Sherlocking: The Effects of Platform-Owner Entry on the Competitive Behavior of Third-Party Firms*

Benjamin T. Leyden†

February 22, 2026

I study how third-party firms respond when a platform owner enters its own marketplace, analyzing Apple's entries into App Store submarkets from 2016–2021. Using text embeddings to define markets and a staggered difference-in-differences design, I find that Apple's entry deters new competitors and shifts incumbents' monetization strategies, but effects vary widely: many markets show no meaningful response, while others move in opposing directions across a host of monetization and quality outcomes. Responses depend on how Apple enters and apps' competitive proximity to Apple. This heterogeneity suggests targeted oversight rather than categorical restrictions on platform-owner entry.

JEL: L13, L86, L40

*I am grateful to Qihong Liu, Chiara Farronato, Devesh Raval, Mike Walker and participants at the 20th Annual International Industrial Organization Conference, the NYU Economics of Strategy Workshop, the NUS Economics of Platforms Workshop, ASSA 2023, the US DOJ Economic Analysis Group, the Cornell Strategy and Economics of Digital Markets Conference, HKUST workshop, the 2025 CESifo Economics of Digitization Conference, the University of Sydney, APIOC 2025, and the Digital Competition Conference 2026 for helpful comments. I thank Yujie Feng for excellent research assistance, and I acknowledge the Cornell Center for Social Sciences for financial support.

†Cornell University; CESifo; leyden@cornell.edu

1 Introduction

Platform-owner entry into digital marketplaces has become a central issue in antitrust policy and the regulation of digital markets (De Chant, 2022). The practice—where marketplace operators compete with third-party sellers on their own platforms—has prompted significant regulatory debate worldwide. The European Union's Digital Markets Act, enacted in 2022, explicitly restricts "gatekeepers" from engaging in "any form of differentiated or preferential treatment... in favour of products or services it offers itself" (EUR-Lex, 2022). In the United States, the previously proposed American Innovation and Choice Online Act would have similarly banned platforms from using "nonpublic data... to offer... products or services... that compete... with products or services offered by business users" (Congress.gov, 2022). In some cases, politicians have called for an outright ban on the practice (Warren, 2019). These blanket restrictions reflect growing concerns that platform owners exploit their unique position to discriminate against competitors, leveraging privileged information and control over marketplace infrastructure to advantage their own offerings (Crémer, de Montjoye, and Schweitzer, 2019; Scott Morton et al., 2019).

Yet the economic effects of platform-owner entry remain theoretically ambiguous (Anderson and Bedre Defolie, 2024; Hagiu, Teh, and Wright, 2022). While entry introduces a typically well-resourced competitor that may engage in self-preferencing or restrict access to essential features, it can also generate positive spillovers through technology diffusion, market expansion, and complementary infrastructure investments. The debate over appropriate regulatory responses—ranging from outright bans to conduct remedies to laissez-faire approaches—hinges critically on understanding how large these effects are, how much they vary across markets and competitive contexts, and what mechanisms drive them.

To help inform this debate, I study how Apple's release of competing apps and software functionality affects third-party developers. This "Sherlocking" phenomenon—named after Apple's incorporation of the third-party search app *Watson*'s functionality into its own *Sherlock* product—is a first-order concern for developers, who routinely track Apple's entry into App Store submarkets (see Figure 1). Developers face substantial uncertainty about whether and when Apple might enter their product space, complicating long-term investment decisions. For those whose products are targeted, the stakes are high: Apple's entry has the potential to render years of development effort obsolete overnight.

In the context of mobile software platforms, platform-owner entry takes two distinct forms that have received insufficient attention in both academic literature and policy discussions. Apple's entry into submarkets within its App Store marketplace occurs through either *standalone entry*—developing and releasing a new separately installable application—or *integrated entry*—incorporating competing functionality directly into the operating system. Standalone entry involves Apple creating a new app that users can discover and choose to download from the App Store, competing alongside third-party offerings. Examples include Measure (an augmented reality measurement tool), Translate (language translation), and Magnifier (visual assistance). These apps appear in search results, have product pages with descriptions and ratings, and can be uninstalled

Figure 1: Developer Reactions to "Sherlocking"

Note: Tweets from developers reacting to Apple's WWDC 2022 announcements of new iOS features that compete with existing third-party apps.

by users—though some may come pre-installed by default. The competitive dynamic is relatively transparent: consumers actively choose whether to adopt Apple's offering over alternatives.

Integrated entry involves Apple integrating functionality directly into iOS that automatically becomes available to all users. Examples include health tracking features integrated into the Health app, Apple Music enhancements, Notes app improvements, and new camera features. This functionality cannot be separately uninstalled and often comes bundled with new APIs that third-party developers can leverage, representing a more fundamental alteration of the platform's capabilities. Users receive these features automatically, without making an active adoption decision.

This distinction matters for competitive dynamics. Integrated entry creates unavoidable competition—third-party apps must compete with functionality that every user already possesses. It also has the potential to generate stronger technology spillovers, as OS-level integration often introduces new hardware capabilities or software frameworks that benefit the broader developer ecosystem.

This paper provides comprehensive evidence on platform-owner entry effects by studying 23 instances of Apple entering submarkets in its App Store between March 2016 and September 2021. The analysis encompasses 14 standalone app releases and 9 integrated OS integrations, enabling systematic comparison across entry types and market characteristics.

I make three key contributions to the literature on platform competition and regulation. First, by examining many entry events, I formally characterize cross-market heterogeneity in competitive responses. Applying statistical heterogeneity diagnostics to market-specific treatment effects reveals that markets frequently disagree even on the direction of response. Second, I develop a novel empirical approach using continuous distance measures derived from text embeddings to capture within-market heterogeneity based on competitive proximity. This methodology recognizes that not all apps in an affected market are equally "treated" by platform entry. Third, I show that entry implementation matters: integrated OS features generate competitive effects on key pricing margins substantially larger than standalone app releases.

My empirical strategy leverages a staggered difference-in-differences design, using comparable Google Play Store apps as controls for App Store apps affected by Apple's entry, with within-market cross-platform comparisons providing the primary source of identification. Market definitions rely on semantic similarity of app descriptions, with apps exceeding a similarity threshold considered part of the competitive market around each entrant.

Pooled estimates show that Apple's entry deters new competitors by 23%, shifts incumbent pricing from free to paid, and increases rating count growth, a proxy for user demand and engagement. These averages, however, mask striking heterogeneity that challenges both pro-platform entry and anti-platform entry narratives. Formal quantification confirms that this variation is genuine: high I^2 values across outcomes indicate that pooled average effects poorly represent any individual market's experience. Perhaps most surprisingly, a substantial proportion of markets experience no statistically significant effects from Apple's entry across many of the outcomes considered in this analysis. This prevalence of null effects, particularly pronounced for quality metrics, contradicts claims of either universal harm or universal benefit from platform competition, though effects can be concentrated among close competitors as the distance analysis reveals.

When effects do occur, they vary across markets and space. Entry type proves crucial. Integrated OS features generate systematically larger effects than standalone app releases. Price increases are substantially larger for integrated entry ($0.65) than for standalone entry ($0.04), while the proportion of free apps declines by 10.3 percentage points for integrated entry but shows essentially no change for standalone entry. Moreover, apps close to Apple's offering in semantic space experience substantial, often negative impacts, while more distant competitors remain largely unaffected. This creates a limited "competitive radius" around each entrant which can vary in size by market and outcome.

These patterns of heterogeneity—across markets, between entry types, and within markets by distance—suggest that platform-owner entry effects depend critically on contextual factors that current regulatory proposals largely ignore. The evidence strongly supports targeted, context-specific oversight rather than the blanket restrictions.

This paper contributes to several strands of literature on platform competition and digital markets. Most directly, it extends empirical work on platform-owner entry effects. Wen and Zhu (2019) study Google's entry into Android submarkets, finding evidence of reduced innovation and increased prices among affected developers. Zhu and Liu (2018) document that Amazon targets successful product categories and that its entry increases demand but discourages third-party investment. Foerderer et al. (2018) find that Google's entry into photo management increased both demand and third-party innovation. My analysis expands this literature by examining a larger set of entry events, enabling systematic analysis of heterogeneous effects. While I focus on competitive effects, complementary work by Halckenhaeusser et al. (2025) examines what drives platform-owner entry decisions. They find that Apple's iOS entries target platform categories characterized by low user satisfaction, limited innovation, and high market concentration.

