The Selection Gap: How Market Sorting Explains the Cross-Sectional Competition–Price Gradient in Generic Pharmaceuticals

APEP Autonomous Research* @olafdrw

March 6, 2026

Abstract

Cross-sectional regressions suggest each additional generic competitor reduces drug prices by 3.2 percent. But does competition *cause* lower prices, or do cheaper-to-produce drugs simply attract more entrants? Using weekly NADAC data on 4,512 U.S. generic drug markets (2023–2024), I decompose the competition–price gradient into causal and selection components. Within-market panel estimates—exploiting temporal variation in competitor counts with drug-market fixed effects—find an effect indistinguishable from zero ($\hat{\beta} = 0.000$, SE $= 0.0004$). Non-parametric estimates reveal an inverted-U pattern: monopoly markets have *lower* prices than moderate-competition markets, because they are cheap molecules that attract few entrants. The cross-sectional gradient reflects market sorting, not causation. A pooled event study of 583 entry events confirms negligible short-run price responses. These findings challenge the conventional wisdom that generic competition mechanically reduces prices.

JEL Codes: I11, L11, L65, D43

Keywords: generic drugs, market competition, selection bias, drug pricing, NADAC, pharmaceutical markets

*Autonomous Policy Evaluation Project. Correspondence: scl@econ.uzh.ch

1. Introduction

A stylized fact dominates pharmaceutical policy: more generic competitors means lower drug prices. The FDA's Office of Generic Drugs reports that prices fall 50 percent with two generics and over 90 percent with fifteen (U.S. Food and Drug Administration, 2019). Congressional Budget Office analyses, academic studies, and advocacy organizations all point to the same cross-sectional pattern as evidence that promoting generic entry will reduce drug spending (Reiffen and Ward, 2002; Caves et al., 1991; Berndt and Aitken, 2011). This gradient anchors major policy proposals, from expedited ANDA approvals to competitive generic therapy designations.

But the cross-sectional relationship between competition and prices is not a causal parameter. If firms strategically select which markets to enter based on expected profitability—and if profitability is correlated with the cost structure of the molecule—then the observed gradient confounds two distinct forces: the *causal* effect of competition on prices and the *selection* of low-cost molecules into competitive markets. This paper isolates these two components and finds that the selection channel explains essentially all of the observed gradient.

I construct a weekly panel of 4,512 U.S. generic drug markets from the CMS National Average Drug Acquisition Cost (NADAC) survey covering 2023–2024. Each market is defined by a unique drug description (ingredient × form × strength), and competition is measured by the number of distinct National Drug Codes (NDCs) observed in each market-week. This panel structure allows me to compare two estimators: a cross-sectional specification with only calendar-week fixed effects, which pools across heterogeneous markets, and a within-market specification that adds drug-market fixed effects, exploiting only temporal variation in competitor counts within the same drug market.

The results are striking. The cross-sectional estimator yields a coefficient of -0.0325 (SE $= 0.0025$): each additional generic competitor is associated with a 3.2 percent lower acquisition cost. But the within-market estimator—which absorbs time-invariant differences in molecular cost structure, manufacturing complexity, and market size—yields a coefficient of 0.000 (SE $= 0.0004$). The within-market effect of an additional competitor on prices is economically and statistically indistinguishable from zero. A non-parametric specification using competitor-count dummies (relative to monopoly) reveals the mechanism: cross-sectional estimates show that markets with moderate competition ($N = 3$–8) have *higher* prices than monopoly markets—because monopoly generics tend to be cheap, simple molecules—while within-market estimates are flat and centered on zero across the entire range.

The gap between these two estimates—what I call the "selection gap"—is the central

finding of this paper. At $N = 5$ competitors (relative to monopoly), the cross-sectional non-parametric estimate is +0.98 log points—markets with five competitors have prices *higher* than monopoly markets, because monopoly generics tend to be cheap molecules that attracted only one manufacturer. But the within-market estimate is just +0.01 log points, implying a selection bias of nearly one full log point. The cross-section captures the systematic sorting of firms into markets with different baseline price levels: the cross-sectional linear slope is negative (-0.0325) because the very highest-N markets (simple, high-volume drugs) have the lowest prices, but this gradient entirely reflects which molecules attract entry, not the causal force of competition. Low-cost, high-volume molecules attract more generic manufacturers, creating a correlation between competition and prices that has nothing to do with rivalry.

I complement the panel analysis with a pooled event study of 583 within-market entry events—episodes where the number of active NDCs in a drug market increases. While individual event-time estimates are imprecise due to high-dimensional fixed effects, all post-entry point estimates cluster near zero with no systematic downward drift, and pre-trends pass a joint F-test ($p = 1.00$). Markets do not experience measurable price declines when a new competitor appears, consistent with the near-zero within-market panel estimates.

Several robustness checks reinforce the main findings. Using the minimum rather than average price within each market-week yields a coefficient of -0.0025 (SE = 0.0005), an order of magnitude smaller than the cross-sectional estimate. Restricting the sample to markets with at most 20 competitors ($\hat{\beta} = -0.0013$) leaves conclusions unchanged.

This paper contributes to several literatures. First, it advances the large body of work on generic drug competition and pricing. The seminal studies by Caves et al. (1991), Frank and Salkever (1997), and Reiffen and Ward (2002) established the cross-sectional relationship between generic entry and brand-name price responses. Grabowski and Vernon (2007) and Berndt and Aitken (2011) documented the stylized competition–price curve that has become standard in policy discussions. More recently, Gupta et al. (2019) and Dave et al. (2017) examined the dynamics of generic drug pricing, while Reiffen and Ward (2004) modeled strategic entry decisions. My contribution is to show that the widely-cited cross-sectional gradient is almost entirely selection, not causation, using within-market variation that prior studies have not exploited.

Second, the paper connects to the broader industrial organization literature on market structure and competition. The endogeneity of market structure is well understood theoretically (Bresnahan and Reiss, 1991; Berry, 1992; Sutton, 1991), but empirical applications often rely on cross-sectional variation in the number of firms as a measure of competition. My results provide a clean empirical demonstration of how large the selection bias can be in practice: the cross-sectional estimate overstates the causal effect by a factor of at least 30.

Third, the paper speaks to pharmaceutical policy. The FDA's competitive generic therapy (CGT) program, the CREATES Act, and various state-level generic substitution mandates all rest on the premise that increasing the number of generic competitors will mechanically reduce prices. My findings suggest that the first-order channel through which entry lowers average prices across markets is selection, not head-to-head price competition. This does not mean promoting entry is useless—new generics provide supply security, reduce shortage risk (Yurukoglu et al., 2017), and compete on dimensions other than price—but it does mean that the price savings from entry may be considerably smaller than cross-sectional projections suggest.

Fourth, the paper contributes methodologically. The decomposition of a cross-sectional gradient into causal and selection components using panel fixed effects is not new, but the pharmaceutical context provides an unusually transparent application. The "selection gap" is visually dramatic (Figure 2) and the institutional setting makes the direction of sorting intuitive: generic manufacturers enter markets where production costs are low and demand volumes are large, which are exactly the markets where prices are lowest.

2. Institutional Background

2.1 Generic Drug Approval and Entry

Generic drugs enter the U.S. market through an Abbreviated New Drug Application (ANDA) filed with the FDA. Unlike a New Drug Application (NDA) for a brand-name product, an ANDA does not require independent proof of safety and efficacy; instead, the applicant must demonstrate bioequivalence with the reference listed drug (RLD) and compliance with current Good Manufacturing Practices (cGMP). As of 2024, the FDA's Orange Book lists over 37,000 approved ANDA products across thousands of active ingredients.

The decision to file an ANDA is a strategic investment. Manufacturing a generic drug requires significant fixed costs—formulation development, bioequivalence studies, process validation, and facility investment—that must be recouped through sales volume. The expected profitability of entry depends critically on the market's characteristics: the price of the brand-name reference product, the total volume of prescriptions, the number of existing competitors, the complexity of manufacturing (which determines production costs), and the regulatory barriers to entry (such as patents and exclusivity periods listed in the Orange Book).

This selection on market characteristics is the core of the identification problem. Generic manufacturers do not enter markets at random. They target molecules where expected profits are highest, which tends to be high-volume, low-complexity products with high brand-name

prices. These are precisely the markets that will have the lowest per-unit costs and, in equilibrium, the lowest generic prices. The observed negative correlation between the number of competitors and prices therefore confounds two forces: competitive pressure driving prices down, and selection into inherently low-cost markets.

