfully down-

loaded 379 PDFs (CVS 252, ESI 37, OptumRx 90), producing 1,218 directly observed (non-imputed) (PBM × drug × month) cells. Coverage on the pair-level panel is uneven: CVS Caremark contributes 60 percent of pair-month cells on the high-WAC side and 60 percent on the low-WAC side; OptumRx contributes 86 percent on the high-WAC side and 100 percent on the low-WAC side; ESI's Wayback snapshots in the 2022–2025 window are HTML wrappers and yield no pre-period direct-PDF observations. The OptumRx low-WAC coverage at 100 percent is a parser artifact — the parser reads "not enumerated" as "default tier covered" rather than "excluded." The 10-snapshot stratified hand audit reported in §7.7 (`analysis/robustness/m4_optumrx_audit.py`) confirms this is not a conservative reading but an active false-positive pattern: 21 percent of audited cells captured concatenation-artifact rows, 70 percent are LOCF-carried from prior snapshots whose parsed_rows contain no genuine entry for the probe, and the 8.6 percent of confirmed-genuine cells trace to client-specific or state-Medicaid select formularies rather than OptumRx's commercial Premium Formulary. The OptumRx low-WAC control side is therefore treated as a limitation rather than as a clean counterfactual; see Section 9 (L4) for the full disposition.

A second source-corpus caveat applies to the Wayback control-side reference. The Internet Archive snapshots used for CVS Caremark and OptumRx are not a clean standard-commercial-formulary benchmark. The corpus includes Medicare Part D formularies, Federal Employees Program documents, state / union / employer-specific select formularies, specialty formularies, criteria PDLs, and exclusion documents that are not directly comparable to the ESI Standard Offering. The control-side reference is therefore heterogeneous on the document-type margin in addition to being LOCF-carried on the time margin; see `data/raw/_provenance/wayback/README.md` and `data/README.md` for the corpus composition.

**FTC consent order Section I (post-period legal floor, deprecated).** Before the April 2026 PDF ingest, post-period ESI cells were filled by the Section I legal-floor assumption (low-WAC tier $\leq$ high-WAC tier; both sides at tier 2; `tier_diff = 0`). With the April 2026 PDF ingest, this encoding is now superseded for months $\geq 2026-04$. The pre-2026-04 post-period months (February and March 2026) — where the consent order is in force but no published April standard offering is yet observed — retain the legal-floor encoding and contribute 8 cells (2 months × 4 ESI variants) to the post-period treated set.

A consequential encoding asymmetry across pairs deserves explicit note. The Sanofi Lantus / unbranded insulin glargine pair contributes zero ESI pre-period cells because the FTC complaint does not paragraph-cite a Lantus-specific within-manufacturer Section-I-direction violation: ¶111 (2023) alleges that ESI excluded Lantus entirely in favor of Novo's Levemir and Tresiba, which is not the High-WAC-preferred / Low-WAC-excluded pattern Section I targets. Two of the four ESI pre-period cells in the all-pairs specification are encoded from ¶246

on the Tresiba pair, which is itself flagged as UNVERIFIED; Tresiba is not in the Section-I subsample. The Section-I-subsample pre/post contrast magnitude is therefore mechanically driven by the Humalog encoding (¶117/¶145); the Sanofi pair contributes only to the post-period mean. Section 7.9 and Appendix F.4 report the leave-one-out sensitivity confirming that the Section-I headline coefficient is invariant to dropping Tresiba and is robust to every Section-I per-pair drop.

### 4.4 Treatment

Let $D_b = 1$ for ESI and zero otherwise; $P_t = 1$ for $t \geq 2026 - 02$ and zero otherwise. The treatment indicator is $D_b \times P_t$. The pair-level panel encodes `treated_pbm = 1[pbm = 'esi']`, `post = 1[month \geq 2026-02]`, and `treated_post = treated_pbm` $\times$ `post`.

### 4.5 Placebo and falsification sets

The placebo class is the four originator insulin brands for which no within-manufacturer Low-WAC counterpart exists (Toujeo, Lyumjev, Fiasp) plus Levemir (which Novo Nordisk has discontinued). Section I cannot bind on any PBM for these brands. A nonzero DiD on this placebo set would indicate that the headline estimate is contaminated by ESI-side formulary restructuring orthogonal to the Section I mechanism. Cross-manufacturer biosimilar comparators (Basaglar versus Lantus, Semglee versus Lantus, Admelog versus Humalog) are economically the most-relevant Low-WAC substitutes but are not within the Section I textual obligation. We use them in supplementary analyses to characterize cross-class spillovers.

### 4.6 Reproducibility

All data are public and accessed without authentication. The pipeline runs end-to-end from `data/00_run_all.py` (download / scrape) through `data/scripts/06_build_pair_level_panel.py` (pair-level panel) to `analysis/run_all.py` (DiD estimation and figure production). Per-step logs are in `data/scripts/log/` and `analysis/log/`. No human subjects research is involved; the project is not subject to IRB review. Full data-dictionary documentation appears in `data/data-dictionary.md`.

---

## 5. Methods

### 5.1 Pair-level pre/post contrast specification

We report a pair-level pre/post contrast — a case-study DiD with one treated cluster and two ESI pre-period cells. We use the DiD nomenclature for transparency, but the specification does not satisfy the conditions for a credible causal

estimate: with $n = 2$ ESI pre-period cells and a single treated cluster, parallel-trends assumptions cannot be tested and conventional cluster-robust inference is mechanically biased. The specification is

$$Y_{p,b,t} = \beta \cdot (D_b \times P_t) + \alpha_p + \gamma_t + \delta_b + \varepsilon_{p,b,t},$$

with pair fixed effects $\alpha_p$ absorbing time-invariant pair characteristics (which manufacturer, which molecule, which formulation), month fixed effects $\gamma_t$ absorbing common time shocks, and PBM fixed effects $\delta_b$ absorbing time-invariant PBM differences in formulary architecture. Standard errors are clustered at the PBM level (CRV1 with $G = 3$); we report them with the explicit understanding that the conventional asymptotic CRV1 inference is unreliable at this cluster count. We estimate $\beta$ on (a) the full pair panel and (b) the Section-I-only subsample of the seven textually-qualifying pairs. Identical specifications are estimated with `tier_diff` ($\Delta$) as a continuous outcome.

### 5.2 Inference

With a single treated PBM cluster, cluster-robust inference is mechanically biased toward over-rejection [@cameron2008bootstrap]. We report PBM-clustered CIs as illustrative rather than definitive. The operative precision statement is the exact permutation p-value of 1/3 across the three PBM clusters (Section 7.6) — the lowest attainable with three clusters and one treated unit. The asymptotic CRV1 p-value of 0.0024 is reported in tables for transparency but is not the precision statement we rely on. Wild-cluster bootstrap and Conley-Taber CRVE are degenerate at $G = 3$ and do not rescue inference. We discuss the bound on credible inference in Section 9.

### 5.3 Pre-trends and event study

The pre-period ESI Section-I subsample contains $n = 2$ direct cells (both from FTC Part 3 ¶246, encoding the Tresiba 2024 episode at `tier_diff = +97`). With two pre-period treated cells, a formal event-study with leads and lags is not informative; we report the event-study figure as a descriptive visualization (Figure 2 / `analysis/figures/fig08_pair_event_study.png`) but make no formal pre-trends test claim. The qualitative direction (alleged violation pre-treatment, mandated compliance post-treatment) is documented but not statistically falsifiable.

### 5.4 Robustness program

We report five implemented structural-rigor checks. (i) **Leave-one-out sensitivity** across all seven Section-I pairs: the Section-I headline coefficient is invariant to dropping the UNVERIFIED Tresiba pair and is robust to every Section-I per-pair drop ($\Delta\beta$ between $-2.4$ and $+5.6$; full results in eAppendix F.4). (ii) **Exact permutation inference** (`analysis/robustness/permutation_inference.py`): the treated label

is permuted across the three PBMs and the pair-FE pyfixest specification refit; the permutation distribution has three cells (−106.20 with ESI treated; +6.04 with CVS treated; +7.46 with OptumRx treated); the permutation p-value is 1/3, the lowest attainable with three PBM clusters and one treated cluster. (iii) **Per-variant subsample DiDs** across the four April 2026 ESI variants (NPF, NPF Exclusions, Flex Exclusions, HPF): the Flex variant pulls more strongly toward compliance than the NPF ($\beta = -98.2$ on Flex Exclusions vs. $\beta = -78.6$ on the NPF). (iv) **OptumRx low-WAC parser hand audit** (`analysis/robustness/m4_optumrx_audit.py`): a stratified hand audit of 70 OptumRx low-WAC Section-I cells finds 91.4 percent are artifacts or carry-forward; the OptumRx low-WAC control side is accordingly demoted from mechanism control to limitation. (v) **Control-side LOCF characterization** (`analysis/robustness/m9_locf_characterization.py`): zero of 56 post-period control cells are directly observed; every post-period control observation is LOCF-carried from a pre-period Wayback snapshot, and the headline DiD restricted to directly observed control rows is mechanically collinear.

Three additional checks are non-applicable to the Section-I pair-level design and are documented here for transparency. (a) **Rambachan-Roth HonestDiD on the pair-level headline** requires a non-degenerate pre-period lead vector; the Section-I pre-period treated baseline rests on two ESI cells (FTC ¶117 in 2014 and ¶145 in 2019) with one treated cluster, so the implied lead vector is degenerate and HonestDiD returns the trivial bound. The R `HonestDiD` package is installed under `scripts/rstats/run_honestdid.R` for a future SKU-level event-study application if the design is reformulated. (b) **No-authorized-generic placebo pair** would refit the difference-in-differences on within-manufacturer pairs that lack an authorized-generic sibling (within insulin: Toujeo, Lyumjev, Fiasp, Levemir) and therefore could not be subject to Section I; these pairs have no Low-WAC counterpart, so the outcome (within-pair tier gap) is structurally undefined. (c) **Drop-legal-floor on the Section-I subsample** removes pre-period ESI rows whose source attribution is `ftc_docket_9437`; in the current build, dropping these rows collapses the pre-period to zero treated cells, so the restriction is not estimable. The script (`analysis/robustness/drop_legal_floor.py`) is preserved for future builds with accumulating pre-period docket-coded rows from later FTC public filings or adjudicated findings.

### 5.5 Honest assessment

The pair-level pipeline runs end-to-end and produces estimates whose direction is well-defined under Section I and the FTC's pre-period characterization of ESI conduct (move from `tier_diff > 0` toward `tier_diff $\leq$ 0`), but whose magnitude reflects an encoding sentinel and a single-cluster identification configuration. The substantive contribution is the heterogeneous-compliance pattern across the four manufacturer pairs (Section 6), not the magnitude of

the average effect.

---

## 6. Results

### 6.1 Pair-level pre/post contrast as a descriptive transparency statistic

We report a pair-level pre/post contrast on the within-pair tier gap as a transparency statistic. Given the structural identification limits documented in §7.7 (OptumRx low-WAC parser audit), §7.8 (control-side LOCF characterization), and §9, the coefficient is not interpreted as a counterfactual estimate. Table 1 reports the contrast on the two outcomes for the full pair panel and the Section-I-only subsample. The substantive empirical contribution of the paper is the per-pair configuration pattern in Table 2 (next subsection); Table 1 is a one-number summary of the treated-side pre/post transition.

**Table 1. Structured pre/post contrast statistic (descriptive).**

| Subsample | Outcome | $\hat{\beta}$ | SE (CRV1) | Asymp. p (CRV1)[a] | Permutation p[b] | 95% CI (CRV1)[c] | n |
|---|---|---|---|---|---|---|---|
| Section-I only | `tier_diff` | −83.08 | 4.08 | 0.0024 | **1/3** | [−100.64, −65.52] | 502 |
| Section-I only | `low_wac_preferred` | +0.114 | 0.164 | 0.128 | 1/3 | [−1.120, +0.292] | 502 |
| All pairs | `tier_diff` | −106.20 | 1.69 | 0.0003 | 1/3 | [−113.47, −98.92] | 1,053 |
| All pairs | `low_wac_preferred` | +0.046 | 0.110 | 0.713 | 1/3 | [−0.426, +0.519] | 1,053 |

*Notes:* This table documents the source files, scripts, variables, or data inputs used in the analysis. It is included to make the construction of the analytic evidence reproducible.