A related literature examines specific mechanisms of platform advantage. Studies of self-preferencing document that platforms favor their own products in search and recommendations, with mixed welfare implications (Teng, 2025; Lee and Musolff, 2025; Lam, 2023; Raval, 2023). Work on "insider imitation" explores how platforms leverage proprietary data to select entry markets (Madsen and Vellodi, 2025; Hagiu, Teh, and Wright, 2022). While I do not isolate specific mechanisms, my reduced-form estimates capture the net effect of all channels through which platform entry affects competition.

The paper also relates to the "platforms as regulators" literature, which recognizes that platform owners shape competitive dynamics through rules, design choices, and technical standards (Boudreau and Hagiu, 2011). Recent work shows how platform policies on privacy (Kesler, 2023; Li and Tsai, 2025; Johnson et al., 2025; Cheyre et al., 2025), ratings systems (Jabr et al., 2020; Leyden, 2025), and product categorization (Ershov, 2024) affect within-platform competition. My findings suggest that entry decisions represent another regulatory lever, with effects varying by how entry is implemented.

Methodologically, the paper advances techniques for defining and measuring competition in digital markets. Building on Hoberg and Phillips (2016) and Leyden (2023), I use transformer-based language models to construct product spaces from text descriptions, enabling more nuanced market definitions than traditional classification systems. The continuous distance measure allows for gradient effects within markets, recognizing that competitive pressure varies with product differentiation.

The remainder of the paper proceeds as follows. Section 2 describes the data sources, market definition methodology, and sample construction. Section 3 presents the empirical analysis, beginning with pooled average treatment effects before formally quantifying cross-market heterogeneity using meta-analytic diagnostics, examining the role of entry type, and analyzing how competitive proximity shapes within-market responses. Section 4 concludes with a discussion of policy implications and directions for future research.

2 Data

I study platform-owner entry effects using the complete population of apps in Apple's App Store and Google's Play Store over a five-year period. I use similar apps on the Play Store as a control group for App Store apps affected by Apple's entry.[1] The Play Store represents the other major mobile marketplace with a comparable developer ecosystem and user base. The use of the Play Store as a control is particularly compelling because Google does not systematically enter the same markets at the same times as Apple, and both platforms experienced similar technological and demand shocks during the sample period. Below, I describe the data sources, Apple's entry events,

[1] Play Store in-app purchase data is only available starting September 2016. As a result, 5 markets (Schoolwork, Breathe, Health 10.0, Homekit 10.0, and Clock 10.0) are excluded from all IAP analyses because they entered before sufficient pre-period data was available.

the market definition procedure, and the construction of the two analysis samples.

2.1 Data Sources

I use a monthly panel dataset of all apps available in the US App Store and Google Play Store provided by the data analytics company AppMonsta (AppMonsta, 2016). The full sample runs from January 2015 to June 2022 and consists of 14.4 million unique apps across both platforms. For each app-month, I observe product characteristics including name, price, average rating, rating count, and whether the app was updated that month. The data also includes developer-written text descriptions that serve as apps' primary marketing content and provide a reliable indication of product functionality and purpose. Additionally, I use daily ranking lists for product sales and revenue to construct sample restrictions focused on actively competing apps.

Key variables for the analysis include monetization outcomes: the download price (Price), whether the app is free to download (Free), whether the app offers any in-app purchases (IAP), and the app's price conditional on it being non-zero (Price-Paid). I measure product quality using apps' average ratings, which are expressed on a 1-5 star scale (Average Rating), whether an app was updated by its developer in the given month (Update), and I use a measure of the arrival of new ratings as a proxy for consumer demand and engagement (Δ Log Rating Count).[2] Finally, I measure the frequency of entry and exit using log market-level counts of new app entries and exits.[3]

2.2 Apple's Entrants

I study 23 instances of Apple entering submarkets within its App Store ecosystem between March 2016 and September 2021. Apple implements entry through two distinct strategies. In *standalone entry*, Apple releases a new separately installable application that users can discover and download from the App Store; my sample includes 14 such events. These apps compete directly alongside third-party offerings with their own product pages, ratings, and the ability to be uninstalled, though some may come pre-installed on devices by default. Examples include Schoolwork (classroom management), Breathe (meditation), and Translate (language translation). In the second case, *integrated entry*, Apple integrates competing functionality directly into its operating system; my sample includes 9 such events. This functionality is automatically available to all users, cannot be uninstalled, and often introduces new APIs that third-party developers can leverage. Examples include Health app enhancements, Apple Music features, camera improvements, and Maps functionality updates.

Table 1 provides the complete list of all 23 entry events with their dates and types. This broad set of entry events, spanning diverse product categories and entry modes, enables systematic analysis of heterogeneous platform-owner entry effects.

[2] The rating count growth rate is calculated as $\Delta \ln(\text{Rating Count} + 1)$.

[3] Specifically, $\ln(\text{Entry or Exit Count} + 1)$.

Table 1: Apple Entry Events and App Counts by Platform

			Incumbent Sample		Market Dynamics Sample	
Market	Entry Date	Type	App Store	Play Store	App Store	Play Store
Schoolwork	2016-03	Standalone	97	30	246	65
Breathe	2016-09	Standalone	18	11	97	12
Clock 10.0	2016-09	Integrated	180	113	308	194
Health 10.0	2016-09	Integrated	19	9	23	15
Homekit 10.0	2016-09	Integrated	83	63	151	93
Camera 11.0	2017-09	Integrated	123	70	180	96
Development 11.0	2017-09	Integrated	202	11	134	19
Notes 11.0	2017-09	Integrated	238	94	285	90
Animoji	2017-11	Standalone	71	23	180	81
Health 11.3	2018-03	Integrated	116	67	169	58
Ibooks 11.4	2018-05	Integrated	50	8	24	12
Measure	2018-09	Standalone	61	19	81	26
Walkie Talkie	2018-09	Standalone	47	28	80	46
Apple Music 13.0	2019-09	Integrated	78	41	46	33
Ecg	2019-09	Standalone	39	16	25	11
Radio	2019-09	Standalone	3	4	8	4
Reality Composer	2019-09	Standalone	67	16	80	25
Apple Research	2019-11	Standalone	22	8	34	15
Sleep	2020-09	Standalone	114	60	81	47
Translate	2020-09	Standalone	327	152	201	101
Blood Oxygen	2020-10	Standalone	9	4	20	9
Find Items	2021-04	Standalone	29	19	22	14
Magnifier	2021-09	Standalone	41	18	9	19

Notes: App counts at Apple entry date based on $\theta = 0.6$ threshold. The incumbent sample counts apps present throughout the 18-month balanced panel that ranked in the top 200 for at least 10% of days. The market dynamics sample counts apps entering or exiting during the analysis window that ever ranked in the top 200 for at least 1 day. These are distinct populations: incumbent apps are established apps present before entry, while market dynamics apps are new entrants and exiters.

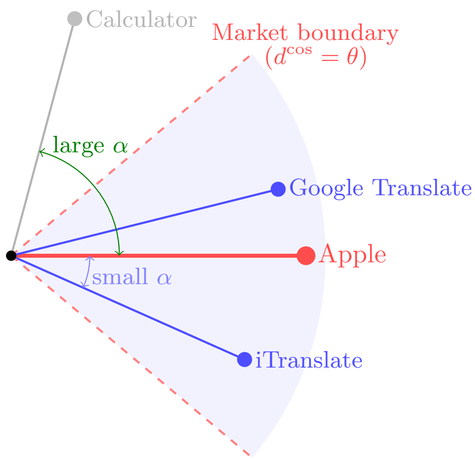
Figure 2: Semantic Similarity in Vector Space: An Illustration

Apple Translate:
"Translate lets you quickly and easily translate your voice and text. . . "

Google Translate:
"Translate between up to 249 languages. . . . "

iTranslate:
"Seamlessly translate text, websites, objects, or start voice-to-voice conversations in. . . "

Calculator App:
"Easily type out math with the basic and scientific calculators for quick solutions. . . "



Note: Left panel: Excerpts from product descriptions used as input to the SBERT model. Right panel: SBERT maps descriptions to vectors in a high-dimensional space where semantic similarity corresponds to geometric proximity (small angles between vectors). Translation apps cluster together despite using different terminology, while the calculator app is distant. The dashed lines represent market boundaries at threshold θ.

