

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

Meg Blume-Kohout[†]

September 25, 2026

Abstract

Blume-Kohout and Sood (2013) demonstrated a rapid increase in pharmaceutical R&D as measured by Phase 1 trials following passage of Medicare Part D. This paper re-examines and expands the data with several extensions to the original design. I recode two Part-A/B-administered classes to zero to provide a zero-treatment reference for the continuous treatment specification, then re-estimate Phase 1 entry with both an unconditional event study and with binned treatment. I then link Phase 1 trials by product-indication to FDA approvals to test the on-the-shelf hypothesis. Using a two-way fixed-effects design that the zero-treatment reference makes feasible, I find similar results to those in 2013 for Phase 1 entries. Conversion from Phase 1 entry to FDA approval falls after passage of Part D, consistent with the on-the-shelf hypothesis that Part-D-induced entries were marginal investments.

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

Keywords: Pharmaceutical innovation, market size, R&D, drug development, prescription drug insurance

*Prepared for presentation at the Annual Health Econometrics Workshop, 2026. The author acknowledges Anthropic's Claude Code for data-engineering assistance on the indication-aware FDA-approval mapping and crosswalk reconstruction, and for the estimation and exhibit-generating code underlying the reported tables and figures; all specification and analytical choices, interpretations, and conclusions are the author's own.

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

1 Introduction

Market size is a major driver of pharmaceutical innovation. Early studies demonstrated that a 1 percent increase in the potential market for a drug category, driven by US demographic change, is associated with a 4 to 6 percent increase in the entry of new molecular entities (NMEs) (Acemoglu and Linn 2004), and policies that expanded utilization of existing vaccines are associated with a 2.5-fold increase in clinical trials for new vaccines (Finkelstein 2004). However, subsequent studies have estimated elasticities of NMEs with respect to market size (as measured by revenues) as low as 0.23 (Dubois et al. 2015). The wide range of estimates across the literature has direct impact for policy projections. For example, the Congressional Budget Office’s estimate of roughly 13 fewer new drugs over 30 years under the Inflation Reduction Act implies an elasticity near 0.2, whereas an analysis of a precursor bill applying an elasticity near 1.5 projected 135 fewer drugs over a comparable time horizon (Maini et al. 2026). Medicare Part D provides a useful setting to investigate this spread, as a large and precisely dated expansion in the market for prescription drugs. Our 2013 study used this policy change to investigate the effects of increased market size on pharmaceutical R&D efforts as measured by clinical trial activity (Blume-Kohout and Sood 2013)

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.

Near-term responses to Part D are well-documented. The program moved prices and utilization in the drugs most used by Medicare beneficiaries (Duggan and Scott Morton 2010), and raised early-stage clinical development in the therapeutic classes with higher Medicare market share (Blume-Kohout and Sood 2013). More recent studies have examined the composition of that increased development (follow-on rather than scientifically novel candidates), its financing through firm cash flows, and its sensitivity to coverage design (Agha et al. 2022; Dranove et al. 2022; Krieger et al. 2022). Whether the induced clinical trials ever resulted in marketed products is unknown, because the follow-up window was not long enough to observe them through approval or discontinuation. That final disposition matters for estimates of elasticities of innovation with respect to market size. Counts of launches cannot distinguish between firms’ undertaking more projects or simply pursuing better ones, and fewer launches may not correspond to lower R&D effort. On the other hand, if a demand expansion increases clinical trial entries while lowering the rate at which entrants convert to approvals, an elasticity measured at entry overstates the elasticity relevant to the drugs

patients eventually receive. This paper addresses this question by testing directly whether the Phase 1 entries observed by Blume-Kohout and Sood (2013) had similar success rates. For each therapeutic class and each Phase 1 entry cohort year from 1998 through 2010, I count the projects that reached FDA approval within fifteen years of entry, and I scale that count by the number of Phase 1 entries in the same class and cohort year. The estimated coefficients test whether the probability that a Phase 1 entrant becomes an approved drug changes after Part D.

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, anticipating market expansion. If this were true, drugs added to the pipeline in response to Part D could be riskier investments with lower probability of success. We observed Phase 1 entry in dual-eligible classes more than doubled, with little short-run Phase 3 follow-through, and cited literature attributing Phase 2 failure to lack of efficacy. However, the hypothesis could not be tested at that time, because many projects were still in active development by our data’s 2010 endpoint. The 15 years elapsed since then now makes the test possible. If quality remained constant, the approval rate per Phase 1 trial should remain unchanged; if marginal products were pushed into clinical testing, the approval rate per Phase 1 trial should fall.