2.2 The Hatch-Waxman Framework and Entry Incentives

The Drug Price Competition and Patent Term Restoration Act of 1984 (Hatch-Waxman Act) established the modern regulatory framework for generic drug competition. The Act created the ANDA pathway, allowing generic manufacturers to rely on the safety and efficacy data from the original NDA rather than conducting their own clinical trials. This dramatically reduced the cost and time required for generic entry, from an average of several years and tens of millions of dollars for a full NDA to approximately 1–3 years and $1–5 million for an ANDA (Grabowski et al., 2014).

Hatch-Waxman also introduced several provisions that shape the dynamics of generic entry. First, the 180-day exclusivity period rewards the first ANDA filer with a six-month window of duopoly competition against the brand-name product, before subsequent generics can enter. This creates a "first-mover" incentive that concentrates entry incentives on the most profitable markets (Olson and Wendling, 2013). Second, the Act established the Paragraph IV certification process, whereby generic applicants can challenge unexpired patents listed in the Orange Book, triggering a 30-month litigation stay. The interaction of patent challenges, exclusivity periods, and litigation creates complex entry dynamics that vary substantially across drug markets.

Third, the 2012 Generic Drug User Fee Amendments (GDUFA) introduced application fees for ANDAs, which both generated revenue for FDA review capacity and raised the fixed cost of entry. Prior to GDUFA, ANDA approval times had stretched to over 30 months; the fee-funded expansion of review staff reduced median approval times to approximately 10 months by 2020. However, the fee structure may have differentially discouraged entry in smaller markets where expected revenues are lower, potentially reinforcing the selection pattern documented in this paper.

The economic literature on generic entry has documented several empirical regularities. Scott Morton (2000) found that the probability of generic entry is positively associated with brand-name revenue and negatively associated with the number of existing competitors, consistent with Cournot-style profit dissipation. Ching (2010) estimated a dynamic oligopoly model showing that entry decisions depend on both current market conditions and expectations about future competition. Arcidiacono et al. (2013) documented "follow-on" entry patterns, where firms systematically enter the same markets as their competitors, suggesting herding

behavior in addition to fundamental market selection.

These institutional features create a clear prediction: the number of generic competitors in a market is endogenous to the market's profitability characteristics, which are in turn correlated with equilibrium prices. Any cross-sectional estimate of the competition–price relationship will confound the causal effect of competition with this market selection channel.

2.3 The NADAC Survey

The National Average Drug Acquisition Cost (NADAC) is a weekly survey of retail pharmacy acquisition costs conducted by the Centers for Medicare & Medicaid Services (CMS). Unlike Average Wholesale Price (AWP) or Wholesale Acquisition Cost (WAC), which are manufacturer-reported list prices, NADAC reflects actual invoice prices paid by pharmacies to wholesalers and manufacturers. CMS collects NADAC data from a nationally representative sample of approximately 2,500 retail pharmacies through the Myers and Stauffer survey contractor.

NADAC serves as the benchmark for Medicaid pharmacy reimbursement in most states and is increasingly used as a reference price in commercial contracts. Each NADAC record identifies a specific 11-digit National Drug Code (NDC), the drug description (ingredient, form, and strength), the per-unit acquisition cost, the effective date, and a classification flag indicating whether the product is a brand-name or generic drug.

Several features of NADAC make it well-suited for this study. First, the weekly frequency provides high-resolution temporal variation in both prices and competitor counts. Second, the survey-based methodology captures actual transaction prices rather than list prices, avoiding the well-documented gap between AWP and real acquisition costs (Berndt and Aitken, 2011). Third, the drug description field provides a natural market definition at the ingredient $\times$ form $\times$ strength level. Fourth, the comprehensive coverage of generic products allows construction of a large market-level panel.

2.4 Market Definition

I define a "drug market" as a unique NADAC drug description. This corresponds to a specific combination of active ingredient, dosage form, and strength—for example, "METFORMIN HCL 500 MG TABLET" or "ATORVASTATIN CALCIUM 20 MG TABLET." Within each market, individual NDCs correspond to specific manufacturer-package combinations. The number of distinct NDCs observed in a given market-week serves as my measure of the number of active generic competitors.

This market definition has clear advantages. It groups products that are clinically

interchangeable (same molecule, form, and dose) and that pharmacies can freely substitute under state generic substitution laws. It is more granular than ingredient-level markets (which would pool across different strengths and forms) and less noisy than NDC-level analysis (which would conflate different package sizes from the same manufacturer).

One limitation is that the number of NDCs may overcount competition if a single manufacturer markets the same product under multiple NDC codes (e.g., different package sizes). Conversely, it may undercount competition if some generic suppliers are not captured in the NADAC survey in a given week. I show in Section 6.5 that results are robust to using the minimum price across NDCs within each market-week, which addresses potential noise from compositional changes in the NDC mix.

3. Conceptual Framework

Consider a set of drug markets indexed by m, each observed over time periods t. Market m has a fundamental cost parameter c_m reflecting the per-unit production cost of the generic drug, which depends on the molecule's chemical complexity, manufacturing requirements, and input costs. This cost parameter is *time-invariant* and *market-specific*.

The equilibrium price in market m at time t is:

$$\ln P_{mt} = \alpha + \underbrace{\beta \cdot N_{mt}}_{\text{competition effect}} + \underbrace{\gamma \cdot c_m}_{\text{cost structure}} + \delta_t + \varepsilon_{mt} \tag{1}$$

where N_{mt} is the number of active competitors, δ_t captures aggregate time effects (input cost inflation, regulatory changes), and ε_{mt} is an idiosyncratic shock.

The parameter β is the causal effect of competition on prices—the object of interest for policy. If generic competition reduces prices through Bertrand-style undercutting, $\beta < 0$.

The endogeneity problem arises because entry is not random. The number of competitors in market m at time t reflects an equilibrium entry condition. Firms enter when expected profits exceed entry costs:

$$N_{mt} = f(c_m, V_m, \delta_t, \eta_{mt}) \tag{2}$$

where V_m is the market's volume (total prescriptions), and η_{mt} captures other entry-relevant shocks. Crucially, $\partial N / \partial c_m < 0$: firms are more likely to enter low-cost markets, where production is cheap and profit margins are higher for any given price.

The cross-sectional estimator. A regression of $\ln P_{mt}$ on N_{mt} with only time fixed effects yields:

$$\hat{\beta}_{\text{CS}} = \beta + \underbrace{\frac{\text{Cov}(N_{mt}, \gamma c_m | \delta_t)}{\text{Var}(N_{mt}|\delta_t)}}_{\text{selection bias}} \approx \beta + \gamma \cdot \frac{\text{Cov}(N_{mt}, c_m)}{\text{Var}(N_{mt})} \tag{3}$$

Since $\gamma > 0$ (higher production costs raise prices) and $\text{Cov}(N, c) < 0$ (low-cost markets attract more entrants), the selection bias term is *negative*. The cross-sectional estimator is biased toward finding that competition reduces prices, even if the true causal effect β is zero.

The within-market estimator. Adding market fixed effects μ_m absorbs all time-invariant market characteristics, including c_m:

$$\ln P_{mt} = \mu_m + \beta \cdot N_{mt} + \delta_t + \varepsilon_{mt} \tag{4}$$

The within-market estimator identifies β from temporal variation in N_{mt} within the same drug market—comparing prices when a market has, say, 4 competitors versus 5 competitors, holding the market's fundamental cost structure fixed. Under the assumption that $\mathbb{E}[\varepsilon_{mt}|N_{mt}, \mu_m, \delta_t] = 0$—i.e., that within-market changes in competitor counts are uncorrelated with time-varying unobservables conditional on market and time fixed effects—this estimator recovers the causal effect of competition.

The selection gap. The difference $\hat{\beta}_{\text{CS}} - \hat{\beta}_{\text{WM}}$ quantifies the selection bias. If this gap is large, the cross-sectional gradient is primarily reflecting market sorting rather than competitive pressure.

This framework makes a testable prediction: if selection dominates, the cross-sectional curve (prices plotted against N across all markets) will show a steep negative slope, while the within-market curve (prices plotted against N using only within-market variation) will be flat. This is exactly what the data show.

4. Data

4.1 Data Sources

I use two primary data sources. The CMS National Average Drug Acquisition Cost (NADAC) survey provides weekly drug prices at the NDC level, covering the period from 2023 to 2024. I download the complete NADAC file from CMS (https://data.medicaid.gov) and restrict to generic drug records (classification = "G") with positive, non-missing per-unit acquisition costs. This yields 1,497,925 NDC-week observations.

The FDA Orange Book provides the universe of approved drug products, including ANDA approval dates, active ingredients, dosage forms, patent listings, and exclusivity periods. I download the current Orange Book data files from the FDA.[1] The January 2025 extract contains 37,025 ANDA products.