2.3 Market Definition

Given the millions of products observed over the sample period, identifying which third-party apps are potentially affected by platform-owner entry requires precise market definitions. Platform-provided categories (e.g., Games, Productivity) are too broad—the Productivity category alone encompasses password managers, shift scheduling apps, document editors, and many other distinct product types.

I define markets using semantic similarity of product descriptions, leveraging recent advances in natural language processing.[4] Specifically, I apply Sentence-BERT (SBERT), a transformer-based neural network model, to embed text descriptions into a 384-dimensional vector space in which semantically similar products are located closer together (Reimers and Gurevych, 2019). These embeddings are trained so that functionally similar apps—regardless of the specific terminology used by the developers—cluster in the embedding space. Markets are defined using the same SBERT process across platforms, ensuring that treated App Store apps and control Play Store apps operate in the same product spaces.

Let J^{AS} and J^{PS} be the sets of apps on the App Store and Play Store, respectively. Each app $j \in J^{\text{AS}} \cup J^{\text{PS}}$ is characterized by a description vector, $desc_j$, constructed using app j's first observed description. For standalone entries, the description vector $desc_a$ used in Equation (2) is simply the app's App Store product page description. For integrated entries, which lack natural product pages, I construct comparable descriptions using Apple's official iOS release notes as the source text. These

[4] There is a growing literature using text documents to characterize markets, following pioneering work by Hoberg and Phillips (2016). See also Hoberg and Phillips (2025), Leyden (2023), Han et al. (2024), and Gugler, Szücs, and Wohak (2024).

release notes document each feature's functionality at the time of the iOS version release. Because release note entries can be terse, I augment brief descriptions with additional context from Apple's developer documentation. I then generate App Store-style descriptions using OpenAI's GPT-5.1 with few-shot learning, providing actual descriptions of Apple's standalone apps as style examples to ensure the generated text matches the tone, structure, and vocabulary typical of Apple.[5] This approach produces description vectors that are semantically comparable to standalone entries and third-party apps, enabling consistent market definition across both entry types.

I measure the similarity between any app j and Apple entrant a using cosine similarity

$$d^{\cos}(desc_j, desc_a) = \frac{desc_j^\top desc_a}{\|desc_j\|_2 \, \|desc_a\|_2} \in [-1, 1], \tag{1}$$

where $\|\cdot\|_2$ denotes the Euclidean norm. Geometrically, cosine similarity is the cosine of the angle α between the two vectors, so vectors pointing in similar directions (small α) have similarity close to 1. Figure 2 illustrates this approach using language translation apps as an example: SBERT maps textually distinct but functionally similar descriptions to nearby vectors, while unrelated apps like calculators are placed farther away. Each app j in market a can thus be characterized by a continuous similarity measure to the entrant $d_j \in [\theta, 1]$, enabling analysis of how treatment effects vary with competitive proximity. This continuous treatment approach, presented in Section 3, represents a methodological advance over binary in/out market definitions.

Given a first-party entrant a with description vector $desc_a$, I define the market J_a^θ as all apps j satisfying:

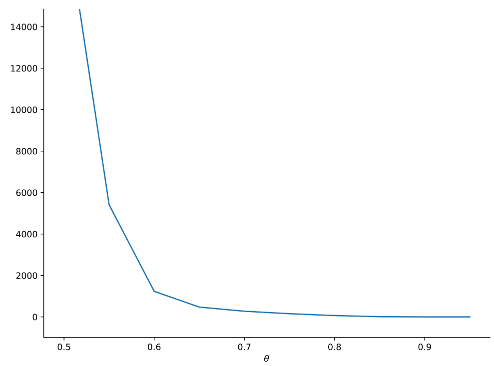
$$J_a^\theta = \{j \in J^{\text{AS}} \cup J^{\text{PS}} : d^{\cos}(\text{desc}_j, \text{desc}_a) \geq \theta\}. \tag{2}$$

The parameter θ governs market size: higher values impose stricter similarity requirements, yielding smaller, more tightly defined markets, while lower values produce larger markets that include more distant competitors. In practice, I set $\theta = 0.6$, which defines relatively tight, more focused competitive sets than lower alternative thresholds. This ensures the treated sample consists of apps that are meaningfully similar to the entrant, reducing concerns about including tangentially related products that might bias estimates toward zero. The threshold choice balances two considerations: setting θ too high yields markets that are too narrow and may miss relevant competitors, while setting it too low includes apps that are not meaningful substitutes.

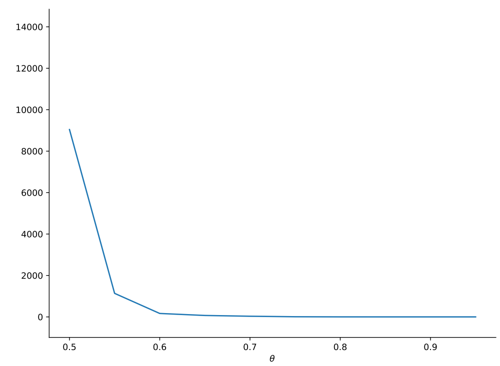
Market sizes vary substantially based on product space density. Figure 3 illustrates this heterogeneity for four representative markets. Each panel shows how many apps would be included in a market as the similarity threshold θ varies from 0.5 to 1.0 (prior to the ranking restriction discussed in the following section). At the baseline threshold of $\theta = 0.6$, market sizes vary substantially, from a dozen apps in specialized markets to several hundred in broader product categories.

[5] The generation process samples 20 Apple app descriptions as style examples for each integrated entry feature. The prompt instructs the model to produce 200–400 word descriptions that focus on user benefits and match the style of the examples. I use Apple's own app descriptions (e.g., Pages, Keynote, GarageBand) as style references because they represent the canonical Apple voice and are semantically neutral with respect to the markets being defined.

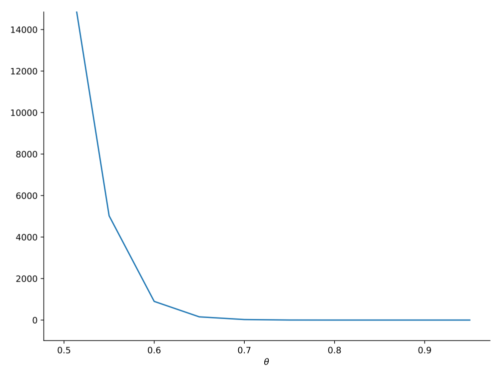
Figure 3: Market Size Heterogeneity: Cumulative App Counts by Similarity Threshold



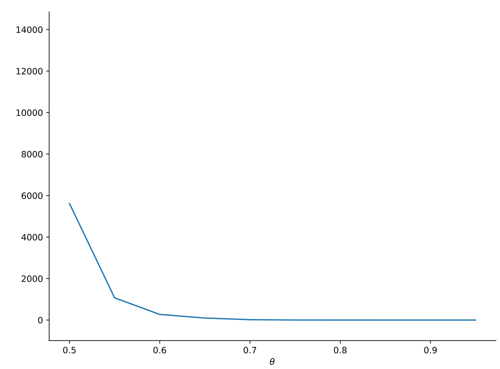
(a) Translate



(b) Measure



(c) Schoolwork



(d) ECG

Note: Each panel plots the cumulative number of apps with cosine similarity to the first-party entrant at or above threshold θ. The first-party entrant is located at $\theta = 1$. As θ increases moving rightward, fewer apps meet the threshold, yielding smaller, more focused markets.

The sharp increases visible at $\theta < 0.6$ demonstrate that loosening the threshold would dramatically expand markets to include many tangentially related apps, motivating the choice of a conservative threshold that focuses on close competitors.

2.4 Analysis Samples

One challenge in analyzing app markets is that many apps are hobby projects, abandoned apps, and experimental products that are not actively competing during the sample period. To address this, and study only competitively relevant products, I restrict my analysis sample based on daily product ranking lists, which rank apps on both platforms based on their number of downloads or the amount of revenue earned.[6]

Specifically, I construct two analysis samples. The *incumbent sample* restricts to apps that ranked 200^{th} or better for at least 10% of days they were available on the platform. This requirement selects for apps with sustained competitive relevance. For each Apple entrant, I construct an 18-month balanced panel centered around the first-party entry date. The resulting dataset contains 52,524 app-months of observations across 23 markets.