Our 2013 paper noted an apparent Medicare-share gradient in the Phase 1 response, similar to the gradient in revenues (Duggan and Scott Morton 2010); subsequent firm-level evidence points instead to a broadly uniform cash-flow lift in approvals (Krieger et al. 2022). That framing applies most cleanly to firm-level activity, which a class-year panel cannot resolve directly. Whether a class-year panel reads as a gradient or as a uniform lift also depends on the specification. This paper’s Medicare share recode, described in Section 2, supplies the zero-treatment reference that a conventional two-way fixed-effects design requires, and Section 3 reports both designs on the same panel.

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 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 controls for the 4th through 15th lags of logged NIH research funding 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, as legacy intermediate files (parsed by project) were not archived, and estimating conversion rates required project-level drug matching to FDA approvals data. 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 data contain 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.

For the Blume-Kohout and Sood (2013) Phase 1 entry replication, the 1998–2010 cohort window is partitioned into five periods aligned with the post-period structure of the 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). To improve precision, the conversion-rate analysis estimates a single coefficient for all post-implementation years 2006-10. 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.

2.3 Coding treatment intensity

The treatment-intensity measure carried forward from Blume-Kohout and Sood (2013) is the survey-weighted class-level share of prescriptions filled by Medicare beneficiaries in the Medical Expenditure Panel Survey (MEPS) in 2004–05, before Part D took effect. It ranges in the cohort universe from 0.007 (Hormones: Contraceptives & Fertility) to 0.865 (Alzheimer’s Disease) and is time-invariant for each class.

Several methodological challenges constrained our 2013 design. First, the treatment variable was continuous. Its construction as a share naturally constrained possible values to the unit interval, but the range of *observed* values included neither zero nor one. This was intentional, as at the time it seemed plausible all therapeutic classes in MEPS would be affected to some degree 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 post-2003 period indicators in a Poisson specification. Recent work has shown that the continuous-treatment TWFE estimator carries a parallel-trends assumption at every dose level, and that weights on the dose-response slopes can flip sign when the response is heterogeneous in the dose (Callaway et al. 2026). That design also lacked a clean control, as every class in MEPS carried a positive

Medicare share. In response to these issues, I recode the Medicare share to zero for two Part-A/B-administered classes, and discretize the recoded share into five bins (Callaway et al. 2024). Outcomes are estimated by Poisson pseudo-maximum-likelihood, which accommodates the count distribution with mass at zero.

The first refinement exploits 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

The discretized Medicare share employs 5 bins, with bin 1 containing the two Part-A/B-administered classes, and bins 2–5 partition the remaining distribution with cutpoints 0.2, 0.4, and 0.6. This partitioning avoids the restrictions imposed by continuous TWFE, which Callaway et al. (2026) 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 Medicare share specification is retained for direct comparison against our 2013 estimates and is used in the conversion-rate analysis.

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. The second is FDA approval within 15 years of Phase 1 entry, aggregated to the class $\times$ cohort year. To count as approval, 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 first

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 is therefore retained at its unrestricted Medicare share value here.

estimating equation follows our 2013 specification:

$$\mathbb{E}[\text{entries}_{ct} \mid X_{ct}] = \exp\left(\alpha_c + \lambda_c t + \sum_{k=1}^4 \beta_k (\text{mshare}_c \times \text{Post}_{k,t}) + \gamma \log pv8_{ct} + \mathbf{X}'_{ct} \boldsymbol{\theta}\right) \quad (1)$$

where c indexes therapeutic class, t indexes calendar year, the α_c are class fixed effects, the $\lambda_c t$ are class-specific linear time trends, mshare_c is the class’s pre-Part-D Medicare share, $\log pv8_{ct}$ is the logged predicted market size, the $\text{Post}_{k,t}$ indicators mark the four post-periods of the cohort partition, and X_{ct} collects the lagged NIH funding controls. This specification matches the 2013 estimating equation directly, and I report it in Column 2 of [Table 2](#) for validation against the original results in Column 1. Columns 3 and 4 then change one thing at a time: Column 3 substitutes the recoded Medicare share under this same specification, and Column 4 drops the class-specific trends $\lambda_c t$ in favor of year fixed effects δ_t in a conventional two-way fixed-effects design holding the recoded Medicare share, predicted market size, and lagged NIH funding fixed.

In this analysis, the construct “market size” comprises two sources of variation. One is demographic: the predicted market measure is the discounted market size eight to twenty years ahead, built from pre-Part-D prescribing rates by age group and demographic projections, following the potential-market construction in Acemoglu and Linn ([2004](#)). The lead window is consistent with evidence that pharmaceutical research responds to anticipated market-size changes. The other source of variation is the policy: discrete, dated, and partly unanticipated. Including the predicted market nets out the demographic component of market growth, so the interaction terms isolate the policy component.

For the approval-rate outcome I estimate

$$\mathbb{E}[\text{approvals}_{ct} \mid X_{ct}] = \text{entries}_{ct} \cdot \exp\left(\beta_0 \text{mshare2}_c + \sum_{k=1}^3 \beta_k \text{mshare2}_c \times \text{Post}_{k,t} + \mathbf{X}'_{ct} \boldsymbol{\theta}\right) \quad (2)$$

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; entries_{ct} is the contemporaneous count of Phase 1 entries in the same (class, cohort-year) cell, entered as an exposure offset; mshare2_c is the recoded Medicare share; $\text{Post}_{k,t}$ marks cohort year t as falling in post-MMA period k (2003, 2004–05, 2006–10); $\mathbf{X}_{ct}$ collects the class-time-varying controls, the logged predicted market and the lagged NIH funding; β_0 is the pre-Part-D Medicare-share gradient in conversion; and β_k identifies the differential approval rate per contemporaneous Phase 1 entry in period k , on the log scale. The unit of observation remains therapeutic class $\times$ cohort year.

The numerator and denominator refer to the same cohort, so β_k is a conversion effect. An offset fixed at pre-Part-D pipeline capacity would instead scale approvals by a quantity the policy itself moved, mixing the entry margin (on which more trials mechanically yield

more approvals) with the conversion margin the on-the-shelf hypothesis is about. Fifty-six of the 650 class-year cells report no contemporaneous Phase 1 entry and leave the estimation sample because the log offset is undefined; however, each of those cells carries zero approvals, so all 77 cohort approvals are retained.

I omit class fixed effects for two reasons. Absorbing them separates 219 of the 594 estimation cells and leaves 30 clusters, discarding the 20 classes that record no cohort approvals at all across 1998–2010. Class fixed effects also absorb the recoded Medicare share, which is class-fixed, so the pre-Part-D gradient β_0 becomes inestimable. The contemporaneous exposure offset controls for differences across classes in pipeline scale directly.

Year and period fixed effects are also omitted for this estimand. Part D arrived as a national policy on a common date. Under the on-the-shelf reading, in which marginal candidates pushed into Phase 1 convert at lower rates, most of the conversion response is common across classes. A time absorber removes that common component and leaves only the cross-class differential, identified against two reference classes that contribute 26 of the 650 class-year cells with a single approval between them. The estimates behave as that argument predicts: the full-sample 2006–10 coefficient is -2.57 with no time absorber, -0.41 with period fixed effects, -0.47 with year fixed effects, and -0.53 with a linear time trend — sign-stable, roughly 80 percent smaller, and imprecise in all three absorbed versions (standard errors near 1.1).

This is the central limitation of this analysis, regardless of the specification chosen. Without time FE, the coefficients in Equation 2 carry any secular decline in success rates along with the Part D response. With time FE, the estimates are too imprecise to determine if the secular decline is small versus non-existent, and differences in pipeline maturity or scientific complexity that correlate with Medicare share could produce similar patterns. The estimates support the claim that approval rates per Phase 1 entry moved in the direction the on-the-shelf hypothesis predicts, but cannot measure how much of that movement Part D caused.

Reported standard errors are asymptotic cluster-robust at the therapeutic class level, with 45–50 therapeutic-class clusters in each regression, and confidence intervals are reported at the 90-percent level throughout.

2.5 Panel reconstruction and pre-trend 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 the legacy code so could not be replayed. However, cell-by-cell residuals are not systematically associated with the Medicare share (the correlation between residual magnitude and a class’s Medicare share is near zero), so the reconstruction gap is uncorrelated with treatment intensity 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	0.037	0.027	0.017	0.018	0.007	0.019
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.

Figure 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 residual middle group is omitted due to its heterogeneity, as it includes both large primarily generic manufacturers (e.g. Teva) and established specialty firms with small portfolios (e.g. Santen, Lundbeck). Total Phase 1 entry across the two reported groups roughly triples from approximately 140 in 1998 to approximately 480 in 2009.

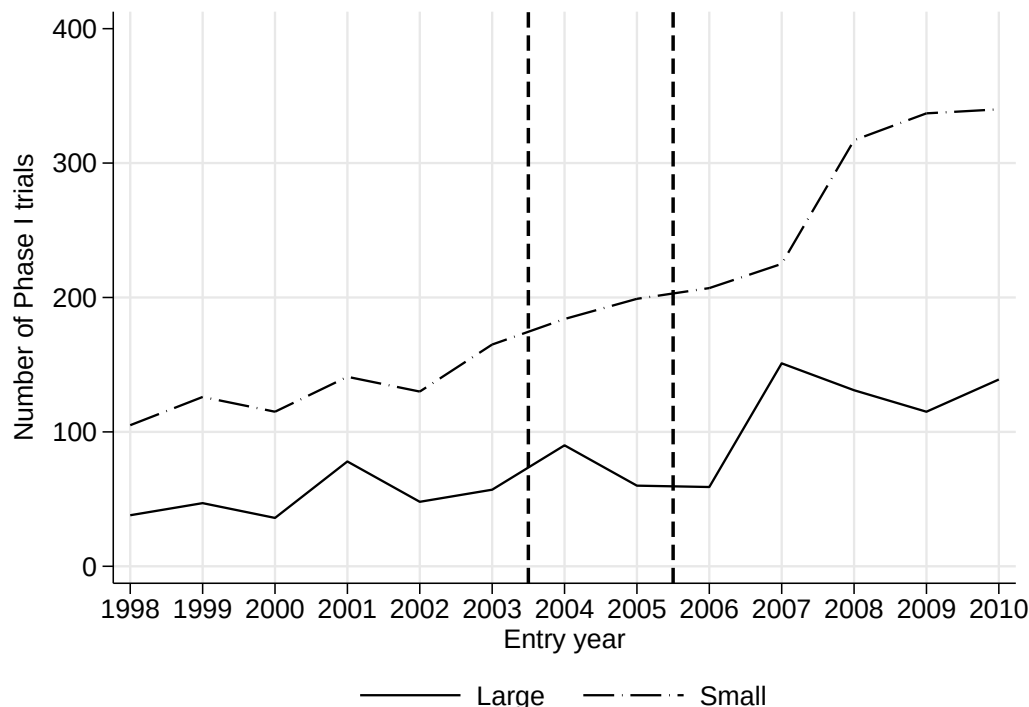


Figure 1: Phase 1 entry counts by year, by Krieger sponsor classification. *Large* = ≥ 10 FDA-approved NMEs in originator’s portfolio at Phase 1 entry minus one (NDA + BLA, M&A-consolidated, per Krieger et al. (2022)); *Small* = no FDA-approved drugs at entry. Mid-sized firms (residual) omitted; see 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

Figure 2 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, the predicted market or the lagged NIH funding: 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.