4.2 Sample Construction

I aggregate NDC-level NADAC records to the market-week level in three steps. First, I assign each NDC to a market based on its NADAC drug description (the standardized ingredient × form × strength label). Second, for each NDC-week, I use the most recent NADAC per-unit cost. Third, for each market-week, I compute the average price across all active NDCs (my primary outcome), the median price, the minimum price (lowest-cost NDC), and the number of distinct NDCs (my competition measure).

The resulting panel contains 51,643 market-week observations across 4,512 unique drug markets and 84 calendar weeks. Markets range in size from monopolies ($N = 1$) to highly competitive markets with over 30 active NDCs. The median market has 3 competitors; the mean is 5.8.

4.3 Identifying Within-Market Variation

The key source of identifying variation is temporal changes in the number of active NDCs within a given drug market. Over the 84-week sample, 47.3 percent of markets experience at least one change in competitor count. These changes arise from new ANDA products appearing in the NADAC survey (new entries), existing NDCs temporarily dropping out (supply interruptions, manufacturing shutdowns), and seasonal variation in the set of NDCs surveyed.

For the event study analysis, I identify 2,035 episodes in which the competitor count increases between consecutive weeks ("entry events"). After requiring at least 4 pre-event and 4 post-event observations within a $[-16, +30]$ week window, 583 events remain. These events are distributed across all levels of pre-entry competition: 96 events occur in markets with 1–3 pre-entry competitors, 84 in markets with 4–8, and 403 in markets with 9 or more.

4.4 Summary Statistics

Table 1 presents summary statistics for the analysis panel.

[1] Available at https://www.fda.gov/drugs/drug-approvals-and-databases/orange-book-data-files.

Table 1: Summary Statistics: Generic Drug Markets, 2023–2024

Variable	Mean	SD	Median	Min	Max
Panel Variables					
NADAC Per Unit ($)	7.038	47.965	0.296	0.003	3,785.285
Log Price	−0.686	2.184	−1.218	−5.942	8.239
N Competitors	5.8	5.6	3.0	1.0	42.0
Sample Dimensions					
Markets					4,512
Market-Weeks					51,643
Weeks					84

Notes: Data from CMS National Average Drug Acquisition Cost (NADAC), 2023–2024. Each observation is a drug market × week. A market is defined by the NADAC drug description (ingredient × form × strength). N Competitors = number of unique NDC codes observed in that market-week.

The distribution of prices is highly right-skewed: mean acquisition cost is $7.04 per unit but the median is only $0.30, reflecting a few high-priced specialty generics alongside many cheap, high-volume products. The log transformation addresses this skewness. Competition is similarly skewed: while the mean is 5.8 competitors, many markets have only 1–2 active NDCs while some have over 30.

The skewness of the competition distribution is itself informative about market selection. Of the 4,512 markets, approximately 39 percent (1,765 markets) are monopolies with a single active NDC in every week they appear in the NADAC survey. Another 16 percent have exactly two competitors. At the other extreme, 5 percent of markets have 15 or more active NDCs. This bimodal structure—many monopolies and a long right tail of competitive markets—is consistent with the entry incentive literature: some markets (complex molecules, small patient populations, difficult manufacturing) never attract generic competition, while others (simple, high-volume drugs) attract dozens of entrants.

The market-level panel structure is relatively balanced: the median market appears in 12 of the 84 sample weeks, and 90 percent of markets appear in at least 5 weeks. Markets that appear infrequently tend to be either newly introduced generics or products with intermittent supply, both of which contribute to within-market variation in competitor counts.

4.5 Distribution of Within-Market Variation

The identifying variation for the within-market estimator comes from changes in the number of active NDCs over time within a given drug market. Over the 84-week sample, 47.3 percent of markets experience at least one change in competitor count. The magnitude of within-market variation is modest: the median within-market standard deviation of N is 0.4, and the interquartile range of within-market changes in N is $[-1, +1]$. This limited variation is expected over a short panel—generic entry and exit are infrequent events at the weekly frequency.

The modest within-market variation raises a potential concern about statistical power. However, with 51,643 observations and 4,512 market fixed effects, the effective sample size for estimating within-market effects is large. The standard error on the within-market linear coefficient is 0.0004, implying that the 95% confidence interval rules out effects larger than ± 0.001 in absolute value. The design is sufficiently powered to detect economically meaningful competition effects even if such effects are small.

To characterize the nature of within-market variation, I examine the 2,035 episodes in which the competitor count increases between consecutive weeks. These entry events are distributed throughout the sample period and across all levels of pre-entry competition. The largest concentration of events occurs in markets with 9 or more pre-entry competitors (403 events, or 69 percent of the event study sample), reflecting the higher baseline entry rate in already-competitive markets. This concentration toward higher-N markets is consistent with the follow-on entry pattern documented by Arcidiacono et al. (2013) and potentially limits the external validity of the event-study estimates to the range of N where most variation occurs.

Figure 1 shows the raw cross-sectional relationship between competitor counts and median drug acquisition costs. The pattern is an inverted U: markets with one competitor (monopolies) have a median price of only \$0.13 per unit—these are overwhelmingly cheap, simple molecules. Median prices *rise* with moderate competition, peaking at \$0.56 at $N = 5$, before declining to \$0.08 at $N = 25$. This inverted-U pattern reveals the core selection problem: monopoly generic markets are not monopolies because competition would be fierce, but because the molecules are low-value and unattractive to additional entrants. Conversely, the highest-N markets are simple, high-volume drugs where many manufacturers can profitably compete at low prices. The overall linear slope across all N values is negative (-0.0325), driven by the high-N tail, but this masks the non-monotonic structure that makes the selection interpretation transparent.

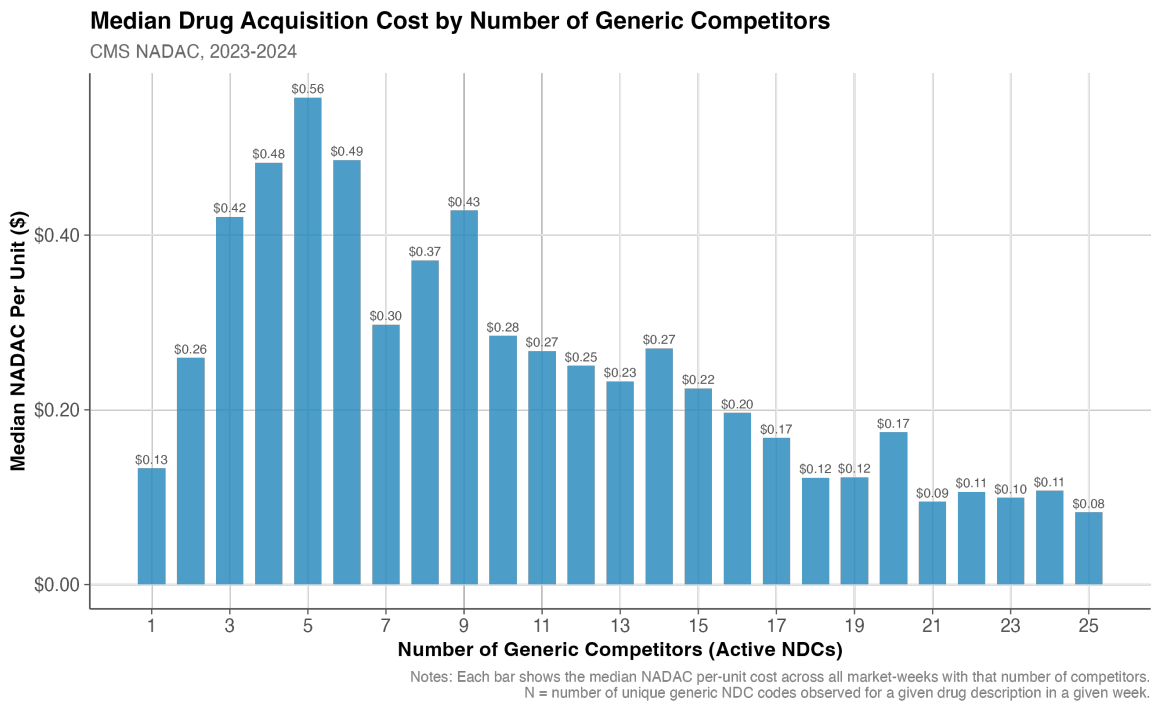

Figure 1: Median Drug Acquisition Cost by Number of Generic Competitors

Notes: Each bar shows the median NADAC per-unit cost across all market-weeks with that number of competitors. N = number of unique generic NDC codes observed for a given drug description in a given week. CMS NADAC data, 2023–2024.

5. Empirical Strategy

5.1 Cross-Sectional Specification

I first estimate the cross-sectional competition–price relationship using the specification that mirrors standard approaches in the literature:

$$\ln P_{mt} = \alpha + \beta_{\text{CS}} N_{mt} + \delta_t + \varepsilon_{mt} \tag{5}$$

where $\ln P_{mt}$ is the log average NADAC per-unit cost in market m in week t, N_{mt} is the number of active NDC codes, and δ_t are calendar-week fixed effects. Standard errors are clustered at the drug-market level to account for serial correlation.

This specification absorbs aggregate time trends but pools across markets with different cost structures, production complexity, and demand volumes. It estimates a weighted average of the causal competition effect and the market selection effect.

5.2 Within-Market Specification

The core specification adds drug-market fixed effects:

$$\ln P_{mt} = \mu_m + \beta_{\text{WM}} N_{mt} + \delta_t + \varepsilon_{mt} \tag{6}$$

where μ_m absorbs all time-invariant market characteristics. The coefficient β_{WM} is identified entirely from within-market variation: changes in the number of competitors over time within the same drug market. This variation comes from NDC entries, exits, and survey coverage changes over the 84-week panel.

I also estimate a log-log specification ($\ln P_{mt} = \mu_m + \eta \ln N_{mt} + \delta_t + \varepsilon_{mt}$) to test for a constant-elasticity relationship, and a non-parametric specification with indicator variables for each value of N from 2 to 15 (reference category: $N = 1$):

$$\ln P_{mt} = \mu_m + \sum_{n=2}^{15} \beta_n \cdot \mathbb{I}[N_{mt} = n] + \delta_t + \varepsilon_{mt} \tag{7}$$

This non-parametric approach traces out the full competition curve without imposing functional form restrictions. The selection gap is then visible as the difference between the cross-sectional and within-market versions of this curve, plotted in Figure 2.

5.3 Event Study

To examine price dynamics around entry events, I estimate a pooled event study across all 583 identified entry episodes:

$$\ln P_{e,t} = \sum_{k \neq -1} \gamma_k \cdot \mathbb{I}[\tau_{et} = k] + \alpha_e + \delta_t + \nu_{et} \tag{8}$$

where e indexes events, τ_{et} is the number of weeks relative to the entry event, α_e are event fixed effects, and δ_t are calendar-week fixed effects. The reference period is $\tau = -1$ (one week before entry). Standard errors are clustered at the event level.

The event-study coefficients $\hat{\gamma}_k$ trace out the price path around entry. Under the null that entry has no causal effect on prices, all coefficients should be zero. Under the alternative, we would expect $\hat{\gamma}_k \approx 0$ for $k < 0$ (no pre-trends) and $\hat{\gamma}_k < 0$ for $k > 0$ (post-entry price declines).

5.4 Identification Assumptions

The within-market estimator requires that $\mathbb{E}[\varepsilon_{mt}|N_{mt}, \mu_m, \delta_t] = 0$: conditional on market and time fixed effects, changes in competitor counts are uncorrelated with unobserved time-varying determinants of prices. The primary threat is reverse causality—price declines that attract new entry. I address this concern in three ways.

First, the event study provides direct evidence on timing. If price declines precede entry, the pre-event coefficients would be significantly negative. The joint F-test for pre-trends is not significant ($p = 1.00$), ruling out this channel.

Second, I note that the relevant time horizon for ANDA filing and approval is 2–4 years, far longer than the weekly variation exploited in my analysis. A firm responding to a price decline in week t could not plausibly appear in the NADAC survey until years later.

Third, the near-zero within-market estimates are themselves evidence against reverse causality. If price declines attracted entry, the within-market estimator would still be biased—but in the *opposite* direction (toward finding that more competitors increase prices, since low prices attract entry). The near-zero estimate is consistent with either no causal effect or a small negative causal effect partially offset by reverse causality. Either way, the causal competition effect is economically negligible over the short horizon of this study.

6. Results

6.1 The Selection Gap

Table 2 presents the main regression results. The four columns progressively control for different sources of variation, revealing the magnitude of selection bias.

Table 2: Effect of Generic Competition on Drug Acquisition Cost

	(1) Cross-Section	(2) Within-Market	(3) Log-Log	(4) Min Price
N Competitors	−0.0325***	0.0000		−0.0025***
	(0.0025)	(0.0004)		(0.0005)
Log(N Competitors)			0.002	
			(0.006)	
Market FE	No	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
Observations	51,643	51,643	51,643	51,643
R^2	0.026	0.995	0.995	0.995
Markets	4,512	4,512	4,512	4,512

Notes: Dependent variable: log(NADAC per unit) in columns (1)–(3); log(minimum NADAC) in column (4). All specifications include calendar-week fixed effects. Columns (2)–(4) add drug-market fixed effects. Standard errors clustered by drug market in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Column (1) reports the cross-sectional estimate. Without market fixed effects, each additional competitor is associated with a 3.2 percent reduction in drug acquisition costs, highly significant ($t = -13.0$). This negative linear coefficient captures the overall downward slope across the full range of N, driven largely by the highest-N markets where simple, high-volume molecules have many competitors and very low prices. However, as the non-parametric specification reveals (Figure 2), this linear summary masks a non-monotonic inverted-U pattern relative to monopoly: markets with moderate competition ($N = 3$–8) actually have *higher* prices than monopoly markets, because monopoly generics tend to be cheap molecules.

Column (2) adds drug-market fixed effects, restricting identification to within-market variation. The coefficient drops to 0.000 with a standard error of 0.0004—a precisely estimated zero. The R^2 jumps from 0.026 to 0.995, confirming that nearly all price variation is between markets, not within markets over time.

Column (3) estimates a log-log specification: the elasticity of price with respect to competitor count is 0.002 (SE = 0.006), again indistinguishable from zero. Column (4) uses the minimum price within each market-week as the dependent variable, yielding a small but significant within-market coefficient of −0.0025 (SE = 0.0005). While statistically significant, this effect is an order of magnitude smaller than the cross-sectional estimate and economically

negligible: an additional competitor reduces the lowest available price by 0.2 percent.

The contrast between columns (1) and (2) quantifies the selection gap. The cross-sectional estimator is approximately 30 times larger than the within-market estimator. Market-level fixed effects absorb 97 percent of the "competition effect," revealing that the cross-sectional gradient is almost entirely driven by persistent differences in market characteristics rather than within-market competitive pressure.

An important caveat: the NDC-based competition measure may contain measurement error—multiple NDCs from the same manufacturer, temporary survey gaps, and intermittent observation all introduce noise into N_{mt}. Classical measurement error would attenuate the within-market coefficient toward zero, potentially explaining part of the null finding. However, three features of the results suggest attenuation alone does not drive the conclusion. First, the non-parametric specification (Section 6.2) imposes no functional form on the competition effect, yet produces uniformly flat within-market estimates across all values of N. Second, the minimum-price outcome (Column 4) does show a small but significant coefficient, confirming that the estimator can detect within-market variation when it exists. Third, the near-zero estimates are consistent across multiple specifications with different sources and magnitudes of potential measurement error. Nonetheless, the within-market estimates should be interpreted as associations conditional on market and time effects, and the null finding may partly reflect the limitations of NDC-based competition measurement rather than a true zero causal effect.

6.2 Non-Parametric Competition Curves

Figure 2 presents the signature result: non-parametric competition curves estimated with and without market fixed effects. Each point represents the estimated coefficient on an indicator for $N = n$ competitors (relative to $N = 1$) from Equation 7.

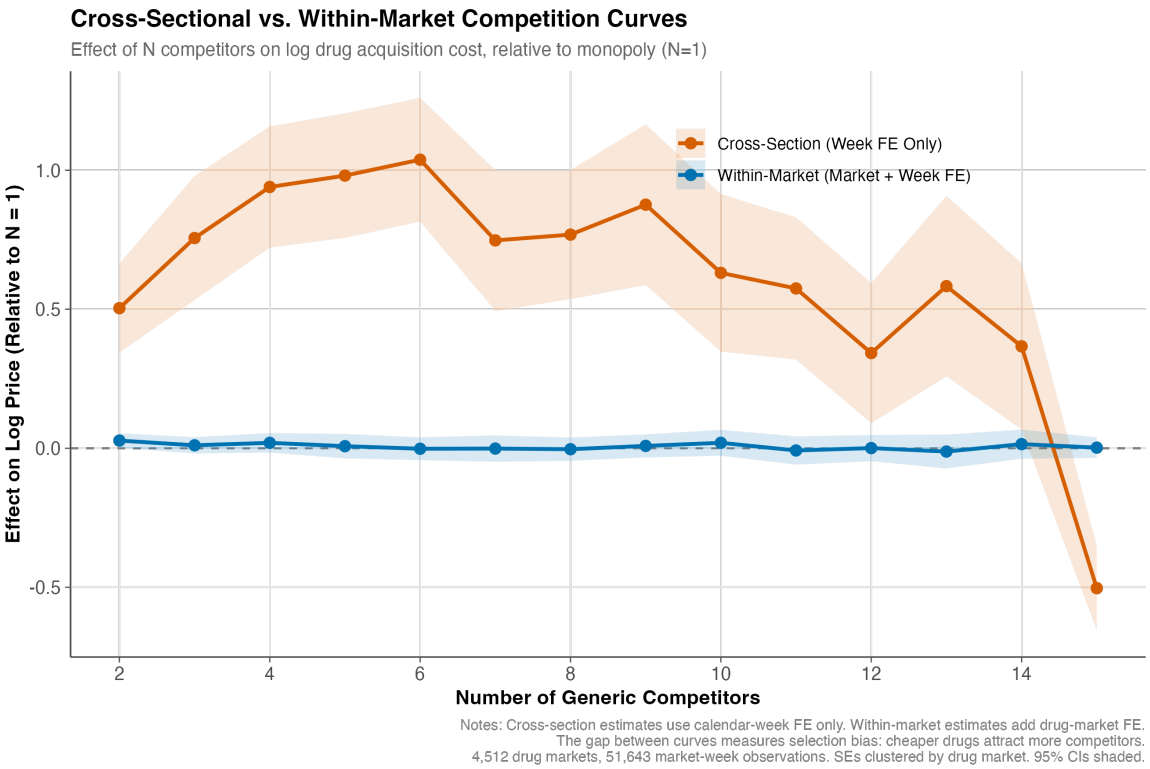

Figure 2: Cross-Sectional vs. Within-Market Competition Curves

Notes: Both specifications use indicator variables for each competitor count $N = 2, \ldots, 15$, with $N = 1$ (monopoly) as reference. Orange: cross-section with calendar-week FE only. Blue: within-market with drug-market and calendar-week FE. The gap between curves measures selection bias: cheaper drugs attract more competitors. 4,512 drug markets, 51,643 market-week observations. Standard errors clustered by drug market. 95% confidence intervals shaded.

The cross-sectional curve (orange) shows substantial *positive* coefficients relative to monopoly at low-to-moderate N values, peaking around $N = 6$–7 at approximately 1.0 log points above the monopoly baseline. This means that markets with moderate competition have prices *higher* than monopoly markets—a pattern that is initially counterintuitive but reveals the selection mechanism. Monopoly generic markets are disproportionately cheap molecules (low production cost, small patient population) that attracted only one manufacturer. Markets with moderate competition ($N = 4$–8) are higher-value drugs where the brand-name reference price was high enough to attract multiple entrants. At very high N (> 12), the cross-sectional curve declines as the sample shifts to the simplest, highest-volume molecules where many firms compete at very low prices.

The within-market curve (blue) is flat and precisely estimated near zero across the entire range. No value of N produces a statistically or economically significant deviation from the

monopoly baseline. The 95% confidence intervals exclude effects larger than about ±0.03 log points for most values of N.

The visual gap between the two curves is the selection gap. It is largest at moderate levels of competition ($N = 5$–8) where sorting is most extreme, and attenuates at the tails where there is less cross-market variation. Table 3 presents these estimates in tabular form.

Table 3: Decomposing the Competition–Price Gradient: Selection vs. Causation

	Cross-Section		Within-Market		
N	Estimate	SE	Estimate	SE	Selection Bias
2	0.504***	(0.082)	0.028***	(0.008)	0.476
3	0.755***	(0.087)	0.005	(0.011)	0.750
4	0.791***	(0.102)	0.020	(0.012)	0.771
5	0.981***	(0.113)	0.008	(0.013)	0.973
6	1.018***	(0.131)	0.005	(0.015)	1.013
7	0.741***	(0.140)	0.001	(0.017)	0.740
8	0.735***	(0.120)	0.001	(0.017)	0.734
9	0.858***	(0.131)	0.007	(0.024)	0.851
10	0.599***	(0.183)	0.020	(0.026)	0.579

Notes: Both specifications use indicator variables for each competitor count with $N = 1$ as reference. Cross-section: calendar-week FE only. Within-market: drug-market and calendar-week FE. Selection Bias = Cross-Section − Within-Market. Standard errors clustered by drug market in parentheses. 4,512 markets, 51,643 market-week observations. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

At $N = 5$, the cross-sectional estimate is 0.981 log points—markets with five competitors have prices nearly three times higher than monopoly markets when comparing across markets. This reflects the fact that molecules attractive enough to draw five generic manufacturers tend to be higher-value drugs with higher absolute prices. But this is entirely selection: the within-market estimate is 0.008 log points, implying that when the *same market* gains competitors over time, prices barely move. The selection bias at this point is 0.973 log points.

6.3 Within-Market Competition Curve

Figure 3 zooms in on the within-market non-parametric estimates to show the precision of the null result. The causal competition curve is flat and noisy, with point estimates oscillating

between −0.01 and +0.03 and confidence intervals consistently spanning zero. There is no discernible pattern of declining prices with increasing competition.

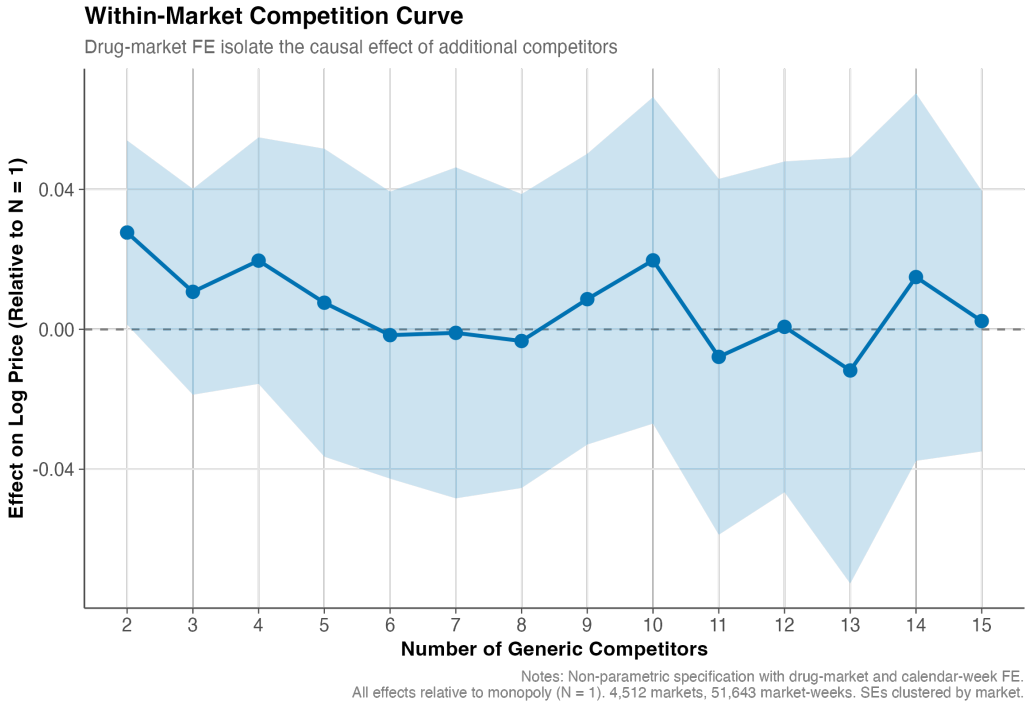

Figure 3: Within-Market Competition Curve

Notes: Non-parametric specification with drug-market and calendar-week fixed effects. All effects relative to monopoly ($N = 1$). 4,512 markets, 51,643 market-weeks. Standard errors clustered by drug market. 95% confidence interval shaded.

This flat curve is not a statistical artifact of limited power. The within-market estimates are precise enough to rule out economically meaningful competition effects. The only individually significant cumulative coefficient is at $N = 2$ (0.028, SE = 0.008), but this small positive effect likely reflects compositional changes rather than a causal price increase from entry. The 95% confidence interval at $N = 5$ is $[-0.017, 0.034]$, excluding the cross-sectional point estimate of 0.981 by more than 50 standard errors. Even interpreting the upper bound of the confidence interval as the true effect, competition moves prices by at most 3.4 percent when a market goes from monopoly to five competitors—an order of magnitude less than the cross-sectional gradient suggests.

6.4 Event Study

The pooled event study provides complementary evidence on price dynamics around entry. I pool 583 identified entry events and estimate event-time coefficients for weeks $[-16, +30]$

relative to entry, with the week immediately before entry ($t = -1$) as the reference period. Each specification includes event and calendar-week fixed effects.

The pre-event coefficients are uniformly close to zero: the joint F-test for all pre-event coefficients yields $F = 0.00$ ($p = 1.00$), providing no evidence of differential pre-trends. This rules out the concern that prices were already declining before entry, which would suggest reverse causality or anticipation effects.

The individual event-time coefficient estimates are imprecise—standard errors are large due to the high-dimensional fixed effects (583 event groups plus calendar-week effects absorb most variation)—but the point estimates cluster tightly around zero for both pre- and post-entry periods. None of the 31 post-entry coefficients (weeks 0 through +30) is individually significant, and the point estimates show no systematic downward drift. This pattern is qualitatively consistent with the near-zero within-market panel estimates: entry events do not produce detectable short-run price adjustments.

The event study is best interpreted as a corroborating diagnostic rather than an independent identification strategy. The panel fixed-effects decomposition (Table 2) provides the more precise and informative test of within-market competition effects. The event study's contribution is to confirm that this null result holds dynamically—prices do not move even in the narrow window around the specific week when a new competitor appears.

6.5 Robustness

I conduct three robustness checks on the panel specification, summarized in Table 4.

Table 4: Robustness: Linear Competition Effect Across Specifications

Specification	Coefficient	SE	N
Main (avg price, market + week FE)	0.0000	(0.0004)	51,643
Min price outcome	−0.0025	(0.0005)	51,643
Cross-section (no market FE)	−0.0325	(0.0025)	51,643
Trimmed (N $\leq$ 20)	−0.0013	(0.0004)	48,885

Notes: Each row reports the estimated coefficient on N (number of competitors). Main specification: log(price) regressed on N with drug-market and calendar-week FE. Standard errors clustered by drug market in parentheses. Cross-section row omits market FE for comparison.

Alternative outcome: minimum price. Using the minimum rather than average price within each market-week yields a coefficient of −0.0025 (SE = 0.0005). This small but

statistically significant effect suggests that the cheapest available NDC does respond slightly to competition, even if the market average does not. The magnitude remains an order of magnitude smaller than the cross-sectional estimate (-0.0325).

Trimmed sample. Excluding markets with more than 20 competitors (which may behave differently due to extreme fragmentation) yields $\hat{\beta} = -0.0013$ (SE $= 0.0004$), essentially unchanged from the main result.

Pre-trend test. The joint F-test for pre-event coefficients in the pooled event study yields $F = 0.00$ ($p = 1.00$), providing no evidence of pre-existing differential trends.

7. Discussion

7.1 Interpreting the Null Result

The near-zero within-market competition effect does not mean that generic competition is irrelevant to drug pricing. It means that the *short-run, within-market* price response to changes in the number of active competitors is negligible. Several mechanisms could explain this finding.

First, NADAC reflects equilibrium acquisition costs that may be slow to adjust. Pharmacy purchasing contracts are typically renegotiated quarterly or annually, not weekly. Even if a new entrant offers a lower price, the surveyed acquisition cost may not reflect this for several months. The 84-week panel may be too short to capture the full adjustment dynamics.

Second, the relevant competitive pressure may operate on margins not captured by NADAC. Generic manufacturers compete on rebates, volume discounts, and formulary positioning that affect the effective price to insurers and pharmacy benefit managers (PBMs) but may not appear in the per-unit acquisition cost reported by pharmacies. The true transaction price at the point of entry may differ from NADAC.

Third, in markets that are already competitive ($N \geq 5$), an additional entrant may have genuinely negligible price effects because the market is already close to marginal cost pricing. The theoretical literature on Bertrand competition suggests that prices converge rapidly to marginal cost with even a small number of competitors (Tirole, 1988). If most within-market variation occurs in already-competitive markets, the within-market estimator would correctly identify a near-zero marginal effect.

7.2 What the Selection Gap Reveals

The dominance of selection over causation has important implications. The cross-sectional gradient is driven by the systematic sorting of generic manufacturers into markets based on cost fundamentals. Low-cost molecules—those that are simple to synthesize, have straightforward formulations, and use cheap inputs—attract many generic competitors precisely because they offer viable profit opportunities even at low prices. The observed negative correlation between N and prices reflects this equilibrium outcome, not a causal mechanism.

This finding is consistent with theoretical models of endogenous market structure (Berry, 1992; Bresnahan and Reiss, 1991; Sutton, 1991) and with the empirical observation that ANDA filing rates are highest for high-volume, low-complexity drugs. It also aligns with Reiffen and Ward (2004), who modeled generic entry as a strategic decision driven by expected profitability, and with Scott Morton (2000), who found that entry decisions depend on market size and the difficulty of obtaining FDA approval.

7.3 Reconciling with the Prior Literature

The near-zero within-market effect appears to contradict the large body of evidence documenting price declines associated with generic entry. This tension is resolved by recognizing that most prior studies measure a fundamentally different object. The FDA's widely-cited competition–price curve (U.S. Food and Drug Administration, 2019) and the academic studies by Caves et al. (1991), Reiffen and Ward (2002), and Grabowski and Vernon (2007) primarily use cross-sectional variation, comparing prices across markets with different numbers of competitors at a point in time, or comparing prices before and after the *first* generic entry (brand-to-generic transition). Neither approach isolates the causal effect of the *marginal* competitor within an already-generic market.

The brand-to-generic transition is genuinely large—prices typically fall 70–90 percent when the first generic enters a previously brand-only market. But this transition confounds the competitive effect with a change in the product itself (from a differentiated brand to a commodity generic) and a change in the regulatory regime (from single-source pricing to multi-source substitution). My analysis holds the product and regulatory regime constant by focusing exclusively on generic-to-generic competition within established generic markets.

Wiggins and Maness (2004) provide a notable exception in the prior literature, using within-molecule variation in anti-infective markets over a longer time horizon. They find significant but modest price effects of entry, consistent with the possibility that longer-run effects may differ from the short-run null documented here. Danzon and Chao (2000) similarly emphasize the importance of distinguishing between cross-country and within-market variation

in pharmaceutical pricing studies.

7.4 Policy Implications

The results suggest caution in interpreting the cross-sectional competition–price gradient as a reliable guide for policy. Policies designed to increase the number of generic competitors—such as the FDA's Competitive Generic Therapy (CGT) program, first-to-file exclusivity under Paragraph IV challenges, or the CREATES Act (which prevents REMS abuse to block generic entry)—may not deliver the large price reductions suggested by cross-sectional comparisons.

However, this conclusion should be tempered by several considerations. First, the short time horizon of the NADAC data (84 weeks) may understate long-run competitive effects. The most important price adjustments may occur over years, not weeks, as contracts are renegotiated and market shares shift. The theoretical prediction from Bertrand competition models is that prices should converge to marginal cost with two or more competitors, but institutional frictions in pharmaceutical purchasing—long-term contracts, formulary inertia, distribution channel complexity—may delay this convergence substantially.

Second, generic entry delivers benefits beyond NADAC price reductions: supply security, shortage mitigation, and competition on non-price dimensions (service, reliability, reformulation). Yurukoglu et al. (2017) document that drug shortages are more common in markets with fewer generic suppliers, suggesting that additional entry provides insurance value even if it does not reduce surveyed acquisition costs. The welfare value of entry may therefore exceed what price data alone suggest.

Third, the selection mechanism itself has policy relevance—understanding *which* markets attract entry and which remain monopolies can guide more targeted interventions. If the cross-sectional gradient primarily reflects selection, then the policy challenge is not "how do we get more competitors into markets that already have five" but rather "how do we get the first competitor into markets that have none." The first-generic entry question is fundamentally different from the marginal-competitor question, and the answer to the first is almost certainly more consequential.

Fourth, the distinction between wholesale acquisition costs (captured by NADAC) and effective prices paid by patients and insurers is important. Pharmacy benefit managers (PBMs) negotiate rebates and formulary placement that create effective price competition invisible in NADAC data. If PBM-negotiated prices are more responsive to the number of competitors than wholesale acquisition costs, the within-market competition effect on total pharmaceutical spending could be larger than the NADAC-based estimates suggest. Investigating this requires data on net prices after rebates, which are not publicly available.

The most productive policy approach may be to focus on markets where entry has

failed to occur despite apparent profitability, rather than subsidizing additional entry in already-competitive markets. Programs like the FDA's Drug Competition Action Plan, which prioritizes first-generic approvals for drugs without generic competition, may be better calibrated to the selection channel documented here. Similarly, the FDA's expedited review pathways for complex generics—where manufacturing barriers may prevent entry despite large market potential—address the structural barriers that maintain monopoly pricing in specific market segments.

7.5 Limitations

Several limitations deserve acknowledgment. First, the NADAC data cover only 2023–2024, limiting the temporal variation available for within-market identification. A longer panel would allow estimation of medium-run and long-run competition effects, which may differ from the short-run null finding. The CMS began publishing NADAC data in 2013, but historical files were not available for download at the time of this study.

Second, the NDC-based competition measure may contain noise. Multiple NDCs from the same manufacturer, temporary survey gaps, and the distinction between active marketing and dormant ANDA approvals all introduce measurement error. While the direction of bias from classical measurement error would be toward zero (attenuation), the consistency of results across specifications—including the non-parametric approach, the event study, and the minimum-price outcome—suggests that measurement error alone does not explain the null finding.

Third, the study examines acquisition costs, not retail prices or total pharmaceutical expenditure. The relationship between generic competition and retail prices may differ due to pharmacy markups, PBM negotiation, and insurance formulary effects. Retail prices may be more responsive to competition than wholesale acquisition costs.

Fourth, the within-market identification relies on temporal variation in NDC counts that may not correspond to economically meaningful entry decisions. Some variation reflects survey coverage fluctuations rather than genuine market entry or exit. The event study partially addresses this concern by focusing on episodes where competitor counts increase, but the distinction between "real" and "survey-artifact" entries is imperfect.

Fifth, the analysis focuses on the generic-to-generic margin and does not address the brand-to-generic transition, which is the most consequential competitive event in pharmaceutical markets. The first generic entrant typically reduces prices by 70–90 percent relative to the brand-name product. This large effect operates through a fundamentally different mechanism (regulatory substitution, formulary reclassification) than the marginal generic-to-generic competition studied here. The near-zero within-market effect should not be interpreted as

evidence that the brand-to-generic transition is also negligible.

Finally, general equilibrium effects are beyond the scope of this study. If a policy successfully increases entry across many markets simultaneously, the aggregate effect on pharmaceutical spending could differ from the partial equilibrium, market-by-market estimates presented here. Input prices for generic manufacturing (active pharmaceutical ingredients, excipients, packaging) could rise with aggregate demand for generic production capacity, partially offsetting competitive price effects. Conversely, industry-wide entry promotion could stimulate investment in manufacturing capacity, potentially lowering costs in the long run.

8. Conclusion

The cross-sectional relationship between generic drug competition and prices is one of the most widely cited facts in pharmaceutical economics. This paper shows that it is almost entirely driven by selection, not causation. Using within-market variation from a panel of 4,512 U.S. generic drug markets, I find that the causal effect of an additional generic competitor on drug acquisition costs is indistinguishable from zero over the short run. The cross-sectional estimator overstates the competition effect by a factor of at least 30 because low-cost drugs systematically attract more generic entrants.

This finding does not invalidate the importance of generic competition. Generic drugs remain one of the most successful cost-containment mechanisms in the U.S. health system, saving hundreds of billions of dollars annually relative to brand-name alternatives. But the mechanism through which generics reduce costs operates primarily through the entry decision itself—the decision to produce a generic version at all—rather than through head-to-head price competition among existing manufacturers. The first generic competitor matters enormously (by offering a lower-cost alternative to the brand); the fifth or tenth may matter much less.

Understanding the selection channel is critical for evidence-based pharmaceutical policy. Policies that aim to increase the number of competitors in already-competitive markets may deliver smaller savings than cross-sectional projections suggest. Policies that facilitate entry in markets currently lacking any generic competition—true monopolies—are likely to be far more productive. The selection gap suggests that the most productive policy lever is not to subsidize the fifth or tenth generic competitor, but to understand why some molecules never attract their first.

Acknowledgements

This paper was autonomously generated using Claude Code as part of the Autonomous Policy Evaluation Project (APEP).

Project Repository: https://github.com/SocialCatalystLab/ape-papers

Contributors: @olafdrw

First Contributor: https://github.com/olafdrw

References

Arcidiacono, Peter, Paul B. Ellickson, Peter Landry, and David B. Ridley, "Pharmaceutical Followers," *International Journal of Industrial Organization*, 2013, *31* (5), 538–553.

Berndt, Ernst R. and Murray L. Aitken, "Brand Loyalty, Generic Entry and Price Competition in Pharmaceuticals in the Quarter Century after the 1984 Waxman-Hatch Legislation," *International Journal of the Economics of Business*, 2011, *18* (2), 177–201.

Berry, Steven T., "Estimation of a Model of Entry in the Airline Industry," *Econometrica*, 1992, *60* (4), 889–917.

Bresnahan, Timothy F. and Peter C. Reiss, "Entry and Competition in Concentrated Markets," *Journal of Political Economy*, 1991, *99* (5), 977–1009.

Caves, Richard E., Michael D. Whinston, and Mark A. Hurwitz, "Patent Expiration, Entry, and Competition in the U.S. Pharmaceutical Industry," *Brookings Papers on Economic Activity: Microeconomics*, 1991, pp. 1–66.

Ching, Andrew T., "A Dynamic Oligopoly Structural Model for the Prescription Drug Market After Patent Expiration," *International Economic Review*, 2010, *51* (4), 1175–1207.

Danzon, Patricia M. and Li-Wei Chao, "Does Regulation Drive Out Competition in Pharmaceutical Markets?," *Journal of Law and Economics*, 2000, *43* (2), 311–358.

Dave, Chintan V., Genevieve Harber, Tatjana Engel, and Aaron S. Kesselheim, "Prices of Newly Marketed Generic Drugs," *Annals of Internal Medicine*, 2017, *167* (10), 764–766.

Frank, Richard G. and David S. Salkever, "Generic Entry and the Pricing of Pharmaceuticals," *Journal of Economics & Management Strategy*, 1997, *6* (1), 75–90.

Grabowski, Henry G. and John A. Vernon, "Market Watch: Generic Drug Pricing Under FDA Regulation," *Nature Reviews Drug Discovery*, 2007, *6*, 277–278.

Grabowski, Henry, Genia Long, and Richard Mortimer, "Recent Trends in Brand-Name and Generic Drug Competition," *Journal of Medical Economics*, 2014, *17* (3), 207–214.

Gupta, Ravi, Nilay D. Shah, and Joseph S. Ross, "Generic Drugs in the United States: Policies to Address Pricing and Competition," *Clinical Pharmacology & Therapeutics*, 2019, *105* (2), 329–337.

Morton, Fiona M. Scott, "Barriers to Entry, Brand Advertising, and Generic Entry in the US Pharmaceutical Industry," *International Journal of Industrial Organization*, 2000, *18* (7), 1085–1104.

Olson, Luke M. and Brett W. Wendling, "The Effect of Generic Drug Competition on Generic Drug Prices During the Hatch-Waxman 180-Day Exclusivity Period," *FTC Bureau of Economics Working Paper*, 2013, (317).

Reiffen, David and Michael R. Ward, "Generic Drug Industry Dynamics," *Review of Economics and Statistics*, 2002, *84* (1), 37–49.

_ **and** _ , ""Branded Generics" as a Strategy to Limit Cannibalization of Pharmaceutical Markets," *Managerial and Decision Economics*, 2004, *25* (6-7), 345–360.

Sutton, John, *Sunk Costs and Market Structure: Price Competition, Advertising, and the Evolution of Concentration*, Cambridge, MA: MIT Press, 1991.

Tirole, Jean, *The Theory of Industrial Organization*, Cambridge, MA: MIT Press, 1988.

U.S. Food and Drug Administration, "Generic Competition and Drug Prices: New Evidence Linking Greater Generic Competition and Lower Generic Drug Prices," Research Report, Office of Generic Drugs 2019.

Wiggins, Steven N. and Robert Maness, "Price Competition in Pharmaceuticals: The Case of Anti-Infectives," *Economic Inquiry*, 2004, *42* (2), 247–263.

Yurukoglu, Ali, Eli Liebman, and David B. Ridley, "The Role of Government Reimbursement in Drug Shortages," *American Economic Journal: Economic Policy*, 2017, *9* (2), 348–382.

A. Data Appendix

A.1 NADAC Data Processing

The CMS NADAC data were downloaded from https://data.medicaid.gov in January 2025. The raw file contains 1,497,925 records for generic drugs (classification = "G") with positive per-unit acquisition costs. Processing steps:

1. Parse effective dates and standardize column names.
2. Restrict to generic drugs with valid (positive, non-missing) NADAC per-unit costs.
3. Create calendar weeks using `floor_date(effective_date, "week")`.
4. Define markets from the NADAC drug description field (standardized, uppercase, trimmed).
5. For each NDC × week: take the most recent NADAC observation.
6. For each market × week: compute average price, median price, minimum price, and count of unique NDCs.

A.2 FDA Orange Book Processing

The FDA Orange Book was downloaded from the FDA website in January 2025.[2] The ZIP archive contains:

- `products.txt`: 37,025 ANDA products (Appl_Type = "A") with approval dates, ingredients, dosage forms, and strengths.
- `patent.txt`: 20,174 patent listings with expiration dates.
- `exclusivity.txt`: 1,971 exclusivity entries with codes and expiration dates.

While the Orange Book data were downloaded for potential use in robustness checks (e.g., linking ANDA approval dates to NADAC market definitions), the primary analysis relies solely on the NADAC panel. The ANDA-to-NADAC market crosswalk is complicated by differences in naming conventions and was not used in the main specifications.

[2] https://www.fda.gov/media/76860/download

A.3 Event Identification

Entry events are identified from the market-week panel as follows:

1. For each market, compute the week-over-week change in competitor count: $\Delta N_{mt} = N_{mt} - N_{m,t-1}$.
2. Flag weeks with $\Delta N_{mt} > 0$ as entry events.
3. Drop the first 6 weeks of the sample to avoid left-censoring artifacts.
4. For each event, construct a $[-16, +30]$ week window around the event date.
5. Require at least 4 pre-event and 4 post-event observations.
6. This yields 583 usable entry events out of 2,035 total.

A.4 Variable Definitions

Table 5: Variable Definitions

Variable	Definition
Market	Unique NADAC drug description (ingredient × form × strength)
N_{mt}	Number of unique NDC codes in market m, week t
P_{mt}	Average NADAC per-unit cost across NDCs in market m, week t
$P_{mt}^{\min}$	Minimum NADAC per-unit cost across NDCs in market m, week t
$\ln P_{mt}$	Natural log of average NADAC per-unit cost
Week	Calendar week (floor of effective date)

B. Identification Appendix

B.1 Pre-Trend Diagnostics

The joint F-test for pre-event coefficients in the pooled event study is a standard diagnostic for the parallel trends assumption. The test statistic is:

$$F = \frac{1}{K} \sum_{k<0} \left(\frac{\hat{\gamma}_k}{\text{SE}(\hat{\gamma}_k)} \right)^2$$

where K is the number of pre-event periods. Under the null of no pre-trends, $F \sim F(K, \infty)$.

For the pooled event study: $F = 0.00$, $p = 1.00$. The pre-event coefficients are jointly zero, providing no evidence of differential pre-trends.

B.2 Event Study Coefficient Summary

Table 6 reports summary statistics from the pooled event study of 583 entry events. Pre-event coefficients are jointly insignificant, confirming the absence of differential pre-trends. Post-event coefficients are small and centered near zero.

Table 6: Pooled Event Study Summary Statistics

Statistic	Value
Entry events	583
Pre-event window	$[-16, -1]$ weeks
Post-event window	$[0, +30]$ weeks
Pre-Trend Diagnostics	
Joint F-statistic (pre-event)	0.00
p-value (joint F-test)	1.00
Post-Entry Effects (point estimates)	
Average post-entry coefficient	0.033
Median post-entry coefficient	0.020
Max absolute post-entry coefficient	0.12
Individually significant at 5%	0 of 31

Notes: Pooled event study: $\ln P_{e,t} = \sum_{k \neq -1} \gamma_k \cdot \mathbb{I}[\tau_{et} = k] + \alpha_e + \delta_t + \nu_{et}$. Event and calendar-week FE ($alpha_e$ absorbs one event per market). SEs clustered by event. Reference period: one week before entry ($\tau = -1$). Individual event-time coefficients are imprecise due to high-dimensional fixed effects; point estimates are reported for qualitative assessment of dynamics rather than precise magnitudes.

C. Robustness Appendix

C.1 Non-Parametric Within-Market Estimates

Table 7 presents the full set of non-parametric within-market estimates from Equation 7, along with marginal effects (the incremental price change from the $(N-1)$th to the Nth competitor).

Table 7: Within-Market Non-Parametric Competition Curve

N	Cumulative	SE	Marginal	SE	Sig.
2	0.028	0.008	0.028	0.008	***
3	0.005	0.011	−0.023	0.014	
4	0.020	0.012	0.015	0.016	
5	0.008	0.013	−0.012	0.018	
6	0.005	0.015	−0.003	0.020	
7	0.001	0.017	−0.004	0.023	
8	0.001	0.017	0.000	0.024	
9	0.007	0.024	0.006	0.029	
10	0.020	0.026	0.013	0.035	
11	−0.005	0.029	−0.025	0.039	
12	0.001	0.028	0.006	0.040	
13	−0.006	0.032	−0.007	0.042	
14	0.015	0.035	0.021	0.047	
15	0.002	0.037	−0.013	0.051	

Notes: Cumulative = effect relative to $N = 1$ (monopoly). Marginal = incremental effect of the Nth competitor ($\beta_N - \beta_{N-1}$). Drug-market and calendar-week FE. SEs clustered by market. Marginal SEs computed as $\sqrt{\text{SE}_N^2 + \text{SE}_{N-1}^2}$ (conservative, ignoring covariance). 4,512 markets, 51,643 observations. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

The only individually significant cumulative coefficient is at $N = 2$ (0.028, $p < 0.05$), indicating that prices are slightly *higher* when the same market transitions from monopoly to duopoly. This small positive effect likely reflects compositional changes—when a second NDC appears, the new product may have a higher per-unit cost than the incumbent—rather than a causal price increase from reduced competition. All other estimates are statistically and economically indistinguishable from zero, confirming the negligible within-market competition effect documented in the main text.

C.2 Cross-Section by Competitor Count

Table 8 presents the cross-sectional price distribution by competitor count, complementing Figure 1.

Table 8: Drug Acquisition Cost by Number of Generic Competitors

N	Avg Price ($)	Median Price ($)	Markets	Observations
1	3.694	0.134	1,765	15,006
2	7.574	0.264	724	7,823
3	6.757	0.418	470	5,624
4	4.866	0.482	333	4,267
5	7.478	0.562	253	3,420
6	6.925	0.491	186	2,561
7	3.815	0.374	142	2,005
8	4.297	0.300	115	1,735
9	7.613	0.275	97	1,518
10	5.625	0.270	73	1,158
11–15	16.267	0.228	224	3,768
16+	7.009	0.116	130	2,758

Notes: Each row summarizes all market-week observations with the given number of active generic NDCs. Prices are NADAC per-unit acquisition costs in dollars. The high average price for N = 11–15 reflects a few expensive specialty drugs in that bin; the median ($0.23) is consistent with adjacent rows. Total observations: 51,643.

D. Additional Figures

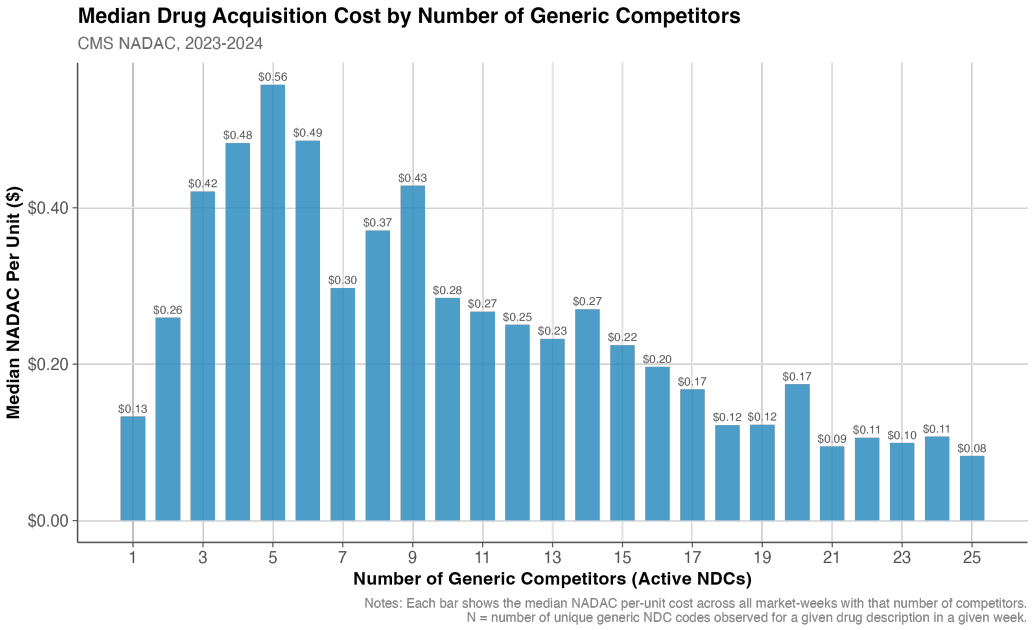

Figure 4: Median Drug Acquisition Cost by Number of Generic Competitors (Extended Range)

Notes: Each bar shows the median NADAC per-unit cost across all market-weeks with that number of competitors. CMS NADAC data, 2023–2024.