For the entry/exit and cohort composition analyses in Section 3.1, I use a less restrictive *market dynamics sample* that includes all apps that ever ranked in the top 200 for at least one day. The incumbent sample's stricter restriction selects on sustained post-entry performance, which is appropriate for the incumbent balanced panel but could bias extensive margin analysis: new entrants are mechanically less likely to satisfy a stringent ranking requirement, and conditioning on post-entry ranking risks selecting on outcomes affected by treatment. The market dynamics sample avoids this concern while still excluding apps with no evidence of market participation, yielding substantially larger samples for the extensive margin analyses: 3,569 entering apps across 23 markets.

Table 2 presents summary statistics for the two datasets used in the analysis, comparing pre-treatment (9 months before Apple's entry) and post-treatment (9 months after) averages. Market-level summary statistics are provided in Appendix A. Panel A reports app-month level statistics for incumbent apps—those present in the market before Apple's entry. The pre-treatment average price of \$0.81 reflects that 79.9% of apps are free; among paid apps, prices average \$4.01. The post-treatment period shows substantial changes: average prices increase to \$1.55, driven by a 15.9 percentage point decline in the share of free apps. In-app purchase adoption increases slightly (from 35.0% to 36.9%), while update frequency declines modestly (from 25.2% to 23.8%). Rating count growth (log-differenced) is positive in both periods, suggesting ongoing user engagement over time.

Panel B reports market-month level log entry and log exit counts. Entry rates decline modestly between periods, while exit rates increase slightly. Panel C reports statistics for the entering cohort analysis, a repeated cross-section where each app is observed once upon entry. Entrants in both

[6] Revenue ranking lists only reflect revenues earned through the platform—namely, download revenues (for non-free apps) and in-app purchase revenues (which can include subscriptions). Advertising revenue and revenue earned through other sources, such as the developer's website, are not reflected in these lists.

Table 2: Summary Statistics

	Pre-Treatment	Post-Treatment	Change
Panel A: Incumbent Apps			
Price	0.807	1.548	+0.741
Price (Paid Apps)	4.014	4.296	+0.282
Free	0.799	0.640	-0.159
In-App Purchases	0.350	0.369	+0.019
Avg. Rating	3.978	3.979	+0.001
Update	0.252	0.238	-0.013
Δ Log # Ratings	0.047	0.030	-0.017
N	26,262	26,262	
Panel B: Market-Level Entry and Exit			
Log Entry	1.340	1.198	-0.142
Log Exit	1.079	1.191	+0.113
N	414	414	
Panel C: Entering Cohorts			
Price	0.392	0.658	+0.266
Price (Paid Apps)	2.938	3.160	+0.221
Free	0.867	0.792	-0.075
In-App Purchases	0.153	0.173	+0.019
Avg. Rating	4.329	4.269	-0.060
Log # Ratings	0.767	0.918	+0.152
Similarity	0.655	0.653	-0.002
N	1,965	1,604	

Notes: Panel A reports app-month level statistics for incumbent apps observed in an 18-month balanced panel around each entry event. Panel B reports market-month level log entry and exit counts. Panel C reports statistics for apps observed once upon entry (repeated cross-section). Pre-Treatment refers to the 9 months before Apple's entry; Post-Treatment refers to the 9 months after.

periods are predominantly free (87% pre-treatment, 79% post-treatment) with substantially lower prices than incumbents, consistent with new entrants using zero-price strategies to compete against established apps. Similarity to the Apple entrant is nearly identical across periods (0.66).

3 Empirical Analysis

I estimate the effects of platform-owner entry on third-party developers using a staggered difference-in-differences design that exploits variation in the timing of Apple's entry across 23 submarkets, comparing App Store apps to similar apps on the Google Play Store as controls. I examine responses along two margins using the two samples defined in Section 2.4. Extensive-margin responses—entry, exit, and the composition of entering cohorts—are analyzed using the market dynamics sample. Intensive-margin responses—monetization and quality outcomes for established apps—are analyzed using the incumbent sample. In Section 3.1, I report pooled average treatment effects, then in Section 3.2 I show that these averages are poor summaries of most markets' experience by formally quantifying cross-market heterogeneity. To explore the sources and structure of this heterogeneity, Section 3.3 examines whether entry type—integrated versus standalone—explains variation across markets and Section 3.4 investigates whether competitive proximity shapes responses within markets.

The fundamental identification challenge is that one cannot observe counterfactual outcomes—what would have happened to affected apps absent Apple's entry. The identifying assumption is that absent Apple's entry, App Store and Play Store apps in the same product market would have followed similar outcome trajectories. The credibility of this design relies on two key institutional features: Google does not systematically enter the same markets at the same times as Apple, avoiding contamination of the control group, and both platforms experience similar technological and demand shocks during the sample period. The first requirement holds for all first-party entrants considered here, and the second requirement is supported by the fact that both platforms run frontier mobile operating systems on comparable hardware and compete for largely the same consumers in the mobile phone market. To further support this assumption, I present evidence regarding pre-trends through event study specifications throughout my analysis.

A potential threat to identification is spillover effects among multi-homing apps—products available on both platforms. If a developer adjusts their strategy in response to Apple's entry, their Play Store version might also change, contaminating the control group. However, platform-specific development teams, different technology stacks, and distinct user bases often lead to independent strategies across platforms. The analysis I present here includes all apps regardless of multi-homing status, maximizing sample size while acknowledging this potential limitation.[7]

[7] In Appendix E, I present results restricting the sample to single-homing apps—those available on only one platform—to verify that cross-platform spillovers do not drive the main findings.

3.1 Average Effects of Platform-Owner Entry

I begin by documenting average treatment effects for each set of outcomes: entry and exit dynamics, the characteristics of entering cohorts, and incumbent app responses.

3.1.1 Entry and Exit

I first examine how Apple's entry affects market structure through changes in the number of third-party apps entering and exiting markets. To do so, I estimate the two-way fixed effects (TWFE) model:

$$y_{apt} = \alpha \cdot D_{apt} + \delta_{ap} + \eta_t + \varepsilon_{apt}, \tag{3}$$

where y_{apt} is the outcome (log entry or exit count) for market a on platform p at month t. The treatment indicator $D_{apt} = \mathbf{1}[p = \text{App Store}, t \geq T_a]$ equals one for App Store observations in periods at or after market a's entry date T_a, and zero otherwise; Play Store observations have $D_{apt} = 0$ throughout, serving as a never-treated control group. δ_{ap} are market-platform fixed effects, η_t are year-month fixed effects, and α is the coefficient of interest. The outcome variables are market-level counts of new app entries and app exits, transformed using $\ln(x + 1)$ to handle zeros. Standard errors are clustered at the market-platform level. I estimate Equation (3) on a balanced market-platform-month panel dataset constructed from the market dynamics sample covering the 9 months before and after each entry event, using OLS.[8]

Table 3: Entry and Exit

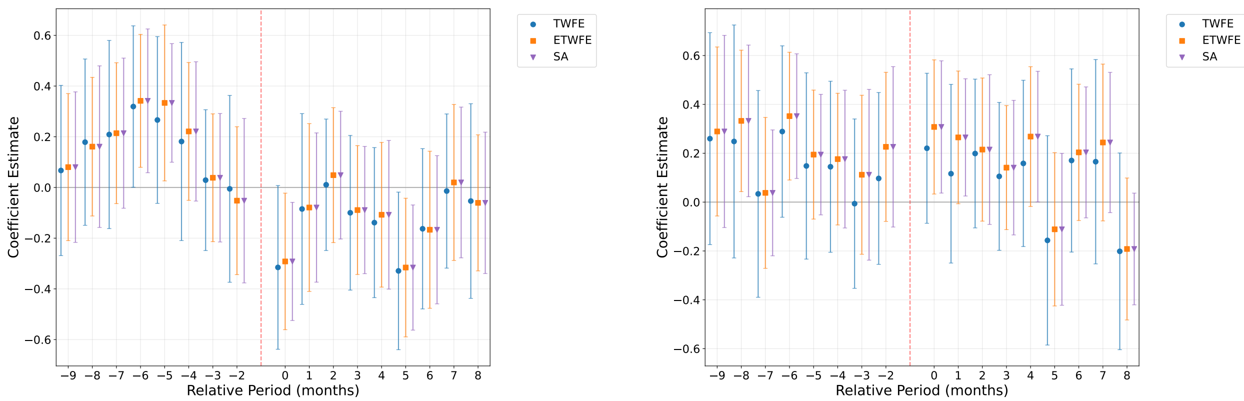
	(1) Log Entry Count	(2) Log Exit Count
ATT	-0.2653*** (0.0714)	-0.0209 (0.0746)
Baseline Mean	1.340	1.079
N	828	828