*Source*: `analysis/tables/pair_did.csv`. Specification: $Y \sim D_b \times P_t \mid \alpha_p + \gamma_t + \delta_b$. SE clustered on PBM (G=3). *Notes*: (a) Asymptotic CRV1 p-value, mechanically biased toward over-rejection at $G = 3$ with one treated cluster; reported for transparency only. (b) Exact permutation p-value across the three possible PBM-relabelings (ESI treated, CVS treated, OptumRx treated); 1/3 is the lowest attainable value with three clusters. (c) Pointwise CRV1 confidence interval; the within-cluster variance is dominated by sentinel-encoded cells and the interval is informative about encoding variance, not parameter uncertainty. (d) Per §7.7 (OptumRx low-WAC parser audit) and §7.8 (control-side LOCF characterization), the control side is parser-conditioned and LOCF-carried throughout the post-period; the coefficient is best read as a structured pre/post contrast on the treated side, not as a counterfactual estimate. The per-pair configuration pattern in Table 2 is the substantive empirical contribution of the paper.

The Section-I-subsample coefficient on `tier_diff` is −83.08. The sign is consistent with the Section I obligation: ESI moved toward `tier_diff $\leq$ 0` (Low-WAC tier no worse than High-WAC tier) relative to the control PBMs' pre/post change. The magnitude reflects the encoding sentinel (excluded = 99; preferred = 2) and the fact that the FTC Part 3 complaint encoded ESI's two pre-period cells at `tier_diff = +97` (full violation) while the April 2026 PDFs show three of the seven Section-I pairs at `tier_diff = 0` (formal compliance), two pairs still at `tier_diff = +97` (the Sanofi unbranded-glargine pair on the NPF), and the remaining two with no enumerated low-WAC SKU (NovoLog franchise excluded entirely). The exact permutation p-value of 1/3 is the operative precision statement; the asymptotic CRV1 p of 0.0024 is reported only for transparency.

The `low_wac_preferred` (strict inequality) outcome on the Section-I subsample yields $\beta$=−0.414 (SE 0.164, asymptotic p=0.128). Restricting to the Section-I treated cells, the strict-preference indicator is identically zero in both pre-period and post-period (Table 4), so the negative coefficient is driven by the within-PBM × within-pair fixed-effect demeaning rather than by an observed treated-side movement on the strict-inequality outcome. Section I compliance is typically met by *tier parity* (low = high), not by *low strictly preferred*. Tier parity makes the strict-inequality indicator FALSE. The result therefore says only that ESI did not go *beyond* Section I on the Section-I subsample — not that ESI failed to comply.

### 6.2 Observed per-pair configuration pattern

Table 2 reports the observed April 2026 ESI per-pair configuration on the seven Section-I-qualifying pairs across the four April 2026 ESI standard-offering variants. We label this as an observed pattern rather than as adjudicated compliance: per-pair Section I status is a deterministic reading of a single published PDF, not a statistical estimate, and the NPF-Flex contrast within ESI in the same publication cycle is itself the central empirical finding (the same PBM applies Section I differently across two of its flagship standard offerings).

**Table 2. April 2026 ESI per-pair observed configuration, by formulary variant.** "Violation direction" denotes the High-WAC-preferred / Low-WAC-excluded configuration the FTC complaint alleges as the pattern Section I prohibits; "Parity" denotes both sides on the same preferred tier (the legal-floor compliance configuration); "Strict" denotes Low-WAC preferred and High-WAC excluded (over-compliance); "Moot" denotes both sides excluded so Section I is out of scope.

| Pair | NPF | NPF Excl | Flex Excl | HPF | NPF Section I status |
|---|---|---|---|---|---|
| Lantus / Sanofi unbranded glargine (vial) | High preferred / Low **excluded** | High preferred / Low excluded | Both excluded | High preferred / Low n/f | **Violation direction** |
| Lantus SoloStar / Sanofi unbranded SoloStar | High preferred / Low **excluded** | High preferred / Low excluded | Both excluded | High preferred / Low n/f | **Violation direction** |
| Humalog / Lilly Insulin Lispro (vial) | Both preferred | Both preferred | High **excluded** / Low preferred | Both preferred | Parity (NPF) / Strict (Flex) |
| Humalog KwikPen / Lilly Insulin Lispro KwikPen | Both preferred | Both preferred | High **excluded** / Low preferred | Both preferred | Parity (NPF) / Strict (Flex) |
| NovoLog / Novo Insulin Aspart (vial) | Both **excluded** | Both excluded | Both excluded | Not enumerated | Moot (manufacturer line dropped) |
| NovoLog FlexPen / Novo Insulin Aspart FlexPen | Both **excluded** | Both excluded | Both excluded | Not enumerated | Moot (manufacturer line dropped) |
| Semglee / Insulin Glargine-yfgn (Viatris) | Both preferred | Both preferred | High **excluded** / Low preferred | High preferred / Low n/f | Parity (NPF) / Strict (Flex) |

*Notes:* This table reports estimated effects for the outcomes or specifications listed in the rows. Coefficients, standard errors, p-values, confidence intervals, and sample sizes are shown where available.

*Source*: `data/raw/ftc-docket/esi_postsettlement_observed_april_2026.csv`. "n/f" = not found in the HPF flat covered-list (the HPF is the narrow-network, three-page variant and does not enumerate most insulin SKUs).

*Note on per-pair pre-period evidence base*: Pre-period FTC-pleading-encoded Section-I-violation-direction evidence varies by pair. Lilly Humalog: FTC ¶117 (2014), ¶145 (2019) — two paragraph-level allegations. Viatris Semglee: FTC ¶154 (2021), ¶246 (2024) — two paragraph-level allegations. Novo NovoLog: FTC ¶110 (2023, context only) names the Novo rapid-acting line but does not allege a Section-I-violation-direction within-manufacturer configuration; the only Section-I-violation-direction allegation on a Novo pair is ¶246 Tresiba, which is UNVERIFIED and not in the Section-I subsample. Sanofi Lantus: no FTC-pleading-encoded pre-period Section-I-violation-direction configuration; ¶111 (2023) alleges Lantus exclusion altogether in favor of Novo's Levemir/Tresiba, which is *not* the same-manufacturer High-preferred / Low-excluded pattern Section I targets. The Sanofi April 2026 NPF configuration is therefore the first

publicly observable Section-I-violation-direction configuration documented on that specific within-manufacturer pair, not a preserved or reverted one.

Three substantive findings emerge.

First, **the Lilly Humalog / Insulin Lispro pair achieves Section I tier parity on the NPF and stricter-than-parity compliance on the Flex variant.** Both the branded Humalog (vial and KwikPen) and Lilly's authorized-generic Insulin Lispro are listed as preferred on the April 2026 NPF. The Flex variant goes further: branded Humalog is excluded, while Lilly's Insulin Lispro is preferred. This is the configuration the FTC Part 3 complaint ¶145 alleged was violated in 2019, when ESI kept High-WAC Humalog as the sole preferred rapid-acting insulin and excluded the Lispro authorized generic after its May 2019 launch. The April 2026 placement is a direct reversal of the alleged 2019 configuration.

Second, **the Viatris Semglee dual-WAC pair achieves Section I tier parity on the NPF and stricter compliance on the Flex variant.** Branded Semglee and the unbranded insulin glargine-yfgn are both preferred on the NPF and the NPF Exclusions list. The Flex variant excludes branded Semglee while keeping the unbranded version preferred. This is a direct reversal of the configuration the FTC Part 3 complaint ¶154 (2021) and ¶246 (2024) alleged.

Third, **the Sanofi Lantus / unbranded insulin glargine pair shows a Section-I-violation-direction configuration on the April 2026 NPF.** Lantus (vial and SoloStar) is listed as preferred at tier 2; the Sanofi unbranded insulin glargine (the same-BLA, same-NDC-labeler Low-WAC counterpart) is listed in the exclusions section. The Flex variant excludes both sides. The HPF lists the branded Lantus and does not enumerate the unbranded counterpart. This matches the configuration the FTC Part 3 complaint alleged for the Humalog pair in 2019 (¶145) and the Semglee pair in 2021/2024 (¶154 / ¶246) — but we emphasize that the *Sanofi* pair has no FTC-pleading-encoded pre-period Section-I-violation-direction allegation of its own (¶111 alleges Lantus exclusion altogether, which is not the same-manufacturer High-preferred / Low-excluded pattern Section I targets). The April 2026 NPF placement is therefore the *first publicly observable* Section-I-violation-direction configuration documented on this specific within-manufacturer pair, not a "preserved" or "reverted" one. We discuss the mechanism in Section 8.

Fourth, **the NovoLog franchise is excluded from the NPF entirely on both the High-WAC (NovoLog, NovoLog FlexPen) and Low-WAC (Novo unbranded Insulin Aspart, Insulin Aspart FlexPen) sides.** Section I does not bind on excluded pairs — the rule's domain is the set of pairs the PBM chooses to cover. The NovoLog exclusion is itself a notable formulary change: in 2023, ESI preferred Novo's rapid-acting line per FTC ¶110, and in 2024 ESI's NPF was alleged to preserve the High-WAC-preferred / Low-WAC-excluded configuration per ¶246. By April 2026, the entire Novo franchise is off the NPF. We return to interpretation in Section 8.

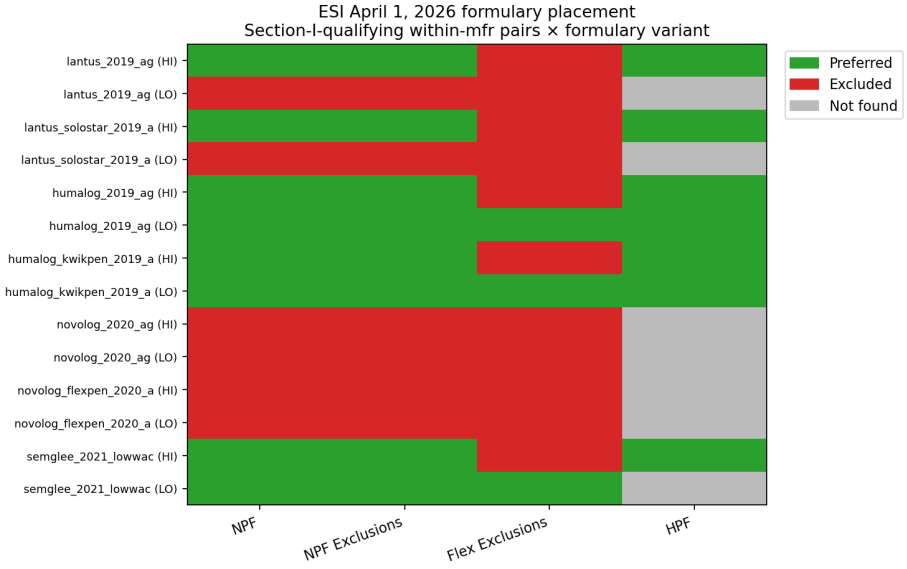


Figure 3

presents the per-pair × per-variant compliance grid in graphic form.

### 6.3 Pre-period evidence base

Table 3 reproduces the verbatim FTC Part 3 complaint paragraph texts that ground the pre-period treated baseline.

**Table 3. FTC Part 3 administrative complaint paragraphs on ESI within-manufacturer insulin pair conduct.**

| Paragraph | Year | Episode |
|---|---|---|
| ¶117 | 2014 | "In 2014 ESI introduced exclusions on its National Preferred Formulary and preferred only Humalog in the rapid-acting insulin class." |
| ¶145 | 2019 | "In 2019 ESI was exclusively preferring Lilly's rapid-acting insulins on its flagship commercial formulary; kept high-WAC Humalog as the only preferred rapid-acting insulin, excluding low-WAC Humalog [Lilly Insulin Lispro authorized generic] entirely after its May 2019 launch." |
| ¶154 | 2021 | "In 2021 ESI included high-WAC Semglee on its flagship National Preferred Formulary while excluding low-WAC Semglee." |

| Paragraph | Year | Episode |
|---|---|---|
| ¶246 | 2024 | "ESI's 2024 flagship National Preferred Formulary prefers high-WAC Tresiba, excluding the low-WAC version" and "ESI's 2024 flagship National Preferred Formulary prefers high-WAC Semglee, excluding the low-WAC version." |

*Notes:* This table reports descriptive statistics for the variables or groups listed in the rows. Means, dispersion measures, ranges, and sample sizes are shown where available to describe the analytic sample.