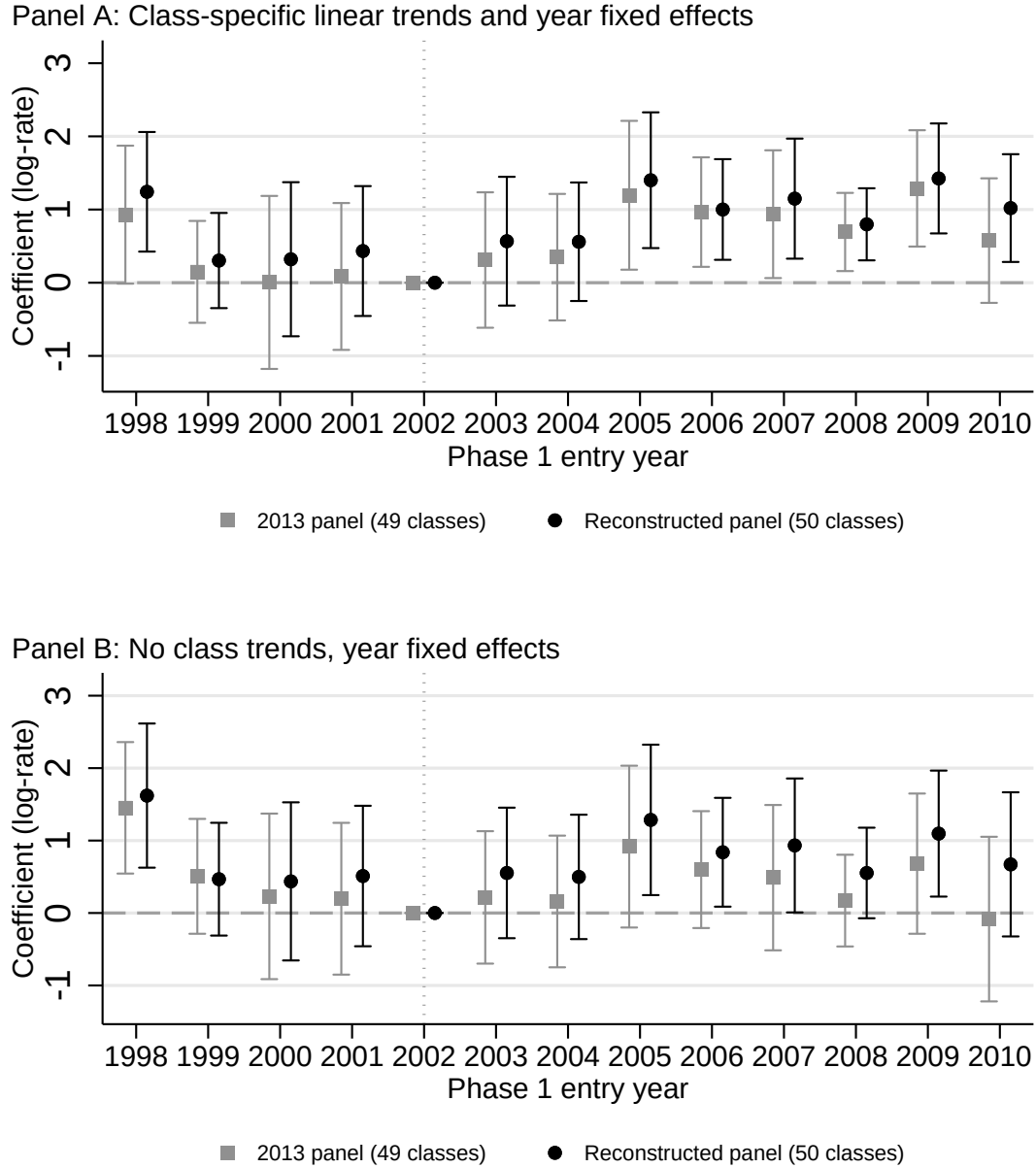


Figure 2: Unconditional event-study coefficients on Medicare share $\times$ cohort-year, with year fixed effects. Two series per panel: legacy 2013 panel (49 classes, original Medicare share) and reconstructed panel (50 classes, recoded Medicare share). Panel A adds class-specific linear trends; Panel B includes only class fixed effects and year fixed effects. Neither panel includes the 2013 paper's covariates, the predicted market or the lagged NIH funding. The 2002 reference year is marked at zero. Bars are 90-percent confidence intervals, cluster-robust at the therapeutic-class level.

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 the original Medicare share 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 the recoded Medicare share 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 recoding is therefore not material at the Phase 1 entry margin.

Column (4) replaces the class-specific linear trends with year fixed effects, retaining the recoded Medicare share, the predicted market and the lagged NIH funding. 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 share-by-period coefficients identify the post-MMA Phase 1 entry response in non-zero-share classes relative to the two zeroed Part-A/B reference classes that the recoding 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 recoding makes feasible, without recourse to class-specific trends.

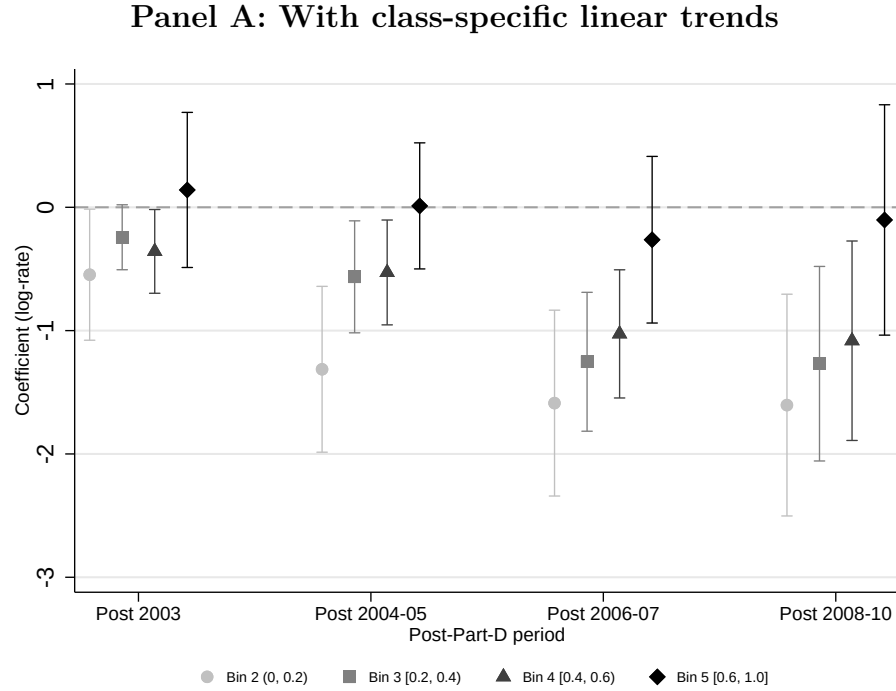
Table 2: Medicare Part D effects on Phase 1 entry: replication and refinement

	(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 lagged NIH funding coefficients, with the standard error of the sum. Brackets report asymptotic 90% confidence intervals (clustered at the therapeutic-class level).

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

Figure 3 reports the five-bin share-by-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 the class-specific trends, tests whether the *trajectory* of share-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 share-driven entry shifted post-Part-D. The four non-zero-share 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

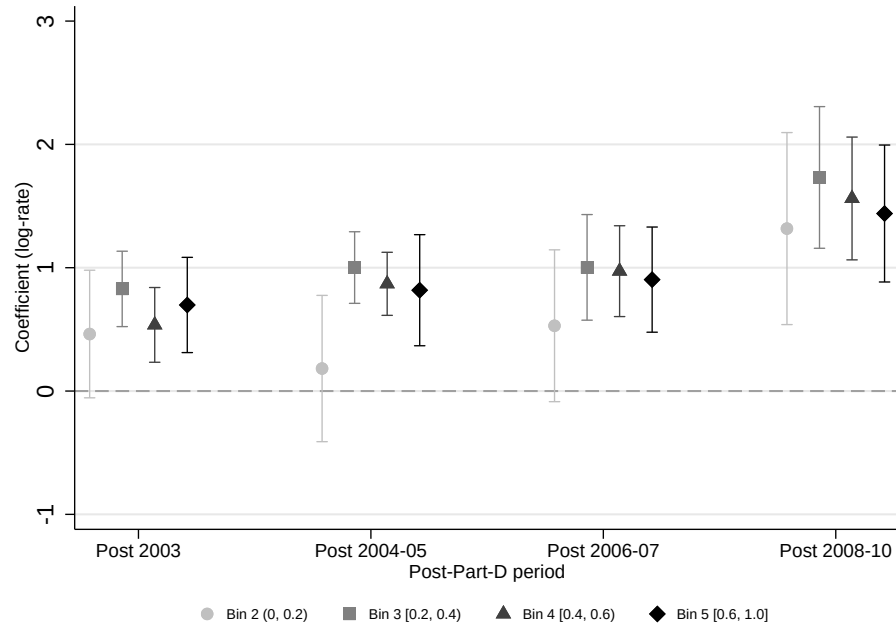


Figure 3: Phase 1 entry: five-bin share-by-post-period coefficients. Bin 1 (the two Part-A/B-administered classes, recoded to a Medicare share of zero) is the omitted reference in both panels. 90% confidence interval bars. Panel A: class-specific linear trends and post-period fixed effects. Panel B: year fixed effects, no class-specific trends. Both panels: Poisson regression with class fixed effects, the predicted market and the lagged NIH funding as controls, and cluster-robust standard errors at the therapeutic-class level.

4 FDA approval rate per Phase 1 entry

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 sponsor-size sample restriction is applied at the project level before collapsing on (class $\times$ year), so each subsample carries its own Phase 1 cohort exposure count. Five classes had no Large-sponsor Phase 1 entry in 1998–2002 (*Acne*, *Antiparasitics & Antimycobacterials*, *GI Ulcer & Antacid*, *Hormones: Menstruation Disorders*, *Ophthalmic: Glaucoma*), and one had no Small-sponsor entry (*Antimalarial*); each leaves the corresponding subsample. 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 explore the on-the-shelf hypothesis. Under quality-constant demand-pull, the share-by-post-period interactions should be near zero; under the on-the-shelf reading, they should turn negative once firms begin pushing pre-existing marginal candidates into Phase 1, in anticipation of the demand expansion and then in response to it.

The first row reports the pre-Part-D gradient. Before the MMA, the Medicare share gradient was at most weakly positive: 1.13 in the full cohort, $[-0.14, 2.39]$; 1.36 among Large sponsors, $[-0.40, 3.11]$; and 0.45 among Small sponsors, $[-1.08, 1.98]$. The post-MMA interactions are differentials against it. The 2003 cohort, entering Phase 1 trials in the year the MMA passed, shows no significant change in approvals.

Conversion falls from the 2004–05 cohorts onward. For the post-passage, pre-implementation cohorts the point estimate is -1.91 in the full sample, $[-3.10, -0.72]$; -3.26 among Large sponsors, $[-6.13, -0.40]$; and -1.07 among Small sponsors, $[-2.94, 0.79]$. For the pooled 2006–10 cohorts the point estimate is -2.57 in the full sample, $[-3.57, -1.58]$; -2.90 among Large sponsors, $[-4.73, -1.07]$; and -1.25 among Small sponsors, $[-2.79, 0.28]$. For small sponsors, the confidence intervals are wide enough that we cannot rule out a null effect; that same imprecision makes it impossible to distinguish the effects for Large vs. Small sponsors as statistically significant. However, a second mechanism may also be operating for Large sponsors in the post-implementation window consistent with the steeper decline. Post-implementation, incumbent firms could be expanding R&D out of sustained revenue from Part D sales, as in the cash-flow-driven expansion documented at the firm level (Krieger et al. 2022). That channel predicts riskier candidates funded by realized cash flows rather than a pre-existing shelf being emptied, but it carries the same implication for conversion, and the -2.90 estimate is consistent with it. For the full sample, a ten-percentage-point higher class Medicare share corresponds to an approval rate per Phase 1 entry roughly 17

Table 3: Medicare Part D effects on the FDA approval rate per Phase 1 entry, by sponsor sample

	(1) Full sample	(2) Large sponsors	(3) Small sponsors
Medicare share (pre-Part-D gradient)	1.127 [-0.14, 2.39]	1.359 [-0.40, 3.11]	0.446 [-1.08, 1.98]
Medicare share $\times$ 2003	-0.024 [-1.47, 1.42]	-0.501 [-2.54, 1.54]	0.807 [-0.97, 2.59]
Medicare share $\times$ 2004–05	-1.908 [-3.10, -0.72]	-3.262 [-6.13, -0.40]	-1.074 [-2.94, 0.79]
Medicare share $\times$ 2006–10	-2.572 [-3.57, -1.58]	-2.903 [-4.73, -1.07]	-1.253 [-2.79, 0.28]
Cumulative NIH funding (4–15 yr lags)	0.158 [0.02, 0.29]	0.180 [-0.01, 0.37]	0.127 [-0.03, 0.29]
Forward market size (log PV)	0.006 [-0.16, 0.17]	-0.074 [-0.28, 0.13]	-0.045 [-0.26, 0.18]
Observations	594	367	541
Classes	50	45	49

Notes: Poisson PML on (class $\times$ year) panel, 1998–2010, with the contemporaneous count of Phase 1 entries in the same (class, cohort-year) cell as an exposure offset. Unit of observation is class $\times$ cohort-year; columns (2) and (3) are not a partition of column (1). 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). “Cumulative NIH funding” is the sum of the lagged NIH funding coefficients. Brackets report asymptotic 90% confidence intervals (clustered at the therapeutic-class level).

percent lower (90% CI: -26.6% , -7.0%) for the 2004–05 cohorts, and roughly 23 percent lower (90% CI: -30.0% , -14.6%) for the 2006–10 cohorts.

The cumulative NIH-funding response is positive in all three samples and similar in magnitude: the sum of the lagged NIH funding coefficients is 0.16 in the full sample, $[0.02, 0.29]$; 0.18 among Large sponsors, $[-0.01, 0.37]$; and 0.13 among Small sponsors, $[-0.03, 0.29]$. Only the full-sample interval excludes zero. Because the dependent variable is the approval count conditional on contemporaneous Phase 1 volume, the summed coefficient is an elasticity: it measures the response of the approval rate per Phase 1 entry to a sustained disease-level increase in NIH R&D funding. The full-sample 0.16 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 1.6-percent higher approval rate per entry. This pattern is consistent with the reading that projects grounded in NIH-funded research have higher probability of subsequent approval.

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. Recoding the two Part-A/B-administered classes to zero moves those coefficients by less than 0.05 in every period, so the refinement that makes a conventional two-way fixed-effects design feasible does not itself change the 2013 answer.

The sign of the entry response survives every specification I estimate; its magnitude does not. Replacing the class-specific trends with year fixed effects increases the 2008–10 coefficient from 1.33 to 1.39, and discretizing the share into five bins leaves the four non-zero bins statistically indistinguishable from one another, so the binned design does not resolve a dose-response.

Conversion from Phase 1 entry to FDA approval declines with Medicare share in the post-Part-D cohorts, in the direction the on-the-shelf hypothesis we flagged in 2013 predicts. In the full cohort the estimate is -1.91 for 2004–05 (90% CI $[-3.10, -0.72]$) and -2.57 for the pooled 2006–10 cohorts $[-3.57, -1.58]$. This suggests a ten-percentage-point higher class Medicare share corresponds to an approval rate per Phase 1 entry roughly 17 percent lower for 2004–05, and 23 percent lower for 2006–10. The point estimates are larger among large established sponsors than among small ones in both post-implementation windows, which would fit incumbents expanding into riskier candidates funded by sustained Part-D cash flows.

That contrast is not identified, however: pooling the two subsamples and interacting every term with sponsor type leaves no differential distinguishable from zero.

The predicted-market control is flat at the conversion margin. It enters the approval-rate equation at 0.006 with a 90-percent interval of $[-0.16, 0.17]$, and the point estimate stays inside roughly $[-0.18, 0.18]$ in every cut I estimate. Whatever moved conversion across these cohorts, the demographic component of market growth was not it.

The elasticity of innovation with respect to market size is therefore not a single number, even holding measurement fixed. As Maini et al. (2026) point out, measurement choices alone — which drugs count as innovation, whose revenue counts as market size — move the estimated elasticity from 0.21 to 0.30 within a single dataset and estimator, and letting the identification strategy vary as well widens the range to 0.20–0.47. The estimates here point to a second dimension: the source of the expansion, and the margin at which changes in innovation are counted.

The same instability is visible inside this paper. Converting the 2008–10 entry coefficient to an elasticity the way we did in 2013 — scaling the Medicare-share response by the revenue increase Part D delivered at that share — gives 4.9 under the 2013 specification and 5.1 under two-way fixed effects. Denominating market size in prescriptions rather than revenue, which is how the predicted-market term in these regressions is built, gives 3.4 and 3.6. Neither of those choices is obviously wrong, and together they move the implied elasticity from about 3.4 to about 5.1.

Of the 1998–2010 Phase 1 entries here, 1.9 percent win a class-matched FDA approval within fifteen years (Table 1); estimates using the same Citeline source put the Phase-1-to-approval rate at 13.8 percent (Wong et al. 2019). Little of that gap comes from the strict indication-level matching: relaxing both the therapeutic-class match and the follow-up window raises the rate only to roughly 2.5 percent. Some of the difference is in the numerator: the approvals are US-only while the Phase 1 entries are worldwide, and approvals outside the 15-year window would be excluded. Some of the difference is in the denominator: the 13.8 percent figure removes programs still in progress from the denominator entirely.

In Blume-Kohout and Sood (2013) we measured R&D effort via entry into each phase of clinical testing. Fifteen years of follow-up adds the rate at which the Phase 1 entries became approved drugs, and that rate moves against the entry response: classes with higher Medicare exposure sent more projects into Phase 1 and converted a smaller share of them. This pattern is consistent with our on-the-shelf hypothesis, though this design cannot rule out pre-existing differences correlated with Medicare share as a source of it. Taken together, the two margins suggest that neither a count of entries nor a count of launches settles the elasticity of innovation with respect to market size.

References

- Acemoglu, D., and J. Linn. 2004. “Market size in innovation: Theory and evidence from the pharmaceutical industry.” *The Quarterly Journal of Economics* 119 (3): 1049–1090.
- Agha, L., S. Kim, and D. Li. 2022. “Insurance design and pharmaceutical innovation.” *American Economic Review: Insights* 4 (2): 191–208.
- 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. 2024. “Event studies with a continuous treatment.” *AEA Papers and Proceedings* 114:601–605. <https://doi.org/10.1257/pandp.20241047>.
- . 2026. “Difference-in-differences with a continuous treatment.” Forthcoming, *American Economic Review*.
- 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>.
- Dranove, D., C. Garthwaite, and M. Hermosilla. 2022. “Does consumer demand pull scientifically novel drug innovation?” *The RAND Journal of Economics* 53 (3): 590–638.
- Dubois, P., O. de Mouzon, F. Scott Morton, and P. Seabright. 2015. “Market size and pharmaceutical innovation.” *The RAND Journal of Economics* 46 (4): 844–871.
- 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.
- Finkelstein, A. 2004. “Static and dynamic effects of health policy: Evidence from the vaccine industry.” *The Quarterly Journal of Economics* 119 (2): 527–564.
- Krieger, J., D. Li, and D. Papanikolaou. 2022. “Missing novelty in drug development.” *The Review of Financial Studies* 35 (2): 636–679.

- Maini, L., M. A. Tran, and J. Feng. 2026. “The effect of market size on pharmaceutical innovation: matching estimates to policy questions.” *Health Affairs Scholar* 4 (7): qxag160. <https://doi.org/10.1093/haschl/qxag160>.
- Wong, C. H., K. W. Siah, and A. W. Lo. 2019. “Estimation of clinical trial success rates and related parameters.” *Biostatistics* 20 (2): 273–286.