Notes: Two-way fixed effects (TWFE) DiD estimates at the platform-market-month level. Standard errors clustered at the platform-market level are in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

Table 3 presents the estimates and Figure 4 presents event study estimates for entry and exit. I find a large and statistically significant decline in third-party entry following Apple's entry. There is a nearly 23% reduction in new app entry relative to the pre-entry baseline. This indicates that Apple's presence substantially deters new development in affected markets. The exit regression, by contrast, shows no evidence of a change in exit rates. The absence of an exit effect, combined

[8] Recent work has shown that conventional TWFE estimators can produce biased estimates under treatment effect heterogeneity with staggered treatment timing. I therefore also estimate these effects using the heterogeneity-robust estimators of Wooldridge (2021) and Sun and Abraham (2021); results are similar across methods. I present event study estimates from these methods in the figures below, and estimates of the average effects in Appendix D.1.

Figure 4: Entry and Exit Dynamics



(a) Log Entry Count

(b) Log Exit Count

Note: Event study estimates showing dynamic treatment effects relative to one month before entry. TWFE is the primary estimator; ETWFE (Wooldridge (2021)) and SA (Sun and Abraham (2021)) shown for comparison. Analysis uses the market dynamics sample at the market-platform-month level. Standard errors are clustered at the market-platform level. Error bars represent 95% confidence intervals.

with the strong deterrence of entry, suggests that while Apple's entry discourages new competitors from entering the market, incumbent apps remain—potentially due to sunk development costs, the low cost of maintaining a listing, or the strategic value of maintaining platform presence. In Section 3.1.3, I consider an alternative measure of developers' participation and effort on the platform, the rate of product updating.

3.1.2 Composition of Entering Cohorts

Platform-owner entry can affect extensive-margin behavior by more than just impacting entry and exit counts—it may alter the composition of new entrants. I examine how entering cohorts differ before versus after Apple's entry using a cross-sectional difference-in-differences design. Each entering app is observed once, at entry. I estimate the model

$$y_j = \beta \cdot P_j + \alpha \cdot D_j + \delta_{ap} + \eta_t + \varepsilon_j, \tag{4}$$

using OLS. y_j represents entry characteristics for app j. $P_j = \mathbf{1}[t_j \geq T_{a(j)}]$ indicates whether app j entered at or after Apple's entry, regardless of platform; β captures the common post-entry shift. The treatment indicator $D_j = \mathbf{1}[j \in \text{App Store}, t_j \geq T_{a(j)}]$ equals one for App Store apps that entered at or after their market's treatment date, and zero otherwise, where t_j is app j's entry date and $a(j)$ denotes its market. Play Store apps have $D_j = 0$ throughout. δ_{ap} are market×platform fixed effects and η_t are calendar month fixed effects. The coefficient α captures the differential change in entry characteristics for App Store apps relative to Play Store apps following Apple's entry. Standard errors are clustered at the market-platform level.

The analyses that follow examine treatment effects along two broad dimensions. Monetization outcomes include the download price (in dollars, including zero-price apps), whether an app is offered for free, and whether it offers in-app purchases. Quality outcomes include the average user rating (1–5 stars), developer update frequency, and measures of consumer engagement: for entering cohorts, the initial level of rating counts and similarity to Apple's entrant; for incumbents, rating count growth (log-differenced).

Table 4 presents the results from estimating the cohort entry characteristics model Equation (4) using the market dynamics sample ($N = 3,569$ entering apps).[9] The results reveal systematic changes in the composition of entering apps following Apple's entry. Prices among new App Store entrants increase by \$0.27, driven by a 6.5 percentage point decline in the probability of offering a free app. This shift on the free/paid margin—arguably the primary monetization decision for app developers—suggests that developers entering after Apple may pursue differentiation strategies that can command a non-zero download price despite Apple's free alternatives. In-app purchase adoption shows no significant change.

Quality metrics at entry show imprecisely estimated effects. Ratings for new App Store entrants are approximately 0.15 stars higher than for Play Store entrants following Apple's entry, though this effect is not statistically significant. Rating count effects are also positive but imprecise. The table also reports effects on new entrants' similarity (measured via cosine similarity) to the Apple entrant, indicating whether developers' strategic product space positioning decisions are affected by Apple's entry. The pooled effect is near zero, which I discuss further in Section 3.2.

Table 4: Cohort ATT Estimates

	(1) Price	(2) Price (Paid Apps)	(3) Free App	(4) In-App Purchases	(5) Avg. Rating	(6) Log Rating Count	(7) Similarity
ATT	0.2701***	0.2855	-0.0652**	0.0015	0.1567	0.1529	0.0002
	(0.1012)	(0.4712)	(0.0264)	(0.0423)	(0.1346)	(0.1244)	(0.0062)
Baseline Mean	0.3918	2.938	0.8667	0.1725	4.329	0.7665	0.6547
N	3,569	596	3,569	2,365	1,032	3,569	3,569

Notes: Cross-sectional DiD estimates for entering cohorts. Each app observed once at entry. Standard errors clustered at market-platform level shown in parentheses. For the IAP outcome, 5 markets are excluded due to insufficient Play Store in-app purchase data in the pre-period. The Rating column has fewer observations because some entrants receive no ratings in their first month. *** p<0.01, ** p<0.05, * p<0.10.

3.1.3 Incumbent Responses

I next examine how incumbent apps respond to Apple's entry across monetization and quality dimensions. All analyses below use the incumbent sample, defined in Section 2.4: a balanced monthly panel of apps operating on the platforms for the 9 months before and after Apple's entry. Because this balanced panel conditions on survival throughout the 18-month window, the incumbent ATTs estimate effects for apps that remain active through the post-entry period; extensive-margin responses including exit are documented above. In my primary analysis, I estimate the two-way

[9] I present event study estimates in Appendix B.

fixed effects (TWFE) model

$$y_{jt} = \alpha \cdot D_{jt} + \gamma_j + \eta_{a(j),t} + \varepsilon_{jt}, \tag{5}$$

using OLS.[10] y_{jt} is the outcome for app j in month t. The treatment indicator $D_{jt} = \mathbf{1}[j \in \text{App Store}, t \geq T_{a(j)}]$ equals one for App Store apps in periods at or after their market's entry date, and zero otherwise, where $a(j)$ denotes the market containing app j. Play Store apps have $D_{jt} = 0$ throughout, serving as a never-treated control group. γ_j are app fixed effects and $\eta_{a(j),t}$ are market-month fixed effects, which absorb market-specific time shocks.[11] The coefficient α represents the average treatment effect on the treated (ATT). Standard errors are clustered at the app level.[12]

I estimate effects on the same monetization and quality outcomes described above, adapted for the panel setting. For rating count, I use the log-difference transformation: $\Delta \ln(\text{Rating Count}+1)$, which captures the growth rate of cumulative ratings. Since new ratings arrive as users download and engage with apps, rating count growth serves as a proxy for consumer demand and ongoing engagement with the product.

Table 5: Incumbent App Outcomes

	(1) Price	(2) Price (Paid Apps)	(3) Free App	(4) In-App Purchases	(5) Avg. Rating	(6) Update	(7) Δ Log Rating Count
ATT	0.3687***	0.0787	-0.0559***	-0.0125**	0.0227***	0.0143*	0.0191***
	(0.0513)	(0.0493)	(0.0097)	(0.0061)	(0.0072)	(0.0084)	(0.0038)
Baseline Mean	0.8069	4.014	0.7990	0.3626	3.978	0.2515	0.0468
N	52,524	14,741	52,524	41,310	50,099	52,524	52,524

Notes: Estimates from the incumbent analysis model at the app-month level. For the IAP outcome, 5 markets are excluded due to insufficient Android in-app purchase data in the pre-period. Standard errors clustered at the app level are in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

Table 5 presents estimates for all incumbent outcomes.[13] Prices increase by an average of \$0.37 following entry (column 1). Similar to what I found for the entering cohorts, this aggregate price increase masks important compositional changes. Apps become approximately 5.5 percentage points less likely to be free (column 3; a 7.0% decline). Conditional on charging a positive price, the price effect is modest (\$0.08) and not statistically significant (column 2), suggesting the overall price increase is driven primarily by apps switching from free to paid rather than existing paid apps raising prices. In-App Purchase (IAP) adoption falls by 1.3 percentage points (column 4).

[10] I report results from alternative, staggered-treatment robust DiD estimators in Appendix D.1. Results are similar across methods.

[11] For robustness, I re-estimate the model using app and month (as opposed to market-month) fixed effects in Appendix D.2.

[12] I cluster at the app level to address within-app serial correlation. Because the analysis covers the full population of apps on both platforms, the design-based case for coarser clustering is weaker than in a sampled setting (Abadie et al., 2023). In Appendix D.3, I report cluster-jackknife (CRV3J) standard errors at the market-platform level, which guard against downward bias with few clusters (MacKinnon, Nielsen, and Webb, 2023; Hansen, 2025). Quality results are robust to this alternative inference; monetization results retain their signs and magnitudes but are less precisely estimated, consistent with the concentration of pricing effects in a subset of markets that I document in Section 3.2.

[13] These estimates use the full incumbent sample including multi-homing apps. In Appendix E, I show that restricting to single-homing apps yields similar results, confirming that cross-platform spillovers do not drive the findings.

One interpretation is that developers simplify monetization strategies under competitive pressure, moving from complex freemium models to straightforward paid downloads, or that they move monetization from on-platform IAPs to off-platform subscriptions, which are not subject to Apple's typical commission.

Columns 5–7 of Table 5 present estimates. Average ratings improve slightly (column 5) and update frequency increases by nearly 6% (column 6). Rating count growth also shows a positive effect: an increase of approximately 0.019 on the log-difference scale, corresponding to roughly 1.9 percent additional growth (column 7). This suggests that Apple's entry may expand overall category visibility or demand, benefiting both incumbents and the platform.

Figure 5 presents event studies for all incumbent outcomes. Estimates are consistent across estimators. The price increase and the decline in the proportion of free apps both show large initial responses that settle at less extreme but still substantial levels. This suggests developers may engage in some experimentation immediately following Apple's entry into their market. IAP adoption begins to decline at entry and shows signs of stabilizing around the fourth post-treatment month.

These pooled estimates provide useful benchmarks, but they may conceal substantial cross-market variation. Understanding whether the averages mask opposing responses across markets—and what drives any such differences—has direct implications for the design of platform regulation. I turn to this question next.

3.2 Cross-Market Heterogeneity

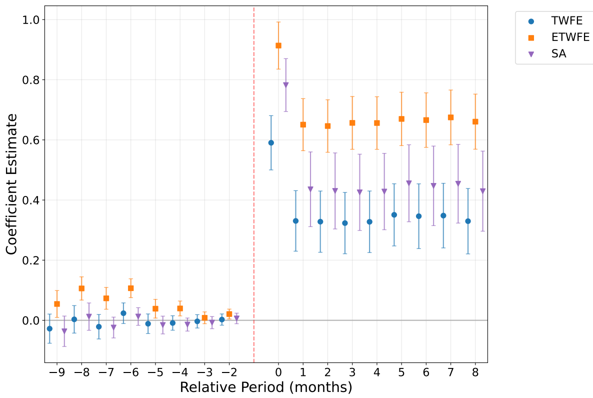
The average effects documented thus far need not describe any individual market's experience. If markets differ in pre-existing competitive conditions, the nature of Apple's product, or the strength of consumer demand, the effect of entry may vary substantially—and markets may even respond in opposing directions. To quantify this cross-market heterogeneity in treatment effects, I generalize Equations (3) to (5) by allowing market-specific ATTs $\alpha_{a(j)}$, interacting the treatment indicator with market dummies. This specification enables me to examine heterogeneity across markets using meta-analytic diagnostics.

To characterize cross-market variation in treatment effects, I apply heterogeneity measures commonly used in meta-analysis to the set of market-specific ATTs. These statistics provide a convenient summary of dispersion across market-level estimates. Specifically, I calculate three diagnostics. Cochran's Q statistic,

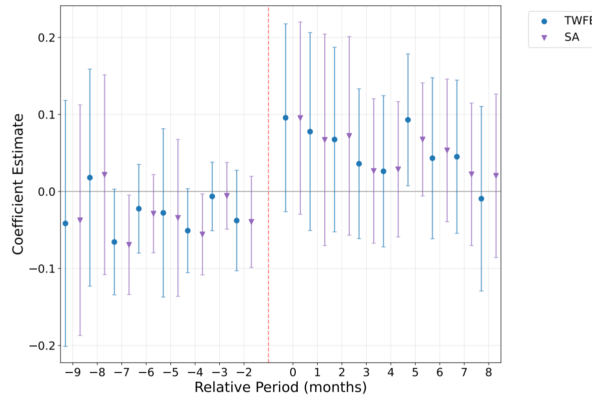
$$Q = \sum_{a=1}^{A} w_a \big(\hat{\alpha}_a - \bar{\alpha}\big)^2,$$

where $w_a = 1/\widehat{\mathrm{Var}}(\hat{\alpha}_a)$ are inverse-variance weights, $\bar{\alpha}$ is the pooled ATT, and A is the number of

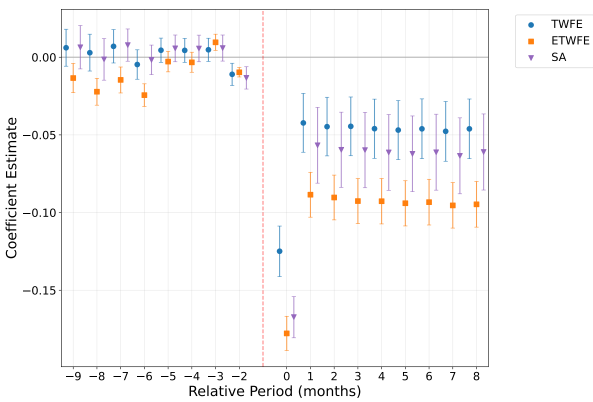
Figure 5: Incumbent Treatment Effect Dynamics



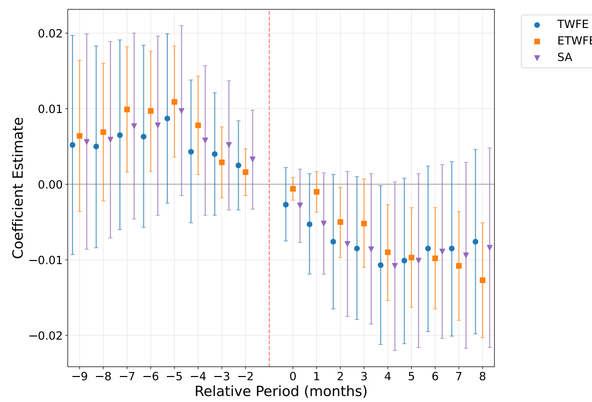
(a) Price (All Apps)



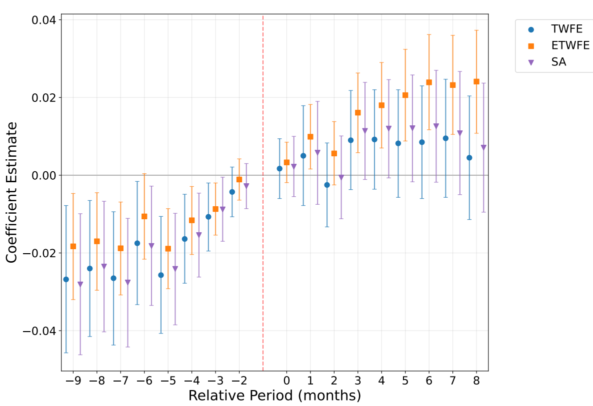
(b) Price (Paid Apps Only)