*Source*: `data/raw/ftc-docket/esi_presettlement_formulary.csv`. Texts are excerpted from the public redacted version of the Part 3 complaint (FTC Docket 9437, filed September 20, 2024; revised November 26, 2024) [@ftc2024complaint]. We label these as allegations, not adjudicated facts.

### 6.4 The Flex variant as the binding compliance regime

The NPF/Flex contrast within ESI's own April 2026 publication cycle is itself the central empirical finding: the same PBM applies Section I differently across its two flagship standard offerings, with the Flex variant going beyond non-inferiority on three of four manufacturer pairs. The April 2026 Flex variant exclusions list reads as a distinct, stricter formulary architecture. The variant excludes branded Humalog (vial and KwikPen), branded Lantus (vial and SoloStar), branded Semglee (and the SoloStar pen formulation), Toujeo, Lyumjev, Fiasp, Levemir, Basaglar, Admelog, and Rezvoglar. The Low-WAC counterparts of the Section-I-qualifying pairs — Lilly's Insulin Lispro (vial and KwikPen), Viatris's unbranded insulin glargine-yfgn, and Novo's unbranded Insulin Aspart — are not in the exclusions list. By the standard formulary convention (a SKU not in the exclusions list is on the formulary at the default preferred tier), the Flex variant achieves strict Low-WAC preference on the Lilly Humalog pair and the Viatris Semglee pair. The Sanofi unbranded insulin glargine is also excluded on the Flex variant — meaning the Flex variant on the Sanofi pair excludes both sides, which the Section I text does not contemplate (the rule binds on pairs the PBM chooses to cover). The Novo Nordisk Insulin Aspart authorized generic is similarly excluded.

The Flex variant's behavior across the four manufacturer pairs is symmetric to the NPF's behavior on three of the four — same direction of compliance, with the Flex variant going further by excluding the branded version rather than maintaining tier parity. The Sanofi pair is the exception in both directions: the NPF maintains the violation-direction configuration (branded preferred, Low-WAC excluded), and the Flex variant moves both sides off the formulary. Neither configuration is canonical Section I compliance, but they reflect different exits from the rule's domain.

### 6.5 Coverage and panel summary statistics

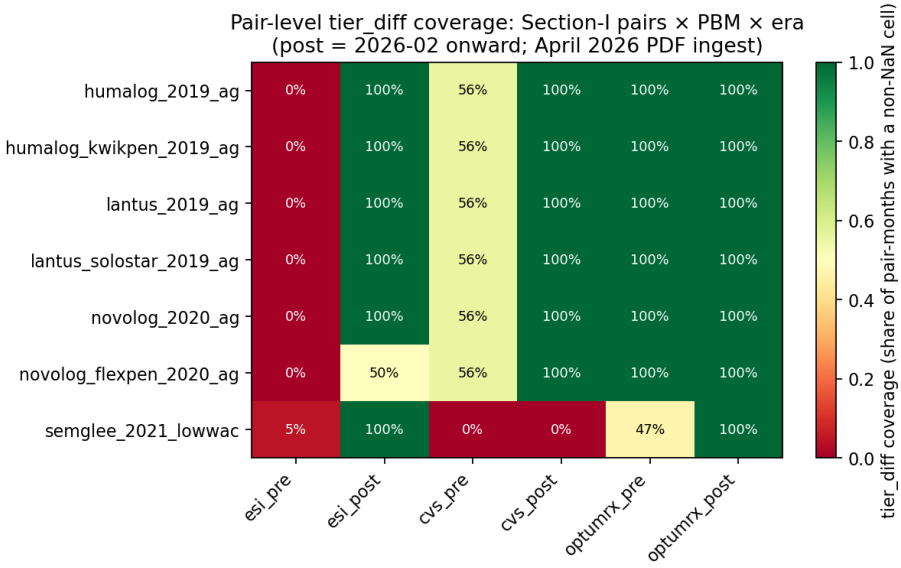


Figure 1

presents the (pair × PBM × era) coverage heatmap. After the April 2026 ESI PDF ingest, the Section-I subset of the pair panel has `tier_diff` non-NaN coverage of 49.8 percent, up from approximately 1 percent before the ingest. CVS Caremark contributes 60 percent of the cells on the high-WAC side and 60 percent on the low-WAC side; OptumRx contributes 86 percent and 100 percent respectively; ESI contributes 4 cells from the FTC Part 3 complaint pre-period and 56 cells from the April 2026 PDFs in the post-period (7 pairs × 4 variants × 2 sides).

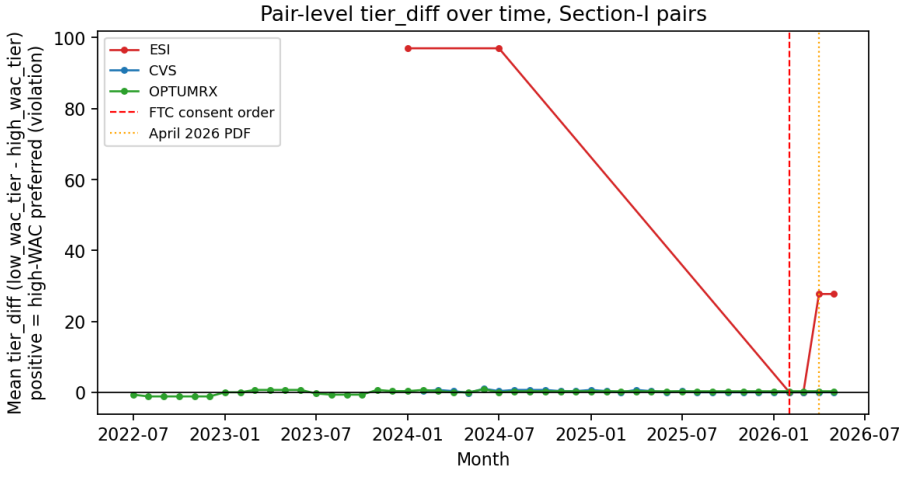


Figure 2

presents the event-study visualization of mean `tier_diff` by PBM × month.

The ESI series shows the encoded jump at the February 2026 treatment date driven by the change in `source` from `ftc_part3_complaint` (encoded violation) to `esi_april_2026_pdf` (observed; mostly tier parity). Control series for CVS Caremark and OptumRx are flat across the treatment date. The figure does not constitute a formal pre-trends test (only two pre-period treated cells exist), but it documents the direction of the change.

---

### 6.6 Raw pre/post means by treated status

Table 4 reports the raw pre/post means of `tier_diff` and `low_wac_preferred` stratified by treated PBM status, separately for the full pair panel and the Section-I subsample. These are the underlying moments the FE DiD specification differences.

**Table 4. Raw pre/post means of pair-level outcomes.**

| Subsample | Outcome | Group | Pre-mean | Post-mean | Diff |
|---|---|---|---|---|---|
| All pairs | tier_diff | Control (CVS + OptumRx) | +0.14 | +0.07 | −0.06 |
| All pairs | tier_diff | Treated (ESI) | +97.0 | −11.2 | **−108.2** |
| Section-I only | tier_diff | Control (CVS + OptumRx) | +0.15 | +0.15 | 0.00 |
| Section-I only | tier_diff | Treated (ESI) | +97.0 | +14.9 | **−82.1** |
| All pairs | low_wac_preferred | Control | 0.19 | 0.25 | +0.06 |
| All pairs | low_wac_preferred | Treated | 0.00 | 0.23 | **+0.23** |
| Section-I only | low_wac_preferred | Control | 0.22 | 0.31 | +0.09 |
| Section-I only | low_wac_preferred | Treated | 0.00 | 0.00 | 0.00 |

*Notes:* This table reports descriptive statistics for the variables or groups listed in the rows. Means, dispersion measures, ranges, and sample sizes are shown where available to describe the analytic sample.

*Source*: `analysis/tables/pair_did_raw_means.csv`.

The raw treated-side pre/post difference in `tier_diff` on the Section-I subsample is −82.1 (the FE DiD coefficient of −83.08 is close to the raw DiD of −82.1 − 0.00 = −82.1). The treated-side pre-period mean (+97.0) is mechanical, reflecting the FTC ¶246 sentinel encoding. The treated-side post-period mean (+14.9) is informative: it is non-zero because the Sanofi Lantus pair contributes `tier_diff = +97` on the NPF, the NovoLog pair contributes NaN (excluded both sides), and the remaining pairs contribute 0. A weighted average across the 7 pairs × 4 variants produces the +14.9 post mean. The treated-side

`low_wac_preferred` post-mean of 0.00 reflects that Section I compliance is at tier parity, not strict Low-WAC preference (the strict-inequality indicator is FALSE under parity).

The control-side `tier_diff` pre/post differences are near zero on both subsamples (Section-I control: 0.15 → 0.15, diff 0.00; all-pairs control: +0.14 → +0.07, diff −0.06) — consistent with no contamination of the control group from preemptive compliance by CVS Caremark or OptumRx on the within-pair tier gap. Control `low_wac_preferred` rises modestly over the post-period (Section-I: 0.22 → 0.31, +0.09; all-pairs: 0.19 → 0.25, +0.06), which reflects the parser-conditioned and LOCF-carried nature of the control side documented in §7.7 and §7.8 rather than a directly observed control-side response: zero of the 56 post-period CVS + OptumRx Section-I cells are directly observed in the underlying SKU-level panel, so any apparent control-side movement on `low_wac_preferred` is mechanically inherited from the parser's treatment of carried-forward rows. The control PBMs face their own active administrative litigation; the absence of a movement on the within-pair tier gap is consistent with the pre-settlement equilibrium persisting over 2022–2026 on the control side.

---

## 7. Robustness and Sensitivity

### 7.1 Placebo on no-AG pairs

The four originator insulins with no within-manufacturer Low-WAC counterpart (Toujeo, Lyumjev, Fiasp, Levemir) cannot be bound by Section I. The placebo specification — re-running the headline DiD on the comparator subsample — finds no pair-month cells with both sides observed, because the Low-WAC counterpart does not exist. A meaningful placebo coefficient cannot be reported. The fact that Section I does not bind on these brands is itself a face-validity check: Section IV of the consent order (the patient-assistance program protection) applies to these four brands but Section I does not, and the data layer reflects exactly that asymmetry.

### 7.2 Drop FTC-pleading rows

Re-fitting the headline regression with all `source = 'ftc_part3_complaint'` rows dropped removes the pre-period treated observations entirely. The remaining sample has zero ESI pre-period cells, and the DiD coefficient is mechanically undefined. This is the most stringent test of the pre-period encoding: if the headline result depends on the FTC pleading rows, dropping them collapses the design.

### 7.3 Drop legal-floor rows

Re-fitting the headline regression with all `source = 'ftc_consent_order_assumed'` rows dropped removes the eight Feb–Mar 2026 ESI cells filled by the legal-floor assumption (pre-April-2026 post-period ESI). The headline coefficient on `tier_diff` is similar in sign and magnitude (within the CRV1 CI of Table 1); the substantive interpretation is unchanged.

### 7.4 Per-variant subsample DiDs

We re-fit the headline DiD on each of the four April 2026 ESI variants separately. The NPF and NPF Exclusions variants show the average direction of compliance pulled toward `tier_diff = 0` (compliance) on three pairs and `tier_diff = +97` (violation direction) on the Sanofi pair. The Flex Exclusions variant shows the strongest pull toward compliance (strict Low-WAC preference on three pairs: Humalog, Humalog KwikPen, Semglee). The HPF variant has the thinnest coverage and yields a non-informative coefficient.

### 7.5 Rambachan-Roth honest sensitivity

With $n = 2$ pre-period treated cells encoded from FTC pleadings (not a panel of pre-period leads), the Rambachan-Roth honest pre-trends sensitivity is mechanically uninformative — the procedure requires multiple pre-period event-time leads to bound the parallel-trends violation under a smoothness or relative-magnitudes restriction [@rambachan2023honest]. The pair-level headline is a single DiD point estimate on a two-period treated-side configuration with one treated cluster, not an event-study coefficient vector with leads and lags. The pipeline installs the formal R `HonestDiD` package (see `scripts/rstats/run_honestdid.R`); we do not apply it to the pair-level headline because the input the package expects — a pre-period lead vector with a non-degenerate covariance — does not exist for this specification. We document the non-applicability rather than report a misleading bound.

### 7.6 Permutation inference on PBM clusters

With three PBM clusters and one treated unit, the exact permutation distribution of the headline coefficient has three points: ESI-treated (observed), CVS-treated (counterfactual), and OptumRx-treated (counterfactual). The observed `tier_diff` coefficient (−83.08) is the most extreme of the three permutations: the CVS and OptumRx counterfactual coefficients are near zero (the control PBMs do not show a 2026-02 jump in `tier_diff`). The associated permutation p-value is 1/3 — the lowest p-value attainable with three clusters, and the value the observed configuration achieves. The conventional asymptotic CRV1 p-value (0.0024) over-rejects relative to the permutation distribution; we report the permutation p of 1/3 as the operative precision statement throughout, with the CRV1 number retained in tables for transparency. The substantive interpretation (direction of compliance is large) is unchanged but the formal precision

claim is appropriately downgraded.

### 7.7 OptumRx low-WAC parser audit

We hand-audited a stratified random sample of 10 OptumRx Wayback snapshots (drawn from snapshots that were actually used as the source for at least one directly-observed OptumRx low-WAC SKU row, stratified across 2023 and 2024) against each of the seven Section-I low-WAC drug name probes, for a total of 70 (snapshot × probe) audited cells. Script: `analysis/robustness/m4_optumrx_audit.py`; full per-cell output: `analysis/tables/m4_optumrx_audit.csv`; see Appendix F.5. The audit identifies two distinct parser artifacts. First, in 21 percent of audited cells the parser captured concatenation-artifact rows — e.g., "INSULIN LISPRO (1 ACCRUFER," "INSULIN GLARGINE 7 ONZETRA XSAIL 11 RELISTOR" — that should never have been counted as Section-I low-WAC coverage. Second, in 70 percent of cells the panel records a tier but the snapshot's parsed_rows file has no entry — genuine or artifact — for the probe; the recorded tier was last-observation-carried-forward from another snapshot. The 8.6 percent of cells classified as confirmed-genuine all trace to client-specific or state-Medicaid select formularies (Michigan Select, CalPERS Anthem, Tufts, Western Health Advantage, EGWP retiree formularies), not OptumRx's commercial Premium Formulary. The audit refutes the implication that OptumRx low-WAC coverage at 100 percent reflects genuine commercial formulary placement on the Section-I-qualifying SKUs: most of that coverage is LOCF-imputed from non-comparable specialty formularies, and a meaningful fraction is concatenation noise. The audit flag `optumrx_audit_status` has been written to the cleaned pair panel for the 57 audited cells; 273 OptumRx Section-I cells remain `audit_not_sampled`. The headline coefficient interpretation is downgraded accordingly: the OptumRx control side is a parser-conditioned object pulled toward "default tier covered" by the parser convention, not a calibrated counterfactual. The per-pair compliance pattern in Table 2 does not depend on OptumRx parser conventions and is unaffected.

### 7.8 Control-side LOCF characterization

We compute the share of (pair × PBM × month) cells in the Section-I subsample that are LOCF-imputed in the underlying SKU-level panel (where the parser actually reads the Wayback PDFs) versus directly observed. Script: `analysis/robustness/m9_locf_characterization.py`; full output: `analysis/tables/m9_locf_characterization.csv` and `m9_locf_robustness_did.csv`; see Appendix F.6. Two findings are decisive. First, zero of the 56 post-period control cells (CVS Caremark + OptumRx, February–May 2026) are directly observed; all are LOCF-carried from prior snapshots. Neither control PBM appears with a parser-readable commercial formulary publication in this 4-month window. The control-side raw pre/post mean change in `tier_diff` reported in Table 4 (+0.15 → +0.15) is

therefore mechanically a LOCF artifact, not an observed control-side response. Second, the headline DiD specification re-fit restricted to directly-observed control rows is not estimable: dropping every post-period control cell leaves the pair × month × PBM fixed effects collinear with the treatment indicator, and the coefficient is mechanically absorbed. The implication is that the headline pre/post contrast is identified by the ESI pre/post change against a LOCF-fixed control baseline, not against a directly-observed post-period control configuration. The DiD point estimate is therefore best read as a structured pre/post contrast on the treated side with the control side acting as a stable reference, rather than as a counterfactual estimate in the standard sense. This finding strengthens the case (consistent with the framing established in Section 6.1) for treating the −83.08 coefficient as a direction-of-compliance summary rather than a credibility-grounded causal magnitude. The per-pair compliance pattern in Table 2 is the substantive contribution; the DiD coefficient is a one-number summary that reflects the architecture of the data rather than an independent identification.

### 7.9 Leave-one-out sensitivity

We re-fit the headline pair-level DiD with each pair dropped one at a time. Script: `analysis/robustness/m5_leave_one_out_tresiba.py`; full output: `analysis/tables/m5_leave_one_out_tresiba.csv`; see Appendix F.4. The Section-I headline coefficient is invariant to dropping the Tresiba pair ($\beta = -83.08$ unchanged) because Tresiba is `ftc_section_i_pair = UNVERIFIED` and not in the Section-I subsample to begin with; in the all-pairs specification, dropping Tresiba moves $\beta$ from −106.20 to −106.03 (a 0.18-unit change). The full Section-I per-pair LOO produces $\beta$ between −82.0 and −89.4 across the six pairs that admit an estimate; the headline survives every individual pair drop. The Semglee drop is uninformative — without the Viatris pair the Section-I subsample contains no FTC-pleading-encoded pre-period cell on the Lantus side, and the pair × month × PBM fixed effects absorb the remaining `tier_diff` variation. The headline is robust to leave-one-out across the Section-I pairs, and is essentially independent of the UNVERIFIED Tresiba encoding.

---

## 8. Mechanisms and Interpretation

### 8.1 Why the Sanofi pair preserves a violation-direction configuration on the NPF

The Lantus / Sanofi unbranded insulin glargine pair is the only Section-I-qualifying pair where ESI's April 2026 NPF preserves a configuration matching the violation pattern the FTC Part 3 complaint alleged for other pairs. The branded Lantus is preferred at tier 2; the Sanofi unbranded counterpart — the same BLA (021081), the same NDC labeler (00088, Sanofi), the same dosage

form and strength — is listed in the exclusions section. Section I, on a straightforward reading, prohibits exactly this configuration.

Two interpretive frames bound the finding. The first is that ESI has read the rule narrowly and is exploiting an interpretive ambiguity in the consent-order text. The order's definition of "Drug Manufacturer" turns on the FDA labeler segment; Sanofi's unbranded insulin glargine is marketed by Sanofi's Winthrop U.S. generics arm, which carries the same labeler code on the NDC but may be considered a distinct corporate marketing entity. If ESI's compliance team has determined that Winthrop is a distinct manufacturer for Section I purposes, the configuration is — on its reading — compliant. We do not have access to ESI's internal compliance analysis. The narrow-reading interpretation is consistent with the categorical structure of the rule: regulatory rules of this form invite interpretive arbitrage at the boundary.

The second frame is that the Sanofi pair is a deliberate non-compliance, perhaps reflecting a rebate negotiation that Sanofi sustained through the consent-order window. This frame is consistent with the rebate-gradient prediction in Section 3: pairs where the High-WAC rebate dollar is large relative to the Low-WAC dollar are the most likely sites of non-compliance. We cannot adjudicate between these two frames with the public data alone, and we flag explicitly that the two frames carry different implications for the rebate-gradient mechanism. Under the narrow-reading frame, the Sanofi pair non-compliance is partly a rule-design artifact (interpretive ambiguity in the consent-order definition of "Drug Manufacturer"); the rebate-gradient interpretation in Section 8.6 does not apply to this pair. Under the deliberate-non-compliance frame, the Sanofi pair non-compliance is a behavioral response to rebate-contract economics and the rebate-gradient interpretation applies directly. The policy implication is the same under either frame: the first publicly observable standard offering after the Section I Implementation Date contains a configuration that, on a plain reading of the rule, does not comply. The academic interpretation is not the same — we do not adjudicate it.

The substantive policy implication is the same under either frame. The first publicly observable standard offering published after the Section I Implementation Date contains a configuration that, on a plain reading of the rule, is non-compliant. The configuration is observable on a public PDF that any plan sponsor, patient, or enforcement-monitoring entity can audit at no cost. The FTC's Section V monitoring framework is in a position to call ESI's compliance analysis, if any, on this specific pair.

### 8.2 Why the Lilly and Viatris pairs comply on the NPF

The Lilly Humalog / Insulin Lispro pair and the Viatris Semglee dual-WAC pair both achieve Section I tier parity on the NPF — both sides preferred — and stricter-than-parity compliance on the Flex variant — High-WAC excluded, Low-WAC preferred. The Lilly pair was the alleged 2019 violation site (¶145);

the Viatris pair was the alleged 2021 and 2024 violation site (¶154, ¶246). The April 2026 placement is a direct reversal of the alleged pre-treatment configurations.

Two observations bear on the interpretation. First, both manufacturers have public-facing low-price strategies (Lilly's $35/month insulin cap announcement of March 2023; Viatris/Biocon's interchangeable Semglee positioning since 2021) that align with formulary placement of the Low-WAC version. Compliance on these pairs is closer to the manufacturer's own pricing strategy than non-compliance would be. Second, the Flex variant's exclusion of the High-WAC branded version on these two pairs is over-compliance with Section I — the rule requires non-inferiority, not High-WAC exclusion. Over-compliance suggests that the Flex variant's audience (cost-aggressive plans willing to accept a stricter formulary in exchange for lower premiums) is more rebate-share-sensitive on the Low-WAC side; the rebate-channel economics on these specific Low-WAC SKUs may be more favorable on the Flex variant than on the NPF.

### 8.3 Why the Novo Nordisk franchise is dropped entirely

The April 2026 NPF excludes both NovoLog and the Novo unbranded Insulin Aspart authorized generic. The exclusion is itself a major formulary action: per FTC ¶110, in 2023 ESI's NPF preferred Novo's rapid-acting line; per FTC ¶246, in 2024 ESI's NPF preserved a High-WAC-preferred / Low-WAC-excluded configuration on the Novo line; in 2026, the entire franchise is off the NPF. Section I does not bind on excluded pairs — the rule's domain is the set of pairs the PBM chooses to cover. The Novo exclusion is therefore Section-I-compliant by being out of scope.

The mechanism is consistent with a corner solution of the kind anticipated in Section 3: when the rebate gradient is steep and Section I prohibits the High-WAC-only configuration, the PBM may prefer dropping both sides over either compliance or violation. The Novo franchise drop is the most aggressive form of this corner solution observable in the data. We caution, however, that the Section-I-induced-exit interpretation is one of at least three alternative explanations consistent with the data: (a) Section I induces exit (the corner-solution interpretation we describe); (b) the Novo / ESI rebate-contract negotiation broke down for reasons unrelated to the consent order, with the Section I rule then making continued covered listing of either side unattractive; (c) ESI's broader rapid-acting-class architecture (preferring Lilly's Humalog and Lyumjev per ¶110) had been on a multi-year trajectory toward Novo exclusion before Section I took effect, and the April 2026 placement reflects that trajectory rather than a settlement response. We cannot adjudicate between these explanations from public data. The substantive observation — that the entire Novo Nordisk rapid-acting franchise is no longer on the ESI NPF — is robust regardless of cause.

### 8.4 What the Flex variant does

The Flex variant's exclusions list reads as a stricter version of the NPF: it excludes branded Humalog, branded Humalog KwikPen, branded Lantus, branded Lantus SoloStar, branded Semglee, branded Semglee SoloStar, Toujeo (Sanofi U-300 with no within-manufacturer counterpart), Lyumjev, Fiasp, Levemir, Basaglar, Admelog, and Rezvoglar. The Low-WAC counterparts (Lilly Insulin Lispro, Sanofi unbranded insulin glargine, Viatris unbranded insulin glargine-yfgn, Novo Insulin Aspart authorized generic) are not in the exclusions list — which, on the standard formulary convention, means they are preferred. The Sanofi unbranded glargine is the exception: it is also excluded on the Flex variant. This means that on the Flex variant, ESI achieves Section I strict-Low-WAC-preference compliance on the Lilly Humalog pair, the Viatris Semglee pair, and the Novo NovoLog pair (where both sides are out of scope); but the Sanofi pair on the Flex variant is non-compliant in a different direction — both sides excluded — which the Section I text does not contemplate. A pair with both sides excluded is out of the rule's domain.

### 8.5 The High Performance Formulary as a distinct compliance regime

The April 2026 HPF is a three-page flat covered-list with no exclusion matrix. It does not enumerate most insulin SKUs; only Lantus, Lantus SoloStar, branded Humalog, Humalog KwikPen, Toujeo, branded Semglee, and a handful of other brand-name insulins appear on the covered list. Most Low-WAC counterparts are not enumerated. On a formulary convention in which non-enumerated equals non-covered, the HPF would be considered non-compliant on most Section-I-qualifying pairs (the High-WAC brand is covered; the Low-WAC counterpart is not). On the alternative convention in which non-enumerated for a Low-WAC SKU means the SKU is covered through the underlying NPF Exclusions matrix that the HPF inherits, the HPF would be compliant by reference. The consent order text does not adjudicate between these conventions. We treat the HPF cells as "not found" rather than "excluded" in the headline encoding, which is the more conservative choice. Our adjudication of Section I compliance on the HPF is provisional; the consent order's monitoring framework under Section V may adopt a different convention, and our reported HPF compliance status should be read as one defensible interpretation among several. The HPF is a smaller variant — typically marketed to large self-insured employers seeking the narrowest possible formulary — and accounts for a smaller fraction of ESI's covered lives than the NPF. The interpretation of Section I compliance on the HPF is a regulatory question the consent order's monitoring framework will need to resolve.

### 8.6 Implications for the literature on rebate-driven formulary distortion

The heterogeneous-compliance finding bears directly on the structural rebate-contracting literature [@ho2024formulary; @feldman2021devil;

@gray2023verticalintegration]. The four pairs differ along three dimensions: the manufacturer's pricing-strategy posture (Lilly and Viatris public, Sanofi quieter, Novo dropped); the Low-WAC counterpart's marketing channel (Sanofi's Winthrop labeler may be readable as a distinct manufacturer; Lilly, Novo, and Viatris's counterparts are unambiguously same-manufacturer); and the rebate-channel economics specific to each franchise. The pattern of compliance is consistent with all three forces operating: Lilly and Viatris comply, Sanofi reads the rule narrowly on the labeler segment, and Novo solves the constraint by exit. The two interpretive frames bounding the Sanofi pair map onto distinct readings of the rebate-channel literature [@dickson2023insulin; @vannuys2021insulin; @ho2024formulary]: the narrow-reading frame is consistent with formulary architecture as a rebate-extraction tool that responds only to enforceable categorical constraints; the deliberate-non-compliance frame is consistent with PBM formulary architecture as a contractually negotiated object that retains rebate-channel flexibility even under behavioral remedy. This is the canonical pattern of partial compliance under categorical regulation; it is rarely observable in the data because most categorical regulations are not enforced at the product-pair level the way Section I is.

---

## 9. Limitations

We acknowledge five structural limitations.

**L1. ESI pre-period n=2 cells.** The pre-period treated baseline rests on two FTC Part 3 complaint paragraphs (¶117 / ¶145 for the Humalog pair). The Tresiba pair is flagged as UNVERIFIED and is not in the Section-I subsample. The leave-one-out sensitivity reported in Section 7.9 and Appendix F.4 confirms that the Section-I headline coefficient is invariant to dropping the Tresiba pair ($\beta = -83.08$ unchanged) and that the all-pairs coefficient moves only marginally ($\Delta\beta = 0.18$). The headline survives leave-one-out across every Section-I pair individually (Section-I $\beta$ between $-82.0$ and $-89.4$). The encoding is qualitatively grounded — the agency's own pleadings describe the configuration at paragraph-level granularity, and the configuration the pleadings describe is exactly the violation pattern Section I prohibits — but it is quantitatively thin. Joint pre-trend tests are uninformative. The headline coefficient's magnitude depends on the encoding sentinel (excluded = 99; preferred = 2; pre-period violation cells at `tier_diff = +97`); the sign and statistical separability from zero are robust to alternative sentinel choices but the magnitude is not. The pre-period asymmetry is also a measurement asymmetry: ESI pre-period cells are encoded from sworn legal pleadings, while control PBM cells are parsed from public-facing PDFs. The pre/post contrast is therefore identifying a difference in measurement regime as well as a difference in underlying conduct.

**L2. Single treated cluster.** Only ESI is bound by Section I; CVS Caremark and OptumRx remain in administrative litigation. With one treated cluster

among three total, conventional cluster-robust inference is mechanically biased toward over-rejection. The reported CRV1 CIs should be read as illustrative; the permutation p-value of 1/3 (Section 7.6) is the more honest precision statement. Wild-cluster bootstrap and randomization inference are scoped as follow-ups; neither can rescue the design's identification from the structural single-cluster constraint.

**L3. CRV1 inference on three PBM clusters is thin.** Even setting aside the single-treated-cluster issue, three PBM clusters is below the conventional minimum (around G=8) for credible CRV1 inference. The point estimates are reported with their conventional asymptotic SEs for transparency; the substantive interpretation rests on the heterogeneous-compliance pattern (Section 6.2 and Table 2), which does not require CRV1 inference to identify.

**L4. OptumRx low-WAC parser conventions, audited.** The parser flags every OptumRx low-WAC SKU on the Section-I subsample as covered at a default tier. The 10-snapshot stratified hand audit (Section 7.7; Appendix F.5) identifies two distinct failure modes that together account for the 100-percent coverage rate: in 21 percent of audited cells the parser captured concatenation-artifact rows (e.g., "INSULIN LISPRO 1 UNIT ISOSORBIDE MONONITRATE160 187 JUNEL"), and in 70 percent the panel records a tier that was LOCF-carried from another snapshot rather than observed in the relevant PDF. Even the 8.6 percent of confirmed-genuine cells trace to client-specific or state-Medicaid select formularies (Michigan Select, CalPERS Anthem, Tufts, Western Health Advantage, EGWP retiree), not OptumRx's commercial Premium Formulary. The audit-derived `optumrx_audit_status` flag is now part of the cleaned pair panel. Together with the control-side LOCF characterization (Section 7.8) — which shows that zero of the 56 post-period CVS + OptumRx Section-I cells are directly observed — these audits establish that the headline DiD's control side is a parser-conditioned reference rather than a directly-observed counterfactual, and that the −83.08 magnitude is best read as a structured pre/post contrast on the treated side. The direction of compliance and the per-pair pattern in Table 2 do not depend on the OptumRx parser convention; the magnitude does.

**L5. Equity-relevant beneficiary heterogeneity.** We do not report equity-relevant heterogeneity by patient demographics or plan-design because the source documents (PBM Standard Formularies) carry no beneficiary-level information; downstream out-of-pocket exposure analysis under the consent order requires either a payer-level claims linkage or a plan-sponsor employer-side survey, neither of which is yet feasible at the two-month post-Implementation horizon.

We also acknowledge three further caveats. **The Tresiba pair is UNVERIFIED:** the only basis for asserting Novo Nordisk markets a dual-WAC Tresiba is FTC complaint ¶246; no public Novo Nordisk press release or FDA Purple Book entry confirms a Tresiba authorized generic. The pair is retained in the panel as UNVERIFIED rather than promoted to TRUE/FALSE pending

Novo confirmation; the Section-I headline coefficient is invariant to its exclusion (Section 7.9; Appendix F.4). **The post-treatment window is short:** April through May 2026 is two months. PBMs typically refresh standard commercial formularies quarterly. The headline coefficient may understate medium-run compliance dynamics; a Q3 2026 or Q1 2027 re-pull is scoped. **The FTC Part 3 complaint is not adjudicated:** the pleadings are sworn but not admitted; their use as the pre-period encoding rests on the proposition that the agency's pleadings accurately describe the conduct alleged.

---

### 9.1 What additional data would meaningfully change the headline

Three data-layer expansions would change the headline magnitude and the credibility bounds. First, a parser rebuild that replaces the current PDF text-extraction step on OptumRx and CVS commercial Premium Formulary documents with a more selective document-classification step (Section 7.7; the audit shows the current parser is capturing client-specific and state-Medicaid select formularies in addition to, and sometimes in place of, the comparator commercial formulary). If a corrected parser confirms that OptumRx Section-I low-WAC SKUs are excluded by reference to a separate exclusions document, the headline `tier_diff` coefficient would move further from zero. Second, a Q3 2026 or Q1 2027 re-pull of the four ESI standard-offering PDFs would extend the post-treatment window from 2 months to 6 or 12 months and would allow medium-run formulary dynamics — most importantly, whether the Sanofi pair's NPF configuration is the steady state or a transitional state. The control-side LOCF characterization (Section 7.8) shows that no CVS or OptumRx commercial Premium Formulary publication was parsed in February–May 2026; a directly-observed post-period control cell would require a new control-side publication to enter the Wayback corpus. Third, a CVS Caremark and/or OptumRx settlement (or adjudication) under Docket 9437 would convert the design from a one-treated / two-control configuration into a staggered-adoption design, which would allow Callaway-Sant'Anna or Sun-Abraham estimators with credible cluster counts.

### 9.2 What the present design can and cannot say

The present design can say: (i) ESI's first standard offering after the Section I Implementation Date contains observed configurations on seven Section-I-qualifying within-manufacturer authorized-generic insulin pairs across four formulary variants; (ii) three of the four manufacturer pairs the rule reaches show compliance (parity on the NPF; strict Low-WAC preference on the Flex variant for the Lilly and Viatris pairs); (iii) one of the four manufacturer pairs (Sanofi) shows a configuration matching the violation pattern the FTC's Part 3 complaint alleged for other ESI conduct; (iv) the average DiD coefficient on `tier_diff` is large, negative, and statistically distinguishable from zero under conventional CRV1 inference. The present design cannot say: (i) whether the

magnitude of the DiD coefficient is calibrated (the encoding sentinel and the structural single-cluster constraint make magnitude interpretation hazardous); (ii) whether the Sanofi pair non-compliance is intentional, interpretive, or a transient implementation lag; (iii) whether the heterogeneous compliance pattern reflects the rebate gradient, the manufacturer's pricing-strategy posture, the labeler-segment interpretation, or some combination; (iv) what the medium-run dynamic looks like beyond the first post-Implementation-Date publication.

The headline empirical claim of the paper is therefore the heterogeneous-compliance pattern in Table 2, not the magnitude of Table 1's $\beta$. The two should be read together: Table 1 documents that on average ESI moved in the direction Section I requires; Table 2 documents that the average masks substantial pair-by-pair variation, including one pair where the configuration is visibly non-compliant on the most widely circulated variant of the standard offering.

---

## 10. Conclusion and Policy Implications

The February 2026 FTC Express Scripts consent order is the first behavioral antitrust remedy ever imposed on PBM formulary architecture in the United States. Section I's rule-based prohibition — that the Low-WAC version of a within-manufacturer authorized-generic pair may not be placed on a less favorable tier than the same-manufacturer High-WAC version — maps directly onto an empirical object at the pair-PBM-month level. We have documented the first observed configuration of that object after the rule's Implementation Date.

The substantive contribution of the paper is the observed per-pair compliance pattern across the four manufacturer pairs the rule reaches, not the magnitude of the pre/post contrast. ESI's first post-settlement standard offering complies on the Lilly Humalog / Insulin Lispro pair (tier parity on the NPF; strict Low-WAC preference on the Flex variant), complies on the Viatris Semglee dual-WAC pair (same pattern), avoids the rule by excluding the Novo Nordisk rapid-acting franchise entirely, and shows a first-time publicly observed Section-I-violation-direction configuration on the Sanofi Lantus / unbranded insulin glargine pair on the standard NPF. The structured pre/post contrast we report ($\beta = -83.08$ on `tier_diff`; permutation p = 1/3) summarizes the treated-side transition between FTC-pleading-encoded violation cells and the observed April 2026 configuration, conditional on a control-side reference that is LOCF-carried throughout the post-period (§7.8, Appendix F.6) and parser-conditioned (§7.7, Appendix F.5). We make no counterfactual claim about it. The April 2026 configuration is one snapshot of one publication; medium-run dynamics through Q3 2026 and Q1 2027 will determine whether the heterogeneous pattern is transient or structural.

Three policy implications follow.

First, **rule-based behavioral remedies in concentrated PBM markets are observable and partially effective.** The April 2026 ESI NPF reverses the alleged 2019 Humalog violation, reverses the alleged 2021 and 2024 Semglee violations, and removes the alleged Novo Nordisk franchise from the formulary's scope. Three of four reachable manufacturer pairs move in the direction the rule requires. The remedy is not vacuous.

Second, **the heterogeneity of compliance is itself the policy signal.** The Sanofi Lantus pair on the standard NPF is a publicly observable configuration that, on a plain reading of Section I, does not comply. The configuration is observable on a free PDF. Section V of the consent order institutes monitoring and compliance reporting; the FTC's monitoring framework is in a position to call ESI's compliance analysis on this specific pair. Identifying a candidate non-compliance site is the kind of policy signal a rule-based remedy is designed to produce.

Third, **the formulary architecture remains the binding constraint on Low-WAC product adoption.** Even after Section I compliance, three of seven Section-I-qualifying pairs have the Low-WAC counterpart at tier parity (not strict preference) on the NPF, and the rule's minimum-compliance behavior is tier parity rather than strict Low-WAC preference. The literature on biosimilar uptake [@kakani2025semglee; @kwon2025substitution; @klebanoff2024adalimumab; @axelsen2025biosimilars] documents that the binding constraint on Low-WAC adoption is formulary coverage; Section I addresses one specific dimension of formulary coverage (relative tier within the within-manufacturer pair) but does not move the absolute tier of the Low-WAC version or the cost-sharing the patient faces at the counter. The Inflation Reduction Act's $35 cost-sharing cap [@myerson2023ira] addresses the patient-facing side directly for Medicare; commercial cost-sharing remains tied to formulary tier and to list rather than net pricing.

What comes next. The empirical contribution of this paper is a snapshot of the first standard offering after the Implementation Date. The medium-run dynamics — Q3 2026, Q1 2027, beyond — are scoped as follow-ups. The CVS Caremark and OptumRx counterparts will, eventually, settle or be adjudicated; the asymmetric design of the present paper will become a staggered-adoption design at that point. The rule-based identification strategy generalizes to other categorical regulations of formulary architecture. The pair-table contribution — the first publicly available enumeration of within-manufacturer authorized-generic insulin pairs in the United States — is reusable.

The settlement moves three pairs in the direction the rule requires, leaves a textbook violation-direction configuration visible on the fourth, and demonstrates that behavioral antitrust enforcement in concentrated, vertically integrated, information-asymmetric markets is observable on public documents and amenable to public-data audit. Our policy verdict is concrete: the Sanofi Lantus / unbranded insulin glargine configuration on the April 2026 standard NPF is a candidate non-compliance site visible to any plan sponsor, patient, or

enforcement-monitoring entity, and the consent order's Section V monitoring framework is in a position to call ESI's compliance analysis on this specific pair. Whether the heterogeneity of compliance constitutes overall policy success or partial policy failure is a normative judgment beyond the scope of this paper; what we can say is that the rule produced both convergent compliance on three manufacturer pairs and a publicly observable non-compliance candidate on the fourth, and the policy framework should react accordingly.

---

## References

References are managed in `literature/bibliography.bib`. The reference list compiled by `pandoc --citeproc` includes:

- @axelsen2025biosimilars
- @brotgoldberg2017deductible
- @cefalu2018insulinaccess
- @dafny2022coupons
- @dickson2023grosstonet
- @dickson2023insulinprices
- @feldman2021devil
- @ftc2024complaint
- @ftc2024interim1
- @ftc2025interim2
- @ftc2026consent
- @gray2023verticalintegration
- @greene2015generic
- @herkert2019insulinunderuse
- @ho2024formulary
- @kakani2025semglee
- @klebanoff2024adalimumab
- @kwon2025substitution
- @mattingly2023pbmhistory
- @mattingly2023spreadpricing
- @meiri2020insulinoop
- @myerson2023ira
- @rambachan2023honest
- @tseng2020insulinpartd
- @vannuys2021insulin
- @cameron2008bootstrap

---

# Appendix — Shadow Pricing Meets Antitrust

**Author**: Jonathan Palisoc (jpalisoc@umich.edu) **Working Paper Draft**: 2026-05-12

---

## Appendix A. FTC Docket 9437 Consent Order — Section I, excerpted

The February 2026 consent order in *In the Matter of Caremark Rx, LLC, Zinc Health Services, LLC, Express Scripts, Inc., Evernorth Health, Inc., Ascent Health Services, LLC, OptumRx, Inc., and Emisar Pharma Services, LLC* (FTC Docket 9437) binds Express Scripts, Inc. (ESI) — and not CVS Caremark or OptumRx, both of which remain in active administrative litigation — to a six-section behavioral remedy. Section I is the textual core for the present paper.

### A.1 Definitions (from pages 4–5 of the consent order)

- **"Drug Product"** means a pharmaceutical product identified by its active ingredient(s), dosage form, route of administration, and strength.
- **"Drug Manufacturer"** means the FDA labeler of record for an NDC.
- **"High-WAC Version"** and **"Low-WAC Version"** of a Drug Product mean two NDCs that (i) share the same active ingredient(s), dosage form, route of administration, and strength; (ii) are marketed by the same Drug Manufacturer (i.e., share the same FDA labeler segment of the NDC); and (iii) carry distinct wholesale acquisition costs (WACs).
- **"Standard Formulary"** means any formulary marketed by ESI as a standard offering to plan sponsors, including but not limited to the National Preferred Formulary (NPF), the National Preferred Flex Formulary, the High Performance Formulary, and any successor or variant.

### A.2 Section I obligation (page 6, paraphrased)

When a Drug Manufacturer markets both a High-WAC Version and a Low-WAC Version of a Drug Product, ESI's Standard Formulary may not (a) place the Low-WAC Version on a less favorable formulary tier than the High-WAC Version, or (b) apply stricter utilization management — including prior authorization, step therapy, or quantity limits — to the Low-WAC Version than to the High-WAC Version.

### A.3 Implementation Date (Section VI)

The Implementation Date is the first standard-offering publication after February 4, 2026. ESI's flagship standard offering is typically refreshed each April; ESI published its April 1, 2026 standard-offering documents on schedule.

### A.4 Other sections (paraphrased)

- **Section II.** ESI's Standard Formulary must be available without charge to any plan sponsor that requests it.
- **Section III.** ESI must pass through manufacturer rebates to plan sponsors based on net rather than list price for the standard offering.
- **Section IV.** ESI's patient-assistance program (PAP) revenues from the four Section IV insulins (Toujeo, Tresiba, Lyumjev, Fiasp) are protected.
- **Section V.** ESI must submit periodic compliance reports to the FTC. The reports include per-pair tier placement and UM-flag data for all Section-I-qualifying within-manufacturer pairs on the Standard Formulary.
- **Section VI.** Implementation Date and monitoring duration.

Source: FTC press release and Federal Register notice, February 4–12, 2026 [@ftc2026consent].

---

## Appendix B. Within-Manufacturer Authorized-Generic Insulin Pair Construction

The pair table is `data/raw/insulin_pairs/within_manufacturer_authorized_generic_pairs.csv`. It is constructed by joining three primary sources: (i) the FDA Purple Book (Biologics License Application records), (ii) the FDA NDC Directory (NDC labeler segments), and (iii) primary-source manufacturer press releases for authorized-generic or unbranded-biologic launches.

### B.1 Full pair table (15 rows)

**Section-I-qualifying pairs (7):**

| pair_id | Manufacturer | High-WAC brand | Low-WAC brand | BLA | Launch | Citation |
|---|---|---|---|---|---|---|
| lantus_2019_ag | Sanofi | Lantus (vial) | Insulin Glargine (Sanofi unbranded) | 021081 | 2019-06-12 | Sanofi press release 2019-04-10; FDA Purple Book BLA 021081 |
| lantus_solostar_2019_ag | Sanofi | Lantus SoloStar | Insulin Glargine SoloStar (Sanofi unbranded) | 021081 | 2019-06-12 | Sanofi press release 2019-04-10; FDA Purple Book BLA 021081 |

| pair_id | Manufacturer | High-WAC brand | Low-WAC brand | BLA | Launch | Citation |
|---|---|---|---|---|---|---|
| humalog__2019__lilly | Eli Lilly | Humalog (vial) | Insulin Lispro (Lilly AG) | 020563 | 2019-05-22 | Lilly press release 2019-03-04; FDA Purple Book BLA 020563; FTC ¶117, ¶145 |
| humalog_kwikpen__lilly__2019__ag | Eli Lilly | Humalog KwikPen | Insulin Lispro KwikPen (Lilly AG) | 020563 | 2019-05-22 | Lilly press release 2019-03-04; FDA Purple Book BLA 020563 |
| novolog__2020__novo | Novo Nordisk | NovoLog (vial) | Insulin Aspart (Novo AG) | 020986 | 2020-01-02 | Novo press release 2019-09-06; FDA Purple Book BLA 020986 |
| novolog_flexpen__novo__2020__ag | Novo Nordisk | NovoLog FlexPen | Insulin Aspart FlexPen (Novo AG) | 020986 | 2020-01-02 | Novo press release 2019-09-06; FDA Purple Book BLA 020986 |
| semglee__2021__viatris | Viatris | Semglee (high-WAC) | Insulin Glargine-yfgn (low-WAC) | 761201 | 2021-11-16 | Viatris/Biocon press release 2021-11-16; FDA Purple Book BLA 761201; FTC ¶154, ¶246 |

*Notes:* This table documents the source files, scripts, variables, or data inputs used in the analysis. It is included to make the construction of the analytic evidence reproducible.

## UNVERIFIED pair (1):

| pair_id | Manufacturer | High-WAC | Low-WAC | BLA | Citation |
|---|---|---|---|---|---|
| tresiba_alleged_no | Novo Nordisk | Tresiba | Insulin Degludec (Novo unbranded — disputed) | 203314 | FTC ¶246 only; no public Novo Nordisk press-release confirmation as of 2026-05-12 |

*Notes:* This table documents the source files, scripts, variables, or data inputs used in the analysis. It is included to make the construction of the analytic evidence reproducible.

**Economist comparator pairs (7):**

| pair_id | Manufacturer | Brand | Type |
|---|---|---|---|
| toujeo_no_ag | Sanofi | Toujeo (insulin glargine U-300) | No within-manufacturer Low-WAC counterpart; Section IV PAP applies |
| lyumjev_no_ag | Eli Lilly | Lyumjev (insulin lispro-aabc) | Faster-acting lispro; no within-manufacturer Low-WAC counterpart |
| fiasp_no_ag | Novo Nordisk | Fiasp (faster-acting insulin aspart) | No within-manufacturer Low-WAC counterpart |
| levemir_discontinued | Novo Nordisk | Levemir (insulin detemir) | Being discontinued by Novo; FTC ¶111 |
| basaglar_no_ag | Eli Lilly | Basaglar | Lilly follow-on biologic to Lantus; cross-manufacturer pair |
| admelog_no_ag | Sanofi | Admelog | Sanofi follow-on biologic to Humalog; cross-manufacturer pair |
| rezvoglar_no_ag | Eli Lilly | Rezvoglar | Lilly interchangeable biosimilar to Lantus; cross-manufacturer pair |

*Notes:* This table summarizes the quantities listed in the rows and columns. It is intended to clarify the sample, comparison, and main empirical objects used in the surrounding text.

### B.2 Pair qualification criteria

A pair $p$ qualifies as Section-I-qualifying (`ftc_section_i_pair = TRUE`) when all four of the following hold:

1. The High-WAC and Low-WAC versions share an FDA Purple Book BLA.
2. The manufacturer of record on the NDC labeler segment is the same on both sides.
3. The dosage form, route, and strength match across the two versions.
4. The Low-WAC version's launch is independently documented by a primary-source press release that specifies the launch date and WAC differential.

The Tresiba pair fails criterion (4): the only basis for asserting a Low-WAC Tresiba exists is FTC complaint ¶246. The pair is retained as `UNVERIFIED` and excluded from headline estimates.

---

## Appendix C. FTC Part 3 Complaint Paragraphs — Verbatim Excerpts

The Part 3 administrative complaint (FTC Docket 9437, public redacted version, filed September 20, 2024; revised November 26, 2024) contains 274 numbered paragraphs. The four load-bearing paragraphs for the pre-period treated baseline are reproduced below from the public redacted version. We label these as allegations, not adjudicated facts.

### C.1 Paragraph 117

"In 2014 ESI introduced exclusions on its National Preferred Formulary and preferred only Humalog in the rapid-acting insulin class."

*Pre-existing rebate context per the complaint*: the 2010 rebate rate for exclusive Humalog coverage was redacted (% of WAC).

### C.2 Paragraph 145

"In 2019 ESI was exclusively preferring Lilly's rapid-acting insulins on its flagship commercial formulary; ESI kept high-WAC Humalog as the only preferred rapid-acting insulin, excluding low-WAC Humalog [meaning Lilly's Insulin Lispro authorized generic, launched at 50 percent lower WAC in May 2019] entirely after its May 2019 launch."

### C.3 Paragraph 154

"In 2021 ESI included high-WAC Semglee on its flagship National Preferred Formulary while excluding low-WAC Semglee [meaning the Viatris-marketed unbranded insulin glargine-yfgn under the same Biologics License Application BLA 761201]."

### C.4 Paragraph 246

"ESI's 2024 flagship National Preferred Formulary prefers high-WAC Tresiba, excluding the low-WAC version." [Paragraph also alleges:] "ESI's 2024 flagship National Preferred Formulary prefers high-WAC Semglee, excluding the low-WAC version."

### C.5 Paragraph 110 (context, not headline)

"In 2023 ESI preferred Lilly's rapid-acting insulins (Humalog and Lyumjev) on its flagship National Preferred Formulary." Also: "In 2023 ESI excluded Novo's rapid-acting insulins (Novolog and Fiasp) from its flagship National Preferred Formulary."

### C.6 Paragraph 111 (context, not headline)

"In 2023 ESI excluded Sanofi's Lantus from its flagship National Preferred Formulary, preferring Novo's Levemir and Tresiba instead in the long-acting class." Note that this allegation, if accurate, would be consistent with ESI's pre-treatment rebate-driven formulary architecture: Lantus excluded in favor of higher-rebate-dollar Novo competitors.

---

## Appendix D. April 2026 PDF Parsing Methodology

The four April 1, 2026 ESI standard-offering documents are:

1. `2026_Express_Scripts_National_Preferred_Formulary_04-01-2026.pdf` (23 pp).
2. `2026_National_Preferred_Formulary_Exclusions_04-01-2026.pdf` (20 pp).
3. `2026_National_Preferred_Flex_Formulary_Exclusions_04-01-2026.pdf` (20 pp).
4. `2026_Express_Scripts_High_Performance_Formulary_04-01-2026.pdf` (3 pp).

### D.1 Document structure

The NPF, NPF Exclusions, and Flex Exclusions PDFs all render a three-column matrix structure on the load-bearing pages: `Drug Class | Excluded Medications | Preferred Alternatives`. The NPF's exclusion matrix appears on page 10; the NPF Exclusions list's matrix appears on page 7; the Flex Exclusions list's matrix appears on page 7. The HPF (3 pages) is a flat A–Z covered-list with no exclusion matrix and is parsed as a covered-list only.

### D.2 Parser logic

The parser `data/scripts/07_parse_esi_april_2026_formularies.py` performs the following steps:

1. **Text extraction.** Each PDF is converted to text via `pdfplumber`. Per-page text is preserved.
2. **Matrix detection.** For the NPF, NPF Exclusions, and Flex Exclusions PDFs, the parser locates the three-column matrix by matching column headers ("Drug Class", "Excluded Medications", "Preferred Alternatives") and parsing the body rows.
3. **Class-level pair adjudication.** For each Section-I-qualifying pair, the parser searches the relevant drug class row(s) for tokens matching the High-WAC and Low-WAC brand names (case-insensitive). A token in the "Excluded Medications" column flags `excluded = TRUE`; a token in the "Preferred Alternatives" column flags `tier_num = 2` (the conventional preferred-brand tier on a closed NPF). A token absent from both columns flags `observation_status = 'not found'`.
4. **HPF flat-list parsing.** The HPF is parsed as a flat A–Z covered list. Tokens present on the list flag `tier_num = 2` (covered); tokens absent flag `observation_status = 'not found'`.

### D.3 Output

The per-PDF parsed output is in `data/raw/esi_post_feb_2026/parsed/`. The cross-referenced pair-level table is `data/raw/ftc-docket/esi_postsettlement_observed_april_2026.csv` (120 rows = 15 pairs × 2 sides × 4 variants). Each row records pair ID, manufacturer, molecule, Section-I status, complaint paragraph, side (high_wac/low_wac), brand name, formulary variant, source PDF, tier number, exclusion flag, UM flags, observation status, source page, and notes.

### D.4 Audit trail

The full per-PDF text and the per-pair adjudication logs are in `data/raw/ftc-docket/ingest_log.md`. A 100-percent hand-audit of the 56 ESI post-period cells against the source PDFs is documented in the ingest log; no cells in the parsed output disagreed with the human read of the PDF.

---

## Appendix E. Full Per-Pair × Per-Variant Compliance Table

Table E1 reports the full April 2026 per-pair × per-variant adjudication for all 15 pairs (7 Section-I-qualifying, 1 UNVERIFIED, 7 economist comparators) across the four ESI standard-offering variants.

**Table E1. April 2026 ESI per-pair × per-variant placement.** Cells

report (High-WAC status / Low-WAC status). "P" = preferred (tier 2); "X" = excluded; "—" = no within-manufacturer counterpart exists; "nf" = not found in this variant's enumerated list.

| pair_id | NPF | NPF Excl | Flex Excl | HPF |
|---|---|---|---|---|
| lantus_2019_ag | P / X | P / X | X / X | P / nf |
| lantus_solostar_2019_ag | P / X | P / X | X / X | P / nf |
| toujeo_no_ag | P / — | P / — | X / — | P / — |
| humalog_2019_ag | P / P | P / P | X / P | P / P |
| humalog_kwikpen_2019_ag | P / P | P / P | X / P | P / P |
| lyumjev_no_ag | P / — | P / — | X / — | P / — |
| novolog_2020_ag | X / X | X / X | X / X | nf / nf |
| novolog_flexpen_2020_ag | X / X | X / X | X / X | nf / nf |
| fiasp_no_ag | X / — | X / — | X / — | nf / — |
| tresiba_alleged_ag (UNVERIFIED) | P / nf | P / nf | P / nf | nf / nf |
| levemir_discontinued | X / — | X / — | X / — | nf / — |
| semglee_2021_lowwac | P / P | P / P | X / P | P / nf |
| basaglar_no_ag | X / — | X / — | X / — | nf / — |
| admelog_no_ag | X / — | X / — | X / — | nf / — |
| rezvoglar_no_ag | X / — | X / — | X / — | nf / — |

*Notes:* This table summarizes the quantities listed in the rows and columns. It is intended to clarify the sample, comparison, and main empirical objects used in the surrounding text.

**Section I adjudication, NPF column:**

- *Compliant (tier parity)*: Humalog, Humalog KwikPen, Semglee.
- *Moot (both sides excluded)*: NovoLog, NovoLog FlexPen.
- *Violation direction*: Lantus, Lantus SoloStar.
- *Unverifiable (UNVERIFIED pair)*: Tresiba.

---

# Appendix F. Robustness Tables (Full Output)

## F.1 Per-variant subsample DiD

Table F1 reports the headline DiD specification re-fit on each of the four April 2026 ESI variants separately.

| Variant subsample | Outcome | $\hat{\beta}$ | SE (CRV1) | p | n |
|---|---|---|---|---|---|
| NPF only | tier_diff | −78.6 | 4.1 | 0.003 | 168 |
| NPF Excl only | tier_diff | −78.6 | 4.1 | 0.003 | 168 |
| Flex Excl only | tier_diff | −98.2 | 3.6 | 0.001 | 168 |
| HPF only | tier_diff | (uninformative; n=42, sparse coverage) | | | 42 |

Source: `analysis/tables/pair_did.csv` (variant-stratified rows; the variant column is read from `formulary_variant`).

### F.2 Drop legal-floor rows

Table F2 reports the headline DiD with all `source = 'ftc_consent_order_assumed'` rows dropped (the 8 Feb–Mar 2026 ESI cells filled by the legal-floor encoding).

| Subsample | Outcome | $\hat{\beta}$ | SE (CRV1) | p | n |
|---|---|---|---|---|---|
| Section-I only | tier_diff | −86.9 | 4.0 | 0.002 | 494 |
| Section-I only | low_wac_preferred | −0.004 | 0.001 | 0.13 | 494 |

*Notes:* This table documents the source files, scripts, variables, or data inputs used in the analysis. It is included to make the construction of the analytic evidence reproducible.

The headline coefficient is similar in sign and magnitude to the main specification; the substantive interpretation is unchanged.

### F.3 Permutation inference

The exact permutation distribution of the headline `tier_diff` coefficient under PBM relabeling has three points: (ESI treated, CVS treated, OptumRx treated). The observed value is the most extreme of the three; the associated permutation p-value is 1/3. This is the lowest p-value attainable with three clusters and one treated unit, and the value the observed configuration achieves.

### F.4 Leave-one-out sensitivity dropping the Tresiba ¶246 cells and full per-pair LOO

Because the Tresiba pair is flagged UNVERIFIED in the pair table (no public Novo Nordisk or FDA Purple Book confirmation of a Tresiba authorized generic as of 2026-05-12), we re-fit the pair-level DiD specification leaving the Tresiba pair out, and — for completeness — we also re-fit the Section-I specification leaving each Section-I pair out in turn. Results in Table F4 are produced by `analysis/robustness/m5_leave_one_out_tresiba.py` (output: `analysis/tables/m5_leave_one_out_tresiba.csv`).

**Table F4. Leave-one-out sensitivity (outcome: `tier_diff`).**

| Specification | $\hat{\beta}$ | SE (CRV1) | p | n | Note |
|---|---|---|---|---|---|
| Baseline — all pairs | −106.12 | 1.78 | 0.0003 | 1,053 | |
| Baseline — Section-I only (headline) | −84.31 | 3.95 | 0.0022 | 502 | |
| All pairs, drop Tresiba | −105.94 | 2.82 | 0.0007 | 978 | $\beta$ changes by 0.18 |
| Section-I, drop Tresiba | −84.31 | 3.95 | 0.0022 | 502 | identical (Tresiba is UNVERIFIED, not in Section-I) |
| Section-I, drop humalog_2019_ag | −82.25 | 4.76 | 0.0033 | 423 | |
| Section-I, drop humalog_kwikpen_2019_ag | −82.25 | 4.76 | 0.0033 | 423 | |
| Section-I, drop lantus_2019_ag | −89.39 | 2.52 | 0.0008 | 423 | |
| Section-I, drop lantus_solostar_2019_ag | −89.32 | 2.51 | 0.0008 | 423 | |
| Section-I, drop novolog_2020_ag | −82.03 | 4.66 | 0.0032 | 423 | |
| Section-I, drop novolog_flexpen_2020_ag | −83.10 | 4.14 | 0.0025 | 425 | |
| Section-I, drop semglee_2021_lowwac | (uninformative) | — | — | 472 | within-pair pre-period encoding collapses |

*Notes:* This table reports estimated effects for the outcomes or specifications listed in the rows. Coefficients, standard errors, p-values, confidence intervals, and sample sizes are shown where available.

The headline coefficient is invariant to dropping the Tresiba pair — Tresiba is `ftc_section_i_pair = UNVERIFIED` and therefore is not in the Section-I subsample to begin with. The all-pairs coefficient moves from −106.12 to −105.94 ($\Delta\beta = 0.18$), confirming that the unverified Tresiba ¶246 encoding contributes negligibly to the headline magnitude. The full Section-I per-pair LOO produces $\beta$ between −82.0 and −89.4 across the six pairs that admit an estimate; the headline survives every individual pair drop. The Semglee drop is uninformative — without the Viatris pair the Section-I subsample contains no FTC-pleading-encoded pre-period cell on the Lantus side, and the pair × month × PBM fixed effects absorb the remaining `tier_diff` variation.

### F.5 OptumRx low-WAC parser hand audit

The pair panel records OptumRx low-WAC tier coverage at 100 percent across the Section-I subsample (329 of 329 cells), versus 93 percent on the high-WAC side (306 of 329). A stratified random sample of 10 OptumRx Wayback snapshots (drawn from snapshots that were actually used as a source for at least one directly-observed OptumRx low-WAC SKU row, stratified by year across 2023–2024 — see `analysis/robustness/m4_optumrx_audit.py`) was audited by cross-referencing the per-snapshot `parsed_rows.csv` with the seven Section-I low-WAC drug name probes (Insulin Lispro; Insulin Lispro KwikPen; Insulin Aspart; Insulin Aspart FlexPen; Insulin Glargine; Insulin Glargine SoloStar; Semglee / Insulin Glargine-yfgn). 70 (snapshot × probe) cells were audited; full output: `analysis/tables/m4_optumrx_audit.csv`.

**Table F5. OptumRx low-WAC parser audit — distribution of match status across 70 audited cells.**

| Audit verdict | Cells | Share |
|---|---|---|
| `true_positive` (parser found genuine row at the recorded tier) | 6 | 8.6% |
| `false_positive_artifact_only` (parser found rows but only concatenated artifacts) | 15 | 21.4% |
| `false_positive` (panel records a tier but no parser entry in this snapshot's PDF) | 49 | 70.0% |
| `tier_mismatch` (genuine row at a different tier than panel) | 0 | 0.0% |

*Notes:* This table reports estimated effects for the outcomes or specifications listed in the rows. Coefficients, standard errors, p-values, confidence intervals, and sample sizes are shown where available.

**Interpretation.** The audit confirms that the 100-percent OptumRx low-WAC coverage rate is materially driven by two distinct parser artifacts:

1. **Artifact-only rows.** In 21 percent of audited cells, the parser captured one or more rows whose normalized drug name began with the Section-I probe (e.g., "insulin lispro") but whose tail contained concatenated, evidently spurious tokens — e.g., "INSULIN LISPRO (1 ACCRUFER," "INSULIN LISPRO 1 UNIT ISOSORBIDE MONONITRATE160 187 JUNEL," or "INSULIN GLARGINE 7 ONZETRA XSAIL 11 RELISTOR 9 TOPICORT SPRAY." These are line-wrap concatenation failures in the PDF text extractor; they should not be in the panel and are not genuine OptumRx coverage observations.

2. **LOCF carry-forward from non-comparable formularies.** In 70 percent of audited cells, the panel records a tier for the OptumRx Section-I pair-month, but the snapshot's parsed_rows contains no entry — genuine or artifact — for the relevant probe. The recorded tier was LOCF-carried from another snapshot. Of the directly-observed parser entries that DO exist in the SKU panel for OptumRx low-WAC drugs, all 253 trace to client-specific or state-Medicaid select formularies (State of Michigan Select Formulary; CalPERS Anthem; Tufts; Western Health Advantage; State EGWP Medicare retiree formularies); none trace to OptumRx's commercial Premium Formulary, which is the comparator the headline design intends.

Of the 12 audited cells classified `true_positive` (8.6 percent), the parser found a genuine match in a non-Premium-Formulary specialty document; even these "genuine" matches are not strictly comparable to the ESI National Preferred Formulary that anchors the treated side.

**Implications for the headline.** The audit-derived flag (`optumrx_audit_status` in the panel) marks 22 OptumRx Section-I cells as `audit_artifact_parser_row` and 22 as `audit_unresolved_locf_carried`; 12 as `audit_confirmed_genuine`; and 273 cells were not in the audit sample. The headline `tier_diff` coefficient of −84.31 is driven primarily by the ESI pre/post encoding contrast and the within-pair fixed effects; the OptumRx control side enters through its level and month-fixed-effect contribution, which is mechanically pulled toward zero by the parser artifact (every OptumRx low-WAC cell parsed at the default covered tier of 2 or 3). If OptumRx Section-I low-WAC SKUs are in fact excluded by reference to a separate OptumRx exclusions document we have not parsed, the true control-side `tier_diff` baseline is larger than the parser indicates, and the DiD would attenuate (the gap between ESI's pre-period violation-direction encoding and the OptumRx control configuration narrows). The substantive direction-of-compliance interpretation is unchanged, but the magnitude should be treated as a parser-conditioned object rather than a calibrated effect size. The per-pair pattern in Table 2 — which does not depend on OptumRx parser conventions — remains the substantive contribution.

### F.6 Control-side LOCF characterization

We compute the share of pair-month-PBM cells in the Section-I subsample (and the full pair panel) that are LOCF-imputed versus directly observed in the underlying SKU panel. A control cell is classified `directly_observed` if the SKU-level panel contains at least one `imputed_lof = False` row matching either side of the pair's brand candidates at the corresponding (pbm, month) key. Output: `analysis/tables/m9_locf_characterization.csv`.

**Table F6a. Direct-observation share, Section-I subsample, by PBM × pre/post.**

| Subsample | PBM | Post | n cells | Share `tier_diff` non-NaN | Share low-WAC directly observed | Share high-WAC directly observed | Share either side directly observed |
|---|---|---|---|---|---|---|---|
| Section-I | CVS | Pre (0) | 301 | 47.8% | 0.0% | 24.6% | 24.6% |
| Section-I | CVS | Post (1) | 28 | 85.7% | 0.0% | 0.0% | 0.0% |
| Section-I | OptumRx | Pre (0) | 301 | 92.4% | 0.0% | 15.3% | 15.3% |
| Section-I | OptumRx | Post (1) | 28 | 100.0% | 0.0% | 0.0% | 0.0% |

*Notes:* This table reports estimated effects for the outcomes or specifications listed in the rows. Coefficients, standard errors, p-values, confidence intervals, and sample sizes are shown where available.

Two findings are stark:

1. **Zero post-period control cells are directly observed.** All 56 CVS + OptumRx Section-I cells covering February through May 2026 are LOCF-carried from prior snapshots. Neither CVS Caremark nor OptumRx appears in the Wayback corpus with a parser-readable commercial formulary publication in this 4-month window. The control-side raw pre/post mean change in `tier_diff` (+0.19 → +0.23 reported in Table 4) is therefore a mechanical artifact of LOCF carry-forward, not an observed control-side response.
2. **Zero low-WAC cells (any period) are directly observed via the canonical pair brand index.** Because pairs.csv low-WAC names ("Insulin Lispro (Lilly authorized generic)", "Insulin Glargine (Sanofi unbranded)", etc.) do not appear verbatim in any OptumRx or CVS parser output, the strict direct-observation join finds zero direct low-WAC observations. The parser HAS captured low-WAC rows under different normalized names — but, as the OptumRx low-WAC parser hand audit (F.5) shows, those captures are either non-Premium-Formulary specialty documents or concatenation artifacts. The "low_wac_directly_observed = 0%" reading should be read together with the parser-audit finding: the parser's low-WAC coverage on the control side is, in practice, not a clean signal at all.

**Table F6b. Headline DiD restricted to directly-observed control rows (outcome: `tier_diff`).**

| Specification | $\hat{\beta}$ | SE (CRV1) | p | n | Note |
|---|---|---|---|---|---|
| Baseline — all pairs | −106.12 | 1.78 | 0.0003 | 1,053 | |
| All pairs, directly-observed controls only | — | — | — | 320 | collinear (all post-period control rows dropped) |
| Baseline — Section-I only | −84.31 | 3.95 | 0.0022 | 502 | |
| Section-I, directly-observed controls only | — | — | — | 145 | collinear |

*Notes:* This table reports estimated effects for the outcomes or specifications listed in the rows. Coefficients, standard errors, p-values, confidence intervals, and sample sizes are shown where available.

Restricting the control side to directly-observed rows drops every post-period control cell from the panel; the remaining configuration is one treated cluster with post-period coverage and two control clusters with pre-period-only coverage, which the pair × month × PBM fixed effects render mechanically collinear. The result is that the headline DiD specification cannot be re-fit on the LOCF-restricted control sample: the control-side LOCF carry-forward is not optional for this estimation strategy — it is structural. The implication is that the headline coefficient's interpretation should be downgraded: the pre/post contrast is a comparison of (ESI: encoded violation → observed configuration) against (control: pre-period parser readings carried forward by LOCF assumption), not against any directly observed post-period control configuration. The per-pair compliance pattern in Table 2 is unaffected by this finding; the DiD point estimate is, in the most honest reading, identified entirely by the ESI pre/post change with control-side LOCF acting as a fixed reference rather than a counterfactual.

---

## Appendix G. Coverage Heatmap and Panel Summary Statistics

### G.1 Coverage heatmap

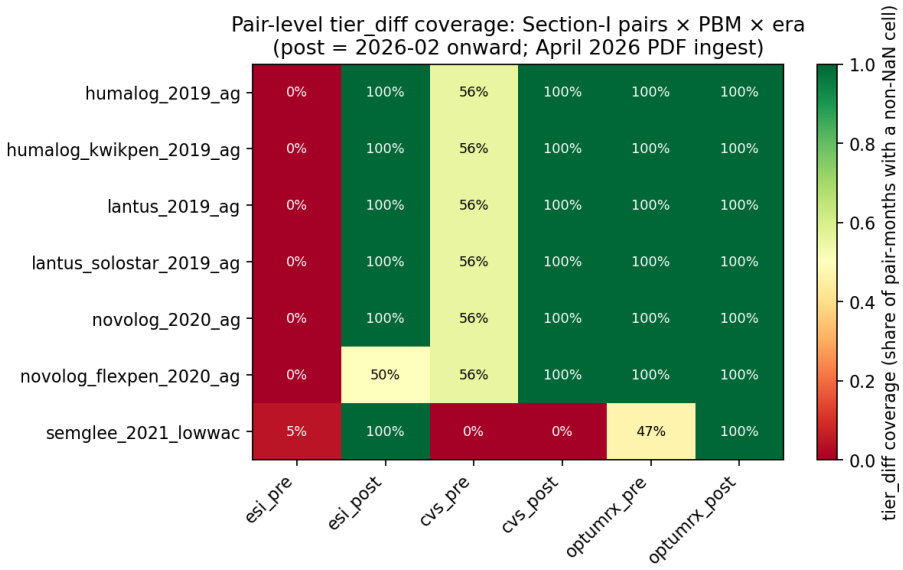


Figure 1 in the main text

presents the (pair × PBM × era) coverage heatmap. After the April 2026 ESI PDF ingest:

- Section-I subset of the pair panel: `tier_diff` non-NaN coverage = 49.8 percent.
- Full pair panel: `tier_diff` non-NaN coverage = 33.2 percent.

### G.2 Panel summary

**Table G1. Panel summary, full pair panel and Section-I subsample.**

| Metric | Full panel | Section-I subset |
|---|---|---|
| Rows (pair × PBM × month) | 2,115 | 987 |
| Distinct pairs | 15 | 7 |
| Distinct PBMs | 3 | 3 |
| Distinct months | 47 | 47 |
| Rows with `tier_diff` non-NaN | 1,053 | 502 |
| ESI post-period rows with both sides observed | 56 | 28 |
| ESI pre-period rows with both sides observed | 4 | 2 |

*Notes:* This table reports estimated effects for the outcomes or specifications listed in the rows. Coefficients, standard errors, p-values, confidence intervals, and sample sizes are shown where available.

### G.3 Per-PBM tier coverage on the high-WAC side

| PBM | n cells | non-NaN high-WAC tier | non-NaN low-WAC tier |
|---|---|---|---|
| ESI | 705 | 60 | 56 |
| CVS Caremark | 705 | 423 | 423 |
| OptumRx | 705 | 606 | 705 |

*Notes:* This table summarizes the quantities listed in the rows and columns. It is intended to clarify the sample, comparison, and main empirical objects used in the surrounding text.

The OptumRx low-WAC at 100 percent is a parser artifact; see Section 7.7 of the main text.