

Did demand-pull deliver? Medicare Part D and pharmaceutical approval rates*

Meg Blume-Kohout[†]

June 8, 2026

Abstract

This paper re-examines the effect of Medicare Part D on pharmaceutical R&D with three refinements to the Blume-Kohout and Sood (2013) design. First, treatment intensity is recoded to zero for two Part-A/B-administered classes, supplying a zero-treatment reference. Second, Phase 1 entry is re-estimated as an unconditional event study and with both continuous and binned treatment, following recent best practices. Finally, an indication-aware FDA-approval mapping links Phase 1 trials to subsequent approvals at the drug-indication grain, testing the on-the-shelf hypothesis. The reconstructed panel entry estimates replicate the 2013 result within 0.08 in every post-period coefficient under the original specification. Using a two-way fixed-effects design (with year fixed effects in place of class-specific linear trends) that the `mshare2` zero-treatment reference makes feasible, I find similar results to those in 2013 for Phase 1 entries. FDA approval rates per Phase 1 trial decline with Medicare share post-Part-D, imprecisely estimated but directionally consistent with the on-the-shelf hypothesis that Part-D-induced entries were marginal investments.

JEL Codes: I18, L65, O31, C23, C25

*Working draft prepared for the Annual Health Econometrics Workshop, 2026. Replication code and reconstructed analytic table available from the author. The author acknowledges Anthropic's Claude Code for data-engineering assistance on the indication-aware FDA-approval mapping and crosswalk reconstruction; all analytical choices, interpretations, and conclusions are the author's own.

[†]Gettysburg College, Department of Economics. Email: mblumeko@gettysburg.edu.

1 Introduction

Medicare Part D was the largest single expansion of US prescription drug coverage since Medicare’s creation; by 2010 it covered roughly 28 million beneficiaries (Blume-Kohout and Sood 2013). The pharmaceutical industry’s contemporaneous advocacy was explicit that demand-side support of this scale was expected to lift pharmaceutical research and development. In our earlier work, we found a strong R&D response, with higher Medicare market share classes exhibiting greater positive departure from their historical growth rates. Whether the early increase in Phase 1 trials we observed translated into an increase in new drug approvals remains an open question.

This paper addresses that question in two ways. First, it tests whether the Phase 1 entries that we attributed to Medicare Part D translated into approved drugs, allowing a 15-year follow-up window for the original paper’s 1998–2010 cohorts to observe projects through FDA approval or discontinued development. Second, it re-estimates the Part D effect in light of the modern continuous-treatment difference-in-differences (DiD) literature, which has clarified identifying assumptions and weighting problems that we did not explicitly address in 2013.

Several methodological challenges constrained our 2013 design. First, our treatment variable, Medicare market share, was continuous, calculated as the survey-weighted pre-Part-D share of prescriptions filled by Medicare beneficiaries in 2004–2005 (prior to Part D’s implementation) using Medical Expenditure Panel Survey (MEPS) data. While this construction naturally constrained possible values to the unit interval, the range of *observed* values included neither zero nor one. This was a feature, not a bug: at the time it seemed plausible all therapeutic classes in MEPS would be affected by Part D. Due to the lack of any clean control, we recognized that including year fixed effects as in a typical TWFE DiD could have absorbed part of the Part D treatment effect. So, we instead parameterized therapeutic class-specific time trends, identifying the Part D effect by deviations from trend across the classes. This secondarily accounted for visibly non-parallel pre-trends for the lowest- and highest-Medicare-share classes, *Hormones: Contraceptives & Fertility* and *Alzheimer’s Disease*, evident in our 2013 paper’s Figure 2.

We interacted the continuous Medicare-share measure with a post-2003 step in a Poisson specification. Recent work has shown that the continuous-treatment TWFE estimator carries a parallel-trends assumption at every dose level (Callaway et al. 2024a), that weights on the dose-response slopes can flip sign when the response is heterogeneous in the dose, and there is no clean control reference when the treatment intensity has no zero (Chaisemartin et al. 2024). The standard remedy is to discretize the continuous dose, which weakens parallel trends to a within-bin form and supplies a control reference. (Callaway et al. 2024a, 2024b). I implement

these remedies first by recoding Medicare market share for two Part-A/B-administered classes (Cancer: Antimetabolite and Cardiovascular: Inotropic), setting their values to zero, then by binning Medicare market share into five groups. Outcomes are estimated by Poisson pseudo-maximum-likelihood, which accommodates the count distribution with mass at zero.

Our 2013 paper documented a rapid Phase 1 response in 2004–2005, before the January 2006 implementation, and offered this "on-the-shelf" hypothesis: firms might be pushing marginal products into clinical testing in anticipation of the market expansion. The implication was that drugs added to the pipeline in response to Part D could be riskier investments with lower probability of success. Two signals in the 2013 estimates pointed this direction: Phase 1 entry in the dual-eligible classes more than doubled while Phase 3 follow-through was negligible, and the Phase 2 failure literature we cited attributed late-stage attrition to lack of efficacy. Neither of these possibilities could be tested at the time, because many projects were still in active development by our 2010 endpoint. The fifteen-year follow-up window now makes the test possible. The null hypothesis under demand-pull with quality-constant entry is that the approval rate per Phase 1 trial is unchanged; under the on-the-shelf conjecture, the approval rate per Phase 1 trial falls.

A second contribution is to revisit our 2013 Medicare-share gradient argument under the specification standards that the recent DiD literature recommends. Our 2013 paper noted an apparent Medicare-share gradient in the Phase 1 response, similar to that found by Duggan and Scott Morton (2010) for revenues. Krieger et al. (2022) subsequently documented a broadly uniform cash-flow lift in approvals at the firm-class level. On the class-year panel, the post-MMA Phase 1 entry response persists under both our 2013 specification (with class-specific linear trends) and a conventional two-way fixed-effects design (year fixed effects, but no trends). The TWFE identification is enabled by the `mshare2` refinement, which provides two true-zero-treatment reference classes (the Part-A/B-administered Cancer: Antimetabolite and Cardiovascular: Inotropic); the 2013 paper, lacking a zero-treatment reference, identified the dose-response gradient via deviations from class-specific trends. Section 3 reports bin magnitudes of comparable size across the two designs. The Krieger uniform-lift framing applies most cleanly to firm-level activity (their primary sample), which a class-year panel cannot resolve directly.

On the approval-rate question, the evidence is directionally consistent with the on-the-shelf hypothesis: approval rates per Phase 1 trial decline with Medicare share across all three sponsor samples in the immediate post-MMA periods, with the largest point estimates in the small-sponsor sample, though no individual coefficient rejects zero at conventional thresholds.

2 Data and Empirical Strategy

2.1 Sources

The analytic dataset derives from four data sources. Project-level R&D pipeline data come from Pharmaprojects (Citeline / Informa), the same extract we used for Blume-Kohout and Sood (2013). For this paper, I re-derived the therapeutic class $\times$ year panel from project-level records using our original crosswalk to a 49-class taxonomy, then split one composite class from the original taxonomy, *Hormones: Menopause and Osteoporosis*, into its two component classes. FDA approval data come from the Orange Book and the CDER Compilation of New Drug Application (NDA), Biologics License Application (BLA), and Approval Documents refreshed through 2025 to provide fifteen-year follow-up for the 2010 cohort of Phase I entrants. Class-level treatment intensity (`mshare`) is held fixed per our original construction from the Medical Expenditure Panel Survey (MEPS), with the exception of the two Part A/B class recodings mentioned above. Disease-specific lagged NIH funding controls (ℓ NIH at lags 4–15 years) are likewise imported from the original panel, constructed by Blume-Kohout (2012) from NIH RePORTER disease-specific funding aggregates mapped onto the Pharmaprojects class structure.

I rebuilt the class-year panel from the original raw HTML project-level records rather than carrying forward the legacy aggregated panel. Section 2.5 reports the cell-by-cell comparison against the 2013 published counts. For each approved NDA 1990–2010, I extracted the drug’s approved indication text from the openFDA drug-label API. The indication text was then scored against per-class keyword vocabularies seeded from mapped entries in the National Library of Medicine’s Medical Subject Headings (MeSH) thesaurus and from Pharmaprojects indication-event history, producing an automated mapping of approved drugs to therapeutic classes. I then manually inspected the assignments for every new molecular entity (NME) approved by the FDA in 1990–2010 and adjusted by hand where the automated mapping was wrong or ambiguous. The cohort universe contains 419 NMEs distributed over 48 of the 50 panel classes, and is the basis for the indication-aware FDA-approval flag used as the second outcome below.

2.2 Cohort universe and partition

The unit of observation for data reported in the descriptive table (Table 1) is (drug $\times$ therapeutic class). The analytic table contains 7,486 observations across 5,588 distinct Pharmaprojects projects and 50 therapeutic classes, of which 4,076 enter Phase 1 in the 1998–2010 window. The 50-class basis reflects the *Hormones: Menopause and Osteoporosis*

split described above. Two of the fifty classes (*Anxiolytics* and *Hormones: Menopause*) have no cohort NME approvals. For the cell-by-cell comparison against the published 2013 Table 3 in [Section 2.5](#), the two split classes are aggregated back to the 49-class basis we used in 2013. For the Poisson models that follow, counts of Phase 1 entries and approvals are aggregated by therapeutic class and year.

The 1998–2010 cohort window is partitioned into five periods aligned with the post-period structure of our 2013 paper: Pre-Part-D (1998–2002); anticipatory effects the year the legislation passed (2003); after passage but before implementation (2004–05); immediate post-implementation (2006–07); and longer-run effects (2008–10). The 2010 endpoint for Phase 1 entry allows FDA approvals up to 15 years later, with the 2025 endpoint of the approval data pull. The 15-year follow-up window is adequate to observe outcomes for the large majority of Phase 1 cohort members. Cho et al. (2025) report mean phase durations by therapeutic area across 90,000+ trials from the Citeline database (Table 2b); summing Phase 1, Phase 2, and Phase 3 means and adding a 12–18 month FDA review period, total development time from Phase 1 entry to approval averages 7–8 years for most therapeutic classes. Oncology is the main exception, with a mean total exceeding 12 years before review; however these classes are largely excluded from the identification or lower Medicare share.

2.3 Treatment intensity: `mshare`, `mshare2`, and the five-bin discretization

The treatment-intensity measure carried forward from Blume-Kohout and Sood (2013), denoted `mshare`, is the class-level Medicare-eligible share of prescription volume in MEPS 2004–05. It ranges in the cohort universe from 0.007 (*Hormones: Contraceptives & Fertility*) to 0.865 (*Alzheimer’s Disease*) and is both class-invariant and time-invariant.

The refinement `mshare2` addresses an institutional feature of two classes whose Medicare patients are billed under Parts A or B rather than Part D. Cancer: Antimetabolite and Cardiovascular: Inotropic consist almost entirely of hospital-administered or physician-administered drugs covered under inpatient (Part A) or outpatient-medical (Part B) reimbursement; for these classes, the demand-pull from the Part D outpatient retail-pharmacy benefit is essentially zero, because Medicare beneficiaries’ utilization was largely already covered before Part D took effect. In 2013 we introduced a separate indicator covering Cancer and Immunosuppressant drugs for the Part B sensitivity table; this paper instead zeroes the Medicare share value for the two classes named above.¹

1. In 2013 we included Immunosuppressants in the same Part-A/B set. On the refreshed panel, Phase 1 entry in this class rises 75 percent between the pre-Part-D period and 2008–2010, dominated by autoimmune indications (rheumatoid arthritis, psoriasis, lupus) that are Part-D- rather than Part-A/B-treated. The class

I additionally discretize `mshare2` into five bins. Bin 1 contains the two Part-A/B-administered classes (`mshare2` = 0); bins 2–5 partition the remaining distribution with cutpoints 0.2, 0.4, and 0.6. Discretizing the treatment serves two purposes. First, it provides a zero-treatment reference group for the four positive-dose bins, which the original continuous `mshare` variable lacked. Second, it avoids the restrictions imposed by continuous TWFE, which Callaway et al. (2024a) show produces a weighted mixture of treatment effects that is difficult to interpret. Instead, identification requires only that each bin trends parallel to the untreated group in the absence of treatment — a separate condition for each bin, but one that places no restrictions on how the response varies across bins. The five-bin specification is the basis for the dose-response figures and the heterogeneous-response check; the continuous `mshare2` specification is retained as the direct comparison against our 2013 estimates.

2.4 Outcomes and the exposure-offset specification

I estimate two outcomes. The first is Phase 1 entry: the class-year count of distinct drug projects entering Phase 1 clinical trials in a given (class, year) cell. This is the outcome on which our 2013 paper identified the Medicare-share gradient, and it is the outcome on which the panel reconstruction (§2.5) is calibrated. The second is FDA approval within 15 years of Phase 1 entry, aggregated to the class $\times$ period. To count as approval, the indication-aware assignment here requires that the project’s therapeutic class at Phase 1 entry match the therapeutic class mapping of its approved indication.

Both outcomes are estimated by Poisson pseudo-maximum-likelihood (PPML). For Phase 1 entry, the estimating equation follows our 2013 specification:

$$\mathbb{E}[\text{entries}_{ct} \mid X_{ct}] = \exp\left(\alpha_c + \delta_t + \sum_{k=1}^4 \beta_k(\text{mshare}_c \times \text{Post}_k) + \gamma \log pv8_{ct} + \mathbf{X}'_{ct}\boldsymbol{\theta}\right) \quad (1)$$

where c indexes therapeutic class, t indexes calendar year, the α_c are class fixed effects, the Post_k indicators are the four post-period dummies of the cohort partition, and X_{ct} collects the lagged-NIH controls (ℓNIH at lags 4–15). The two-way fixed-effects specification with class-specific linear trends matches the 2013 estimating equation directly; I report it in Column 2 of Table 2 for the validation against Column 1. Column 4 specification replaces the class-specific linear trends with year fixed effects, holding the covariates and the rest of the equation fixed in a conventional TWFE design that retains $\log pv8$ and the LL_4 – LL_{15} NIH-funding lags as controls.

is therefore retained at its unrestricted `mshare` value here.

For the approval-rate outcome I estimate

$$\mathbb{E}[\text{approvals}_{ct} \mid X_{ct}] = \overline{p1}_c \cdot \exp\left(\sum_k \beta_k \text{mshare2}_c \times \text{Post}_{k,t} + \mathbf{X}'_{ct} \boldsymbol{\theta}\right) \quad (2)$$

where approvals_{ct} counts drugs that entered Phase 1 in cohort year t in therapeutic class c and received an FDA approval within 15 years; $\overline{p1}_c$ is the class-level 1998–2002 mean of Phase 1 entries, used as a PPML exposure offset (class-fixed across cohort years); $\text{Post}_{k,t} = \mathbb{I}[t \in \text{period } k]$ indexes the four post-MMA periods (2003, 2004–05, 2006–07, 2008–10); $\mathbf{X}_{ct}$ collects the class-time-varying controls ($\log pv_8$ and the LL_4 – LL_{15} NIH-funding lags); and the coefficient β_k identifies the differential approval-rate response per pre-Part-D Phase 1 trial in period k , on the log scale. The post-period structure enters through the period dummies; the unit of observation remains therapeutic class $\times$ cohort year. Class fixed effects would drop the 20 classes with zero approvals via perfect prediction; the pre-period offset $\overline{p1}_c$ controls for differences across classes in pipeline scale. Year or period fixed effects are similarly omitted: the two $\text{mshare2} = 0$ reference classes record zero 15-year approvals from any post-MMA cohort entering 2003–2010, leaving no within-reference variation across post-period years to identify time effects separately from the $\text{mshare2} \times \text{period}$ interactions of interest. One limitation of this identification deserves explicit acknowledgment. The $\text{mshare2} \times \text{Post}$ coefficients in Equation (2) compare approval rates per Phase 1 trial in higher-Medicare-share classes to the zero-mshare2 reference classes across post-MMA cohort periods. This comparison cannot separately identify a Part-D-induced quality decline from pre-existing differences in pipeline maturity or therapeutic-area complexity that are correlated with Medicare share. The estimates are therefore interpretable as reduced-form evidence on whether approval rates move in the direction the on-the-shelf hypothesis predicts, not as a clean causal test.

Standard errors are cluster-robust at the therapeutic class level (50 clusters). Confidence intervals are reported at the 90-percent level throughout. With 46–50 therapeutic-class clusters across the three samples, asymptotic cluster-robust inference is defensible without finite-sample corrections.

2.5 Panel reconstruction and pretrend evidence

The reconstructed (class $\times$ year) panel matches the 2013 Table 3 published cell counts closely. Of the 637 panel cells in the 1998–2010 window (aggregating the two split classes back to the 49-class basis), 90.3 percent match exactly and 99.7 percent differ by no more than five entries. Total reconstructed cohort entries (4,076) run 69 below the legacy panel (4,145), with the largest residuals in Antibiotics (−27), Hormones: Menopause and Osteoporosis

(−26), Respiratory: Other (−16), and GI Inflammatory Disease (+13). Two sources of discrepancy account for these gaps: re-derivation from raw Pharmaprojects project-level exports introduces parsing differences in ambiguous cases, and certain ad-hoc drug-level recodings performed at the MEPS-crosswalk stage for the 2013 paper were not documented in our legacy code so could not be replayed. However, cell-by-cell residuals are not systematically associated with `mshare` (the correlation between residual magnitude and class `mshare` is near zero), so the reconstruction gap is uncorrelated with `mshare` and should not bias the coefficients.

2.6 Sample composition

[Table 1](#) reports the 1998–2010 Pharmaprojects Phase 1 cohort split into the five-period breakdown. Rows report Phase 1 project counts, unique drug counts, mean treatment intensity, and shares for the principal project-level covariates and the approval outcome.

Table 1: Cohort composition by entry-year cohort, 1998–2010

	(1)	(2)	(3)	(4)	(5)	(6)
Phase 1 project entries	1,008	256	579	706	1,527	4,076
Unique drugs	824	212	493	602	1,271	3,275
Average Medicare share	0.405	0.403	0.414	0.414	0.410	0.410
Share approved by FDA within 15 years	0.033	0.023	0.019	0.017	0.007	0.018
Percent sponsored by large firms	0.245	0.223	0.259	0.297	0.253	0.258
Percent sponsored by small firms	0.612	0.645	0.661	0.612	0.653	0.636

Notes: Unit of observation is the (drug × class) project row from the analytic table (7,486 rows; 4,076 in the 1998–2010 cohort). Columns: (1) Pre-Part-D 1998–2002 (base period); (2) 2003 (MMA legislation passed; separated from the base period due to possible anticipation effects); (3) 2004–2005 (post-passage, pre-implementation); (4) 2006–2007 (Part D enacted January 2006); (5) 2008–2010; (6) All 1998–2010. The Medicare share is the Medicare patient share of total prescriptions for the drug’s therapeutic class, with two Part-A/B-administered classes (Cancer: Antimetabolite, Cardiovascular: Inotropic) set to zero because their patients are billed under Medicare Part A or B (institutional / physician-administered) rather than Part D (retail pharmacy). The approval indicator is indication-aware to the cohort therapeutic class. Sponsors are classified by their cumulative count of FDA-approved new molecular entities (NDA and BLA pathways combined, M&A-consolidated) at the project’s Phase 1 entry year minus one: large firms have at least 10 such approvals in portfolio at entry; small firms have none of any type; the residual mid-sized firms are not reported.

Fig. 1 partitions Phase 1 entry counts along Krieger et al. (2022)’s definitions for Large (established novel-drug-developer) and Small (no prior FDA-approved drugs) sponsors. The mid-sized residual is omitted because the sponsor-subgroup analyses contrast Large versus Small. Total Phase 1 entry across the two reported groups roughly triples from approximately 140 in 1998 to approximately 480 in 2009.

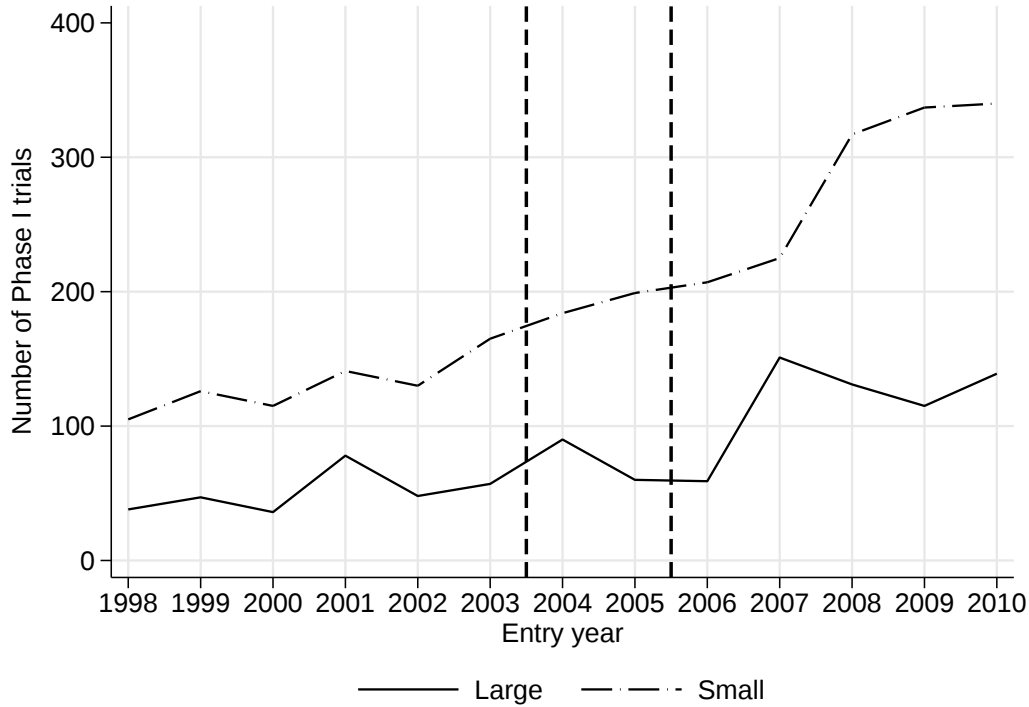


Figure 1: Phase 1 entry counts by year, by Krieger sponsor classification. *Large* = ≥ 10 FDA-approved NMEs in originator’s portfolio at Phase 1 entry minus one (NDA + BLA, M&A-consolidated, per Krieger et al. (2022)); *Small* = no FDA-approved drugs at entry. Mid-sized firms (residual) omitted; see Table 1 for the full distribution. Vertical dashed reference lines mark the post-period boundaries at end-of-2003 (MMA legislation passed) and end-of-2005 (Part D enacted January 2006).

3 Phase 1 entry response

Fig. 2 provides a year-by-year disaggregation of the post-period interaction structure in our 2013 paper, with the same `mshare`-by-year and `mshare2`-by-year interactions used in Table 2 but estimated separately for each calendar year rather than aggregated to the five post-Part-D periods. Without year fixed effects to absorb common-time shocks, this figure is for descriptive purposes, and is not an event study: each plotted coefficient reflects both class-specific differential responses and any year-specific common variation in entry rates that the model cannot otherwise attribute. The substantive observation is the visual flatness of the pre-Part-D period (1998–2002 coefficients cluster near zero, with the exception of a 1998 spike in the no-trends panel driven by a within-year shock in Antithrombotic R&D), followed by a clear post-MMA rise in the differential Medicare-share-weighted entry pattern under the no-class-trends specification (Panel B).

Fig. 3 presents the unconditional event-study for the Medicare share $\times$ cohort-year interaction. Year fixed effects absorb common-time shocks. Panel A adds class-specific linear trends; Panel B (the no-trends specification) drops them. Neither panel includes the covariates ($\log pv_8$ or the LL_4 – LL_{15} NIH-lag controls): this is the unconditional year-by-year response, with only class and year fixed effects.

In Panel B on the reconstructed 50-class panel, the coefficients exceed the 2002 reference at $p < 0.10$ in four post-MMA years: 1.29 in 2005, 0.84 in 2006, 0.93 in 2007, and 1.10 in 2009. The remaining post-MMA years (2003, 2004, 2008, 2010) are positive but imprecise. The pre-period coefficients hover near zero except for a 1998 spike of roughly +1.6 driven by a shock in Antithrombotic R&D. The pre-period coefficients (1998–2001) are not jointly significant.

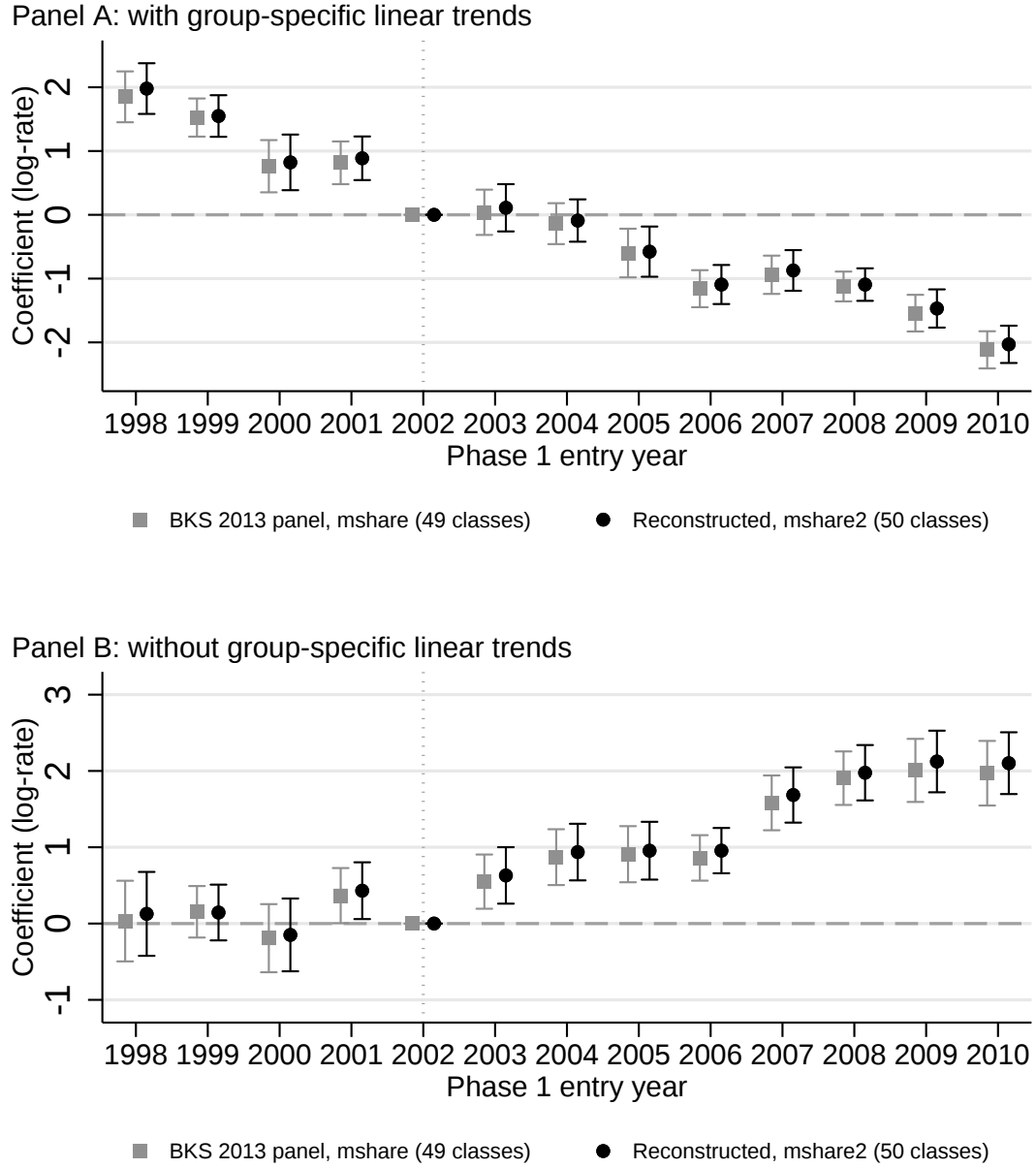


Figure 2: Descriptive changes in Phase 1 entry, 1998–2010, 90% confidence intervals. Two series per panel: our original 2013 panel (49 classes, original `mshare`) and reconstructed panel (50 classes, refined `mshare2`). Panel A includes class-specific linear trends (`i.group × c.time`); Panel B includes class fixed effects only. Neither panel includes year fixed effects. The 2002 reference year is marked at zero.

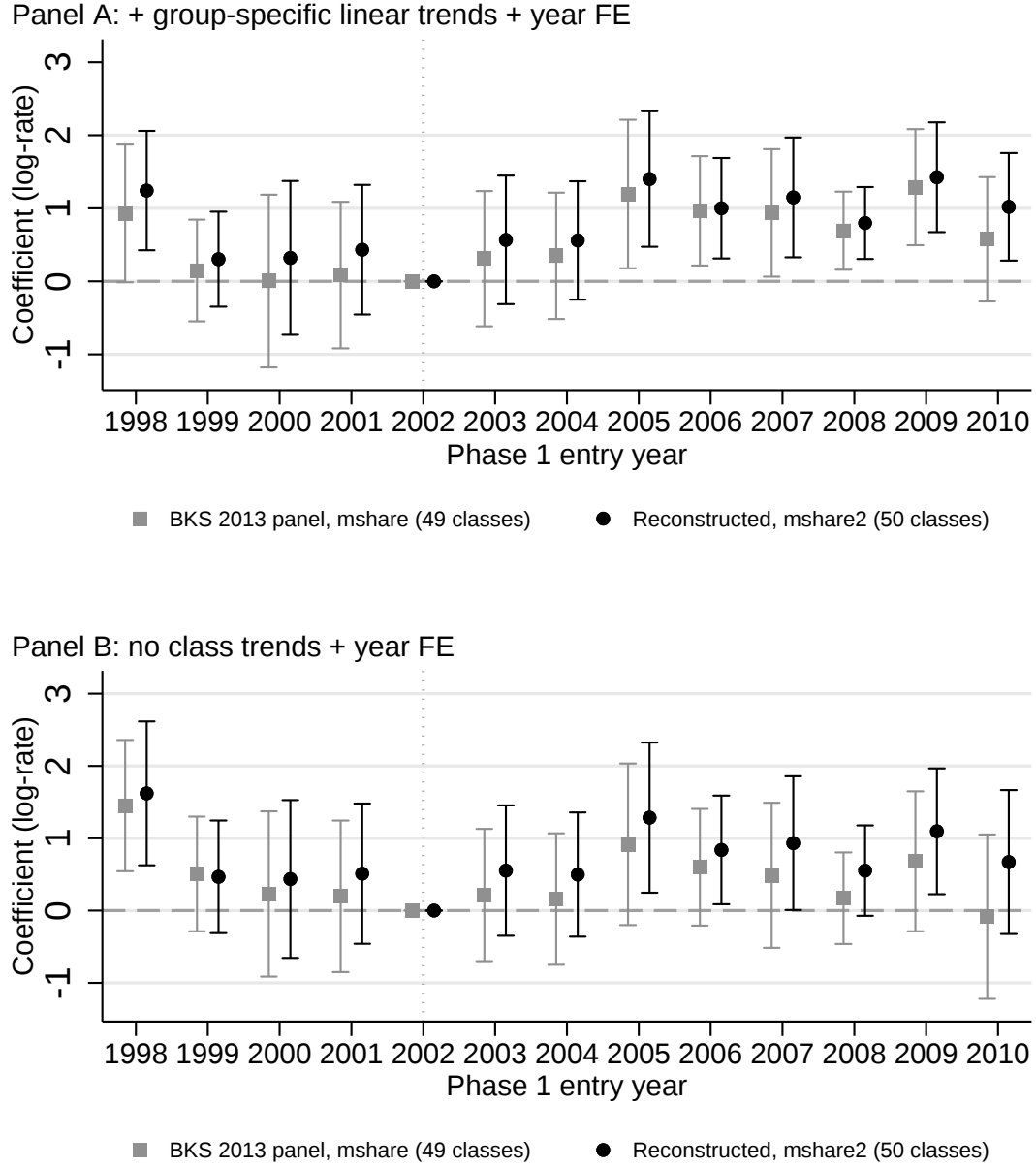


Figure 3: Unconditional event-study coefficients on Medicare share $\times$ cohort-year, with year fixed effects. Two series per panel: legacy 2013 panel (49 classes, original `mshare`) and reconstructed panel (50 classes, refined `mshare2`). Panel A adds class-specific linear trends (`i.group  $\times$  c.time`); Panel B includes only class fixed effects and year fixed effects. Neither panel includes the 2013 paper’s covariates ($\log pv_8$ or the LL_4 – LL_{15} NIH-lag controls). The 2002 reference year is marked at zero.

Table 2 reports the aggregated evidence across four post-MMA sub-periods following Blume-Kohout and Sood (2013). Column (1) reproduces the published 2013 estimates for comparison; Columns (2)–(4) report new Poisson PML regressions on the reconstructed 50-class panel.

Column (2) retains our 2013 specification with `mshare` and class-specific linear trends. The post-period coefficients reproduce the published 2013 values to within 0.08 in every post-period, indicating that the reconstructed panel recovers the 2013 identification at the class-year cell on which the original estimates were computed. Column (3) substitutes `mshare2` for `mshare` under the same specification. The recoding affects only two of the fifty classes, and the coefficients differ from Column (2) by less than 0.05 in every period; the `mshare2` refinement is therefore not material at the Phase 1 entry margin.

Column (4) replaces the class-specific linear trends with year fixed effects, retaining `mshare2` and covariates ($\log pv_8$ and the LL_4 – LL_{15} NIH-funding lags). This is a conventional TWFE design: class fixed effects absorb time-invariant class-level differences, year fixed effects absorb common-time shocks, and the four `mshare2`-by-period coefficients identify the post-MMA Phase 1 entry response in non-zero-`mshare2` classes relative to the two zeroed Part-A/B reference classes that the `mshare2` refinement provides. The estimated coefficients are 0.51 (2003, 90% CI $[-0.23, 1.24]$), 0.99 (2004–05, $[0.28, 1.70]$), 1.29 (2006–07, $[0.57, 2.01]$), and 1.39 (2008–10, $[0.46, 2.32]$). The three later periods reject zero at $p < 0.05$ asymptotically. The Column (4) coefficient magnitudes demonstrate the 2013 post-MMA Phase 1 entry response persists under conventional TWFE identification that the `mshare2` refinement makes feasible, without recourse to class-specific trends.

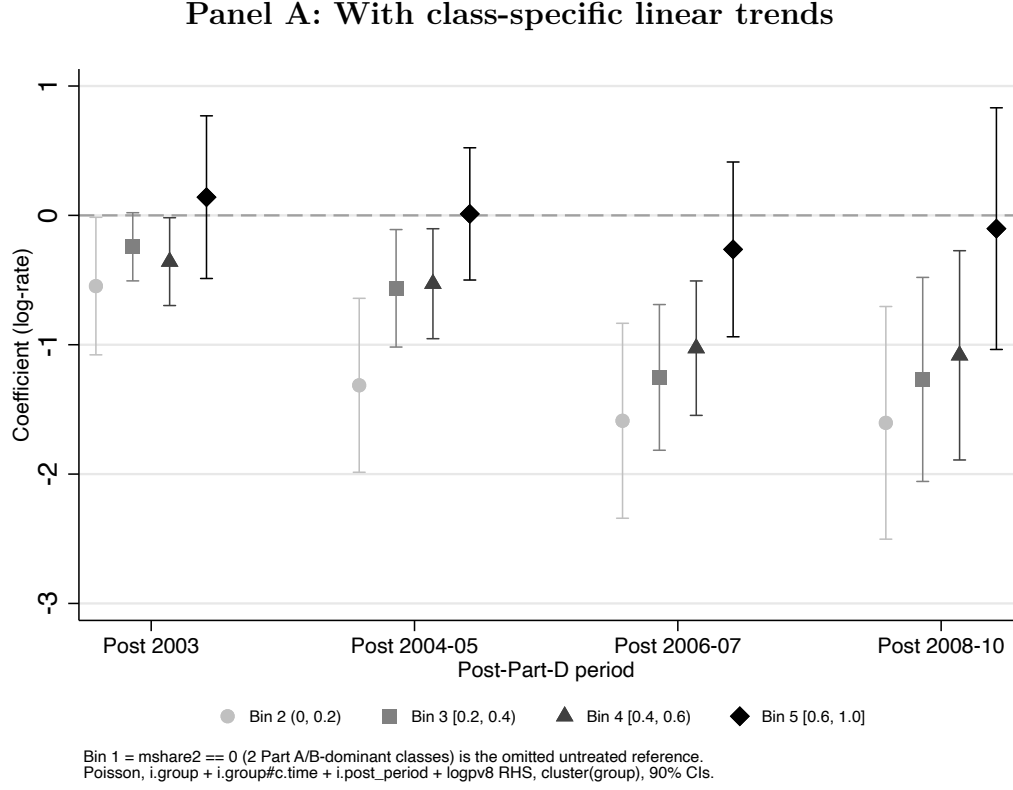
Table 2: Medicare Part D effects on Phase 1 entry: replication and refinement

	(1)	(2)	(3)	(4)
Medicare Share $\times$ Post 2003	0.353 [-0.03,0.73]	0.377 [-0.01,0.77]	0.371 [-0.03,0.77]	0.507 [-0.23,1.24]
Medicare Share $\times$ Post 2004-05	0.640 [0.33,0.95]	0.641 [0.32,0.96]	0.623 [0.30,0.95]	0.987 [0.28,1.70]
Medicare Share $\times$ Post 2006-07	0.810 [0.36,1.26]	0.879 [0.43,1.33]	0.832 [0.37,1.29]	1.294 [0.57,2.01]
Medicare Share $\times$ Post 2008-10	1.251 [0.60,1.91]	1.327 [0.66,1.99]	1.279 [0.60,1.96]	1.392 [0.46,2.32]
Cumulative NIH Funding (4-15 Yr Lags)	0.039 [-0.53,0.61]	-0.023 [-0.60,0.55]	0.012 [-0.55,0.58]	0.232 [-0.11,0.58]
LN(PV(8-20 Yr Lead Market))	4.321 [-11.77,20.42]	2.797 [-13.44,19.03]	2.704 [-13.85,19.26]	-5.845 [-9.76,-1.93]
N	637	650	650	650

Notes: Phase 1 entry count Poisson regression on (class, year) panel. “Cumulative NIH Funding (4–15 Yr Lags)” reports the sum of the LL_4 – LL_{15} NIH-lag coefficients, tested by `lincom`. Brackets report asymptotic 90% confidence intervals (clustered at the therapeutic-class level).

Columns: (1) Original 2013 published estimates, legacy 49-class panel, `mshare`, class fixed effects and class-specific trends; (2) reconstructed 50-class panel, legacy `mshare` treatment variable, same specification with trends as in (1), $N = 650$; (3) reconstructed panel with new `mshare2` treatment variable, same specification with trends as in (1) and (2), $N = 650$; (4) reconstructed panel, `mshare2`, *no class-specific trends*, conventional TWFE with year fixed effects, full 1998–2010 panel ($N = 650$).

Fig. 4 reports the 5-bin `mshare2` $\times$ post-period specification estimated with and without class-specific linear trends. The two panels answer different empirical questions about the post-Part-D Phase 1 entry response. Panel A, which includes `i.group` $\times$ `c.time`, tests whether the *trajectory* of `mshare2`-driven entry shifted post-Part-D, relative to each class’s own pre-existing linear trend as in the 2013 paper. Under this specification, bin 2–4 coefficients turn negative in 2008–10 (−1.6 to −1.1), reflecting that entry in those classes fell below their pre-existing upward trends by the end of the period, not that levels declined in absolute terms. Panel B, which drops the class-specific trends and uses year fixed effects as the common-time absorber, tests whether the *level* of `mshare2`-driven entry shifted post-Part-D. The four non-zero-`mshare2` bins all show comparable post-2008–10 lifts (+1.3 to +1.7), consistent with the broad cash-flow-shock framing in Krieger et al. (2022). Both specifications are empirically valid and answer different questions; the specification choice, not the data, drives the apparent disagreement.



Panel B: Without class-specific linear trends, year fixed effects

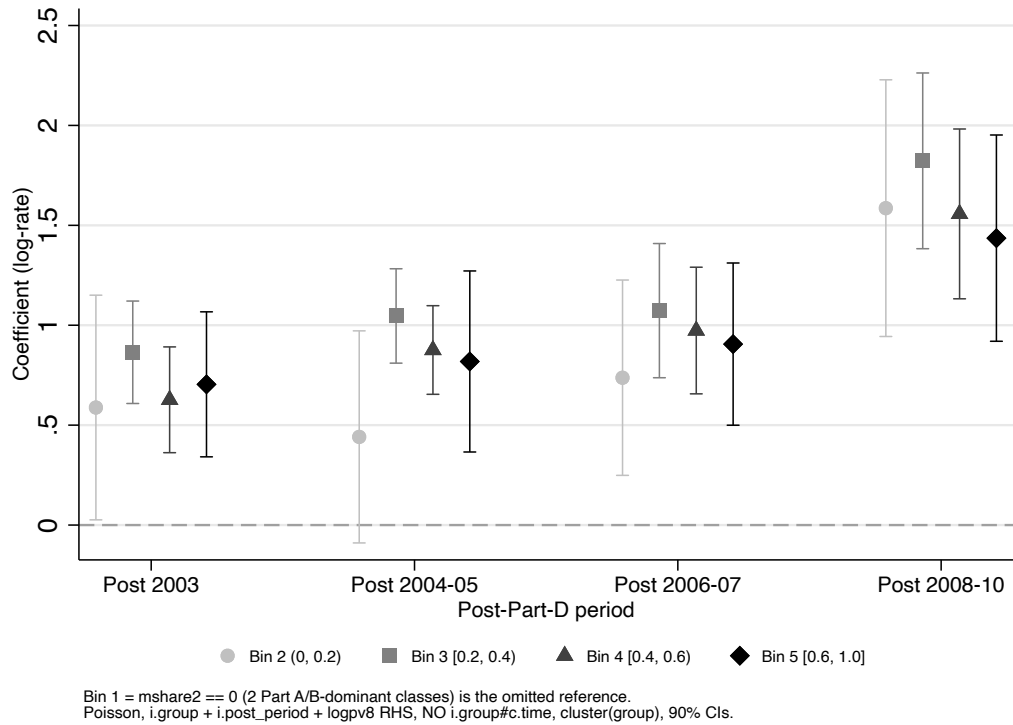


Figure 4: Phase 1 entry: 5-bin `mshare2` $\times$ post-period coefficients. Bin 1 (`mshare2` = 0, two Part-A/B-administered classes) is the omitted reference in both panels. 90% confidence interval bars. Panel A: class-specific linear trends (`i.group#c.time`) and post-period fixed effects (`i.post_period`). Panel B: year fixed effects (`i.year`), no class-specific trends. Both panels: Poisson regression with class fixed effects, $\log pv_8$ and the LL_4 – LL_{15} NIH-lag controls, cluster-robust standard errors at the therapeutic-class level.

4 FDA approval rate per Phase 1 trial

Table 3 reports approval-rate estimates on three samples: the full cohort (Column 1), Large sponsors only (Column 2), and Small sponsors only (Column 3), using the Krieger et al. (2022) sponsor classification refined to FDA-approved NMEs (see Table 1 notes). The sample restriction is applied at the project level before collapsing to (class $\times$ year), so each subsample produces a different `avg_p1` exposure. Four classes are dropped from the Large sample because they had no Large-sponsor Phase 1 entry reported in 1998–2002 (*Antiparasitics & Antimycobacterials*, *Hormones: Contraceptives & Fertility*, *Hormones: Menstruation Disorders*, *Ophthalmic: Glaucoma*). Asymptotic cluster-robust 90% confidence intervals are reported at the therapeutic-class level. The specification omits class fixed effects and year or period fixed effects for the reasons given in Section 2.

The estimates in Table 3 bear directly on the on-the-shelf hypothesis we flagged in 2013. Under quality-constant demand-pull, the `mshare2`-by-Post interactions should be near zero; under the on-the-shelf reading, they should be negative in 2003 through 2007, when pre-existing marginal candidates were being pushed into Phase 1 in anticipation of (or shortly after) the demand expansion.

All three samples produce negative `mshare2`-by-Post coefficients in the 2004 and 2005 cohorts: full sample -0.50 , $[-1.67, 0.67]$; Large sponsors -0.13 , $[-2.54, 2.28]$; Small sponsors -0.98 , $[-2.46, 0.51]$. The directional consistency across samples supports an on-the-shelf reading for the post-passage pre-implementation period, but no individual coefficient rejects zero at conventional thresholds. The Small-sponsor point estimate is the largest in magnitude, consistent with small firms being disproportionately likely to push marginal candidates in response to the anticipated market expansion, though the imprecision allows only a weak claim. The 2006–07 coefficients are similarly negative across samples (full -0.23 , $[-1.55, 1.09]$; Large -0.20 , $[-2.80, 2.40]$; Small -0.67 , $[-2.75, 1.41]$); none rejects zero.

The 2008–10 results may reflect a different mechanism. Post-implementation, incumbent firms could be expanding R&D in response to sustained revenue from Part D sales, as in the cash-flow-driven expansion described in Krieger et al. (2022). The full-sample Post 2008–10 `mshare2` coefficient is -1.22 , $[-2.74, 0.31]$, directionally negative but not rejecting zero. The Large-sponsor coefficient is more strongly negative (-4.51 , $[-9.58, 0.57]$) but also fails to reject zero asymptotically. Both directions are consistent with incumbents expanding into higher-Medicare-share classes with riskier candidates funded by post-Part-D cash flows, rather than with pre-existing pipeline shelves being emptied. This mechanism is consistent with the on-the-shelf hypothesis’s broader prediction that demand-side support attracts lower-quality investments, but operates through a distinct cash-flow channel rather than through

Table 3: Medicare Part D effects on FDA approval rate per pre-Part-D Phase 1 trial, by sponsor sample

	(1) Full sample	(2) Large sponsors	(3) Small sponsors
Medicare share $\times$ 2003	0.534 [-1.36, 2.43]	-0.157 [-3.98, 3.66]	1.190 [-0.87, 3.25]
Medicare share $\times$ 2004–05	-0.497 [-1.67, 0.67]	-0.128 [-2.54, 2.28]	-0.976 [-2.46, 0.51]
Medicare share $\times$ 2006–07	-0.230 [-1.55, 1.09]	-0.200 [-2.80, 2.40]	-0.668 [-2.75, 1.41]
Medicare share $\times$ 2008–10	-1.217 [-2.74, 0.31]	-4.505 [-9.58, 0.57]	-0.198 [-1.60, 1.21]
Cumulative NIH funding (4–15 yr lags)	0.154 [0.03, 0.28]	0.503 [0.29, 0.72]	0.125 [0.00, 0.25]
Forward market size (log PV)	0.094 [-0.07, 0.25]	0.237 [-0.01, 0.49]	0.178 [-0.02, 0.38]
Observations	650	598	650
Classes	50	46	50

Notes: Poisson PML on (class $\times$ year) panel, 1998–2010, with the class-level 1998–2002 mean of Phase 1 entries (**avg_p1**) as an exposure offset. Coefficients are log of the approval rate per pre-Part-D Phase 1 trial. Column (1) uses the full cohort. Column (2) restricts to projects sponsored by Large firms (≥ 10 FDA-approved NMEs at Phase 1 entry minus one). Column (3) restricts to projects sponsored by Small firms (no FDA-approved drugs at entry). **mshare2**: the 2013 paper’s Medicare patient share with two Part-A/B-administered classes (Cancer: Antimetabolite, Cardiovascular: Inotropic) recoded to zero. **Cumulative NIH funding** is the sum of the LL_4 – LL_{15} NIH-lag coefficients, tested by **lincom**. **logpv8**: log present value of forward 8–20 year lead market per the 2013 paper. Class fixed effects dropped to retain the 20 cohort classes with zero approvals across 1998–2010 (which class fixed effects would discard via perfect prediction); group-specific linear trends dropped per pretrend evidence in [Fig. 3](#). Brackets report asymptotic 90% confidence intervals (clustered at the therapeutic-class level).

anticipatory pre-positioning. The point estimates lean strongly negative in the Large-sponsor sample, though the cluster count is too small to identify the differential precisely.

The cumulative NIH-funding response is strongly positive in the Large-sponsor sample (sum of LL_4 – LL_{15} coefficients = 0.50, [0.29, 0.72], $p < 0.001$); the same coefficient is 0.13, [0.00, 0.25], $p = 0.087$ in the Small-sponsor sample and 0.15, [0.03, 0.28], $p = 0.043$ in the full sample. Because the dependent variable is the approval count conditional on Phase 1 trial volume, the summed coefficient is an elasticity: it measures the response of the approval rate per Phase 1 trial to a sustained disease-level increase in NIH R&D funding. For Large sponsors, the 0.50 estimate implies that a 10-percent increase in disease-level NIH funding 4-to-15 years prior to Phase 1 entry is associated with a roughly 5-percent increase in the approval rate per trial, consistent with NIH-funded research raising the probability of subsequent approval, with large established sponsors better positioned to translate that foundation into successful drug launches.

5 Conclusion

Two questions motivated this paper. Did the Phase 1 entry response that we attributed to Medicare Part D in 2013 translate into approved drugs? And does the original Medicare-share gradient survive the specification choices that the post-2012 DiD literature recommends?

For Phase 1 entry, the reconstructed panel reproduces the 2013 estimates within 0.08 in every post-period coefficient under the original specification. Two additional refinements confirm a positive post-MMA Phase 1 entry response. The first is a conventional two-way fixed effects design that replaces the 2013 paper’s class-specific linear trends with year fixed effects, retaining the continuous Medicare-share index as treatment. Its post-period coefficients match or modestly exceed the with-trends estimates, reaching 1.3 in 2006–07 and 1.4 in 2008–10 against 0.8 and 1.3 respectively under the 2013 specification. The second discretizes the Medicare-share index into five bins, following the dose-discretization remedy that the recent continuous-treatment DiD literature recommends, and identifies a separate coefficient for each non-zero bin against the two zero-treatment reference classes. All four non-zero bins show comparable increases relative to the zero-treatment reference classes.

FDA approval rates per Phase 1 trial decline with Medicare share in the post-Part-D cohorts, directionally consistent with the on-the-shelf hypothesis we flagged in 2013. In the 2004–07 cohorts, point estimates are negative across the full cohort and for both sponsor subsamples — large established firms and small firms with no prior approvals — though no individual coefficient rejects zero at conventional thresholds. In the 2008–10 cohorts, the negative point estimate is largest among large sponsors at -4.5 (90% CI $[-9.58, 0.57]$),

directionally consistent with incumbent firms expanding into riskier candidates funded by sustained Part-D cash flows. The imprecision in the estimates may partly be due to the strict indication-level matching for approvals which reduces the apparent "success rate" from Phase 1 well below the 10% expected, leading to an excess of zeros and less range variation in the sample. Several of the highest-Medicare-share classes also have relatively low pre-Part-D Phase 1 entry rates, and the present specification cannot separate demand-pull failure from pipeline-capacity or class-maturity explanations.

Ongoing work examines sensitivity to more permissive approval-matching and to project-level rather than class-aggregate outcomes.

References

- Blume-Kohout, M. E. 2012. “Does targeted, disease-specific public research funding influence pharmaceutical innovation?” *Journal of Policy Analysis and Management* 31 (3): 641–660.
- Blume-Kohout, M. E., and N. Sood. 2013. “Market size and innovation: Effects of Medicare Part D on pharmaceutical research and development.” *Journal of Public Economics* 97:327–336.
- Callaway, B., A. Goodman-Bacon, and P. H. C. Sant’Anna. 2024a. Difference-in-differences with a continuous treatment. NBER Working Paper 32117. National Bureau of Economic Research.
- . 2024b. “Event studies with a continuous treatment.” *AEA Papers and Proceedings* 114:601–605. <https://doi.org/10.1257/pandp.20241047>.
- Chaisemartin, C. de, X. d’Haultfœuille, and G. Vazquez-Bare. 2024. “Difference-in-difference estimators with continuous treatments and no stayers.” *AEA Papers and Proceedings* 114:610–613.
- Cho, J., Q. Xu, C. H. Wong, and A. W. Lo. 2025. “Predicting clinical trial duration via statistical and machine learning models.” *Contemporary Clinical Trials Communications* 45:101473. <https://doi.org/10.1016/j.conctc.2025.101473>.
- Duggan, M., and F. M. Scott Morton. 2010. “The effect of Medicare Part D on pharmaceutical prices and utilization.” *American Economic Review* 100 (1): 590–607.
- Krieger, J., D. Li, and D. Papanikolaou. 2022. “Missing novelty in drug development.” *The Review of Financial Studies* 35 (2): 636–679.