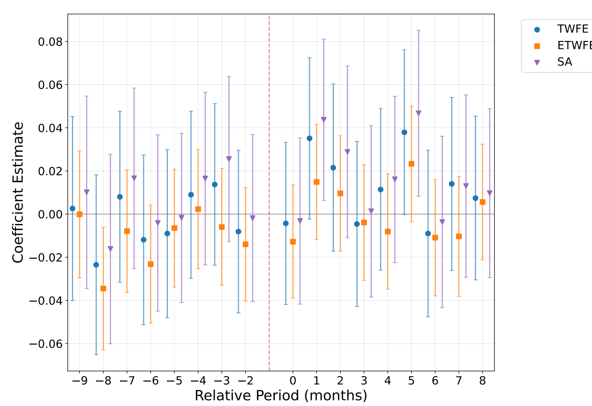
(c) Free



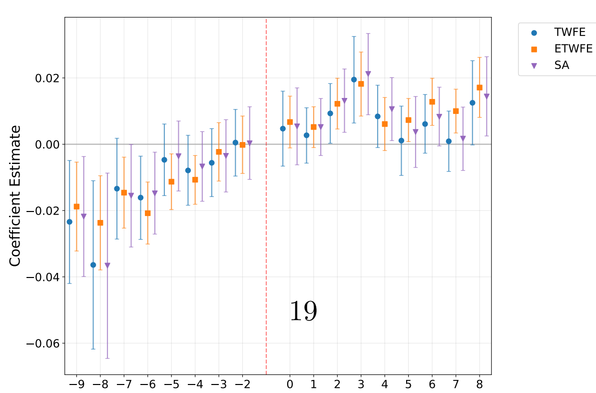
(d) In-App Purchases



(e) Rating



(f) Update Frequency



markets. This tests the null hypothesis that all market-level effects are equal. The I^2 statistic,

$$I^2 = \max\left\{0,\ \frac{Q - (A-1)}{Q}\right\} \times 100\%,$$

quantifies the proportion of total variation attributable to true between-market heterogeneity rather than sampling error (Higgins and Thompson, 2002; Higgins et al., 2003). Finally, the DerSimonian–Laird estimate of τ^2 (DerSimonian and Laird, 1986) provides the between-market variance of true effects. In the tables that follow, I report $\sqrt{\tau^2}$, the between-market standard deviation, because it is directly comparable in scale to the pooled ATTs reported in the preceding tables. Finally, I report sign agreement, which measures the fraction of market-level estimates sharing the same sign as the average effects from Section 3.1, and the percentage of markets where the effect is statistically indistinguishable from zero at the 5% level (% Null).

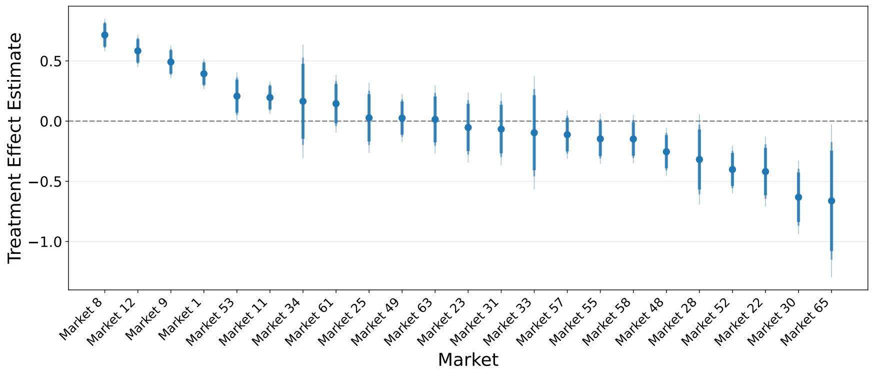
Figure 6 presents market-specific treatment effects for three illustrative outcomes, each drawn from a different margin of analysis.[14] The exit result is perhaps the most striking: the near-zero effect documented in Section 3.1.1 masks genuinely opposing dynamics across markets. With only 52% sign agreement, roughly half of markets experience increased exit following Apple's entry while the other half experience decreased exit—and many of these market-level estimates are themselves statistically insignificant, reinforcing the absence of a common exit response. Similarity to the Apple entrant—for which the pooled cohort effect was essentially zero—displays a similarly divided pattern, with some markets seeing new entrants cluster toward Apple's product position and others shifting away. By contrast, incumbent update frequency tells a different story: most market-level point estimates are positive (70% sign agreement), though 91% are statistically indistinguishable from zero at the individual market level—consistent with a modest, relatively uniform effect that is difficult to detect with market-level precision.

Table 6 reports the full set of meta-analytic diagnostics across all three margins. The pattern is pervasive: most outcomes exhibit substantial cross-market heterogeneity, indicating that the pooled ATTs from Section 3.1 are consistently poor summaries of any individual market's experience. Following the taxonomy of Higgins et al. (2003), monetization heterogeneity is generally high for the primary pricing margins ($I^2 \geq 71\%$ for unconditional price and free/paid status across both cohort and incumbent analyses), though more moderate for in-app purchases among incumbents ($I^2 = 50\%$). Quality heterogeneity ranges from low (incumbent update frequency, $I^2 = 18\%$, p-value for $Q = 0.214$) to very high (cohort rating, $I^2 = 91\%$). Entry and exit effects show substantial heterogeneity, with exit displaying $I^2 = 96\%$.

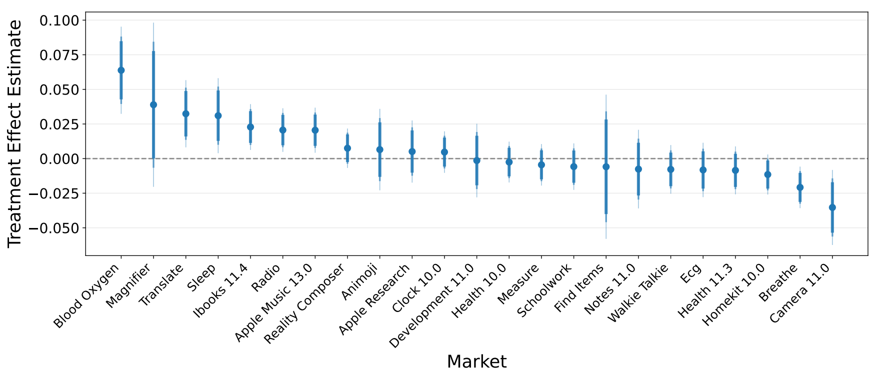
The cohort similarity result deserves particular attention. The near-zero pooled effect (0.0002) with $I^2 = 82\%$ and only 48% sign agreement indicates that this null average masks strongly opposing market-level dynamics. One interpretation is that Apple's entry may serve a coordinating function in some markets—signaling consumer demand, validating product categories, or concentrating developer effort around capabilities Apple has demonstrated or wishes to improve on its

[14] See Appendix C for the remaining figures.

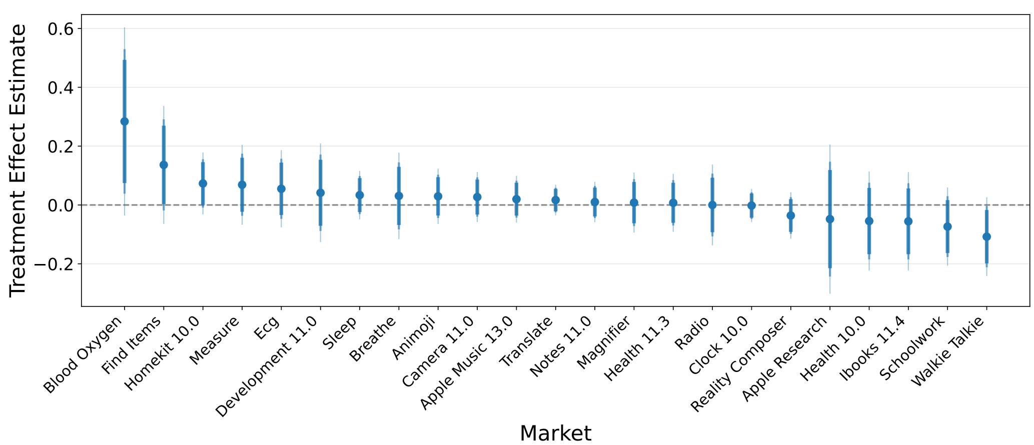
Figure 6: Selected Market-Specific Treatment Effects



(a) Exit Count (market dynamics sample, market-platform-month level)



(b) Similarity to Apple Entrant (market dynamics sample, app-level, observed once at entry)



(c) Incumbent Update Frequency (incumbent sample, app-month level)

Note: Market-specific treatment effects from Equations (3) to (5). Points show estimates with 90%, 95%, and 99% confidence intervals (thick to thin lines). Panel (a) clusters at the market-platform level; panels (b) and (c) cluster at the market-platform and app levels, respectively. Forest plots for all outcomes are presented in Appendix C.

Table 6: Cross-Market Heterogeneity Diagnostics

Outcome	Est. (SE)	Sign Agree.	% Null	I^2 (Q p)	$\sqrt{\tau^2}$
Panel A: Entry and Exit					
Entry Count	-0.265 (0.071)	78%	30%	87% (<0.001)	0.294
Exit Count	-0.021 (0.075)	52%	48%	96% (<0.001)	0.360
Panel B: Entering Cohorts					
Monetization					
Price	0.270 (0.101)	78%	65%	84% (<0.001)	0.242
Free	-0.065 (0.026)	78%	70%	71% (<0.001)	0.063
In-App Purchases	0.002 (0.042)	72%	44%	89% (<0.001)	0.113
Price (Paid)	0.285 (0.471)	50%	36%	90% (<0.001)	2.093
Quality					
Rating	0.157 (0.135)	65%	55%	91% (<0.001)	0.484
Rating Count	0.153 (0.124)	65%	57%	85% (<0.001)	0.461
Similarity	0.000 (0.006)	48%	61%	82% (<0.001)	0.016
Panel C: Incumbent Responses					
Monetization					
Price	0.369 (0.051)	65%	76%	83% (<0.001)	0.075
Free	-0.056 (0.010)	64%	79%	79% (<0.001)	0.023
In-App Purchases	-0.013 (0.006)	64%	86%	50% (0.017)	0.005
Price (Paid)	0.079 (0.049)	58%	92%	13% (0.322)	0.020
Quality					
Rating	0.023 (0.007)	70%	83%	56% (<0.001)	0.031
Update Freq.	0.014 (0.008)	70%	91%	18% (0.214)	0.018
Rating Count Growth	0.019 (0.004)	70%	74%	47% (0.007)	0.015

Notes: This table reports heterogeneity diagnostics applied to market-level ATTs from TWFE regressions. Panel A uses the market dynamics sample at the platform-market-month level. Panel B uses the market dynamics sample at the app level (each app observed once at entry). Panel C uses the incumbent sample at the app-month level with market×month fixed effects. Est. is the pooled TWFE ATT with standard error in parentheses. Sign Agreement is the share of markets whose point estimate matches the sign of the overall estimate. % Null is the share of markets where the 95% confidence interval includes zero. I^2 measures the proportion of total variation due to between-market heterogeneity rather than sampling error; Q p-value tests the null of homogeneous effects. $\sqrt{\tau^2}$ is the DerSimonian–Laird estimate of the between-market standard deviation of true effects.

platform. In these markets, new entrants cluster toward Apple because its entry reduces uncertainty about viable product directions. In other markets, Apple's presence may instead push developers to seek new niches, with new entrants differentiating away from the now-occupied product position. The heterogeneity in similarity effects across markets is consistent with both mechanisms operating in different contexts, a finding that would be obscured by relying solely on the pooled (null) effect.

Finally, update frequency stands out as the one outcome where heterogeneity is low: $I^2 = 18\%$ with the p-value for Q indicating that the null hypothesis of homogeneous effects cannot be rejected. The market-level estimates do not differ much from one another, consistent with a modest, relatively uniform positive effect.

These diagnostics raise two questions for the remainder of this section. First, what explains the cross-market differences? In Section 3.3, I investigate the role of entry type—whether Apple enters via a standalone app or OS integration—in driving cross-market differences. Second, are market-level averages themselves adequate summaries of within-market effects, or does an app's competitive proximity to Apple shape its response? In Section 3.4, I address this question using continuous distance measures derived from product descriptions.

3.3 Entry Type and Competitive Dynamics

The cross-market heterogeneity documented in Section 3.2 motivates examining whether observable market characteristics explain the variation in treatment effects. I investigate one salient dimension—entry type, whether Apple enters via a new standalone app or OS integration—which varies across markets and has clear theoretical implications for competitive dynamics.

To directly test whether entry type affects competitive outcomes among incumbent apps, I estimate a variation of Equation (5) with entry type interactions,

$$y_{jt} = \alpha_{\text{SA}} \cdot D_{jt} + \alpha_{\text{Diff}} \cdot D_{jt} \times \text{Integrated}_j + \gamma_j + \eta_{a(j),t} + \varepsilon_{jt}, \tag{6}$$

where D_{jt} is defined as in Equation (5) and Integrated_j indicates whether app j is in a market where Apple entered via OS integration rather than a standalone app. The coefficient α_{SA} captures the treatment effect for standalone entry (the baseline group), α_{Diff} captures the additional effect for integrated entry, and their sum $\alpha_{\text{SA}} + \alpha_{\text{Diff}}$ yields the total effect for integrated entry. Standard errors are clustered at the app level. This specification directly tests the hypothesis that entry type matters: a significant α_{Diff} indicates systematically different responses. I present the results in Table 7. The "All" column reproduces the pooled ATT from the baseline specification (Equation (5)) for reference, while the remaining columns report estimates from the interaction specification (Equation (6)).

Two key patterns emerge from Table 7. First, integrated entry generates larger price effects than standalone entry. These differences suggest distinct competitive mechanisms: integrated entry forces more aggressive pricing responses, while standalone entry generates more modest adjustments. Second, the entry types exhibit similar demand responses, as proxied by the growth

Table 7: Entry Type Effects on Incumbent Apps: Standalone vs. Integrated

Outcome	All	Integrated	Standalone	Difference
Price	0.369***	0.653***	0.044*	0.609***
	(0.0513)	(0.0911)	(0.0240)	(0.0942)
Free	-0.056***	-0.103***	-0.0006	-0.102***
	(0.0097)	(0.0168)	(0.0032)	(0.0171)
In-App Purchases	-0.013**	-0.0029	-0.012*	0.0087
	(0.0061)	(0.0052)	(0.0067)	(0.0085)
Price (Paid Apps)	0.079	-0.0085	0.063	-0.072
	(0.0493)	(0.0775)	(0.0530)	(0.0938)
Update	0.014*	0.026	-0.0031	0.029
	(0.0084)	(0.0213)	(0.0205)	(0.0295)
Rating	0.023***	0.0076	0.0040	0.0037
	(0.0072)	(0.0072)	(0.0085)	(0.0111)
Δ Log Rating Count	0.019***	0.0099*	0.0043	0.0055
	(0.0038)	(0.0052)	(0.0064)	(0.0082)

Notes: TWFE estimates with standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.10$. All: 23 markets (14 standalone + 9 integrated). Difference: Integrated − Standalone.

rate of rating counts. This suggests that any attention-grabbing demand expansion associated with Apple's entry into a submarket does not necessarily depend on how Apple implements entry.

One possible explanation for the price differences is the unavoidability of integrated releases. OS-integrated features cannot be separately uninstalled or avoided by users, creating fundamentally different competitive pressure from that of marketplace apps. When Apple releases a standalone app like Translate, users must actively discover it in the App Store, evaluate it against alternatives, and choose to download it. Market forces can operate normally—if third-party alternatives are superior, users may never adopt Apple's offering.[15]

Integrated entry eliminates this choice architecture. When Apple integrates translation capabilities directly into iOS, every user automatically possesses this functionality upon updating their device. Third-party translation apps must now compete against a feature that users already have, cannot remove, and may discover through system prompts or Siri suggestions. This unavoidability likely explains the larger monetization responses to integrated entry: facing an unremovable competitor, third-party apps must differentiate more aggressively through pricing and business model changes.

[15] Self-preferencing by the platform—such as favorable search placement or default settings—may blunt this competitive discipline even for standalone entries.

3.4 Product Space Dynamics

Notably, not all apps in an affected market are equally "treated" by platform entry. An app providing nearly identical functionality to Apple's entrant may face different competitive pressures than a tangentially related app at the market periphery. I exploit the continuous similarity measure from the SBERT embeddings described in Section 2.3 to estimate how treatment effects vary with competitive proximity. I use B-splines to flexibly estimate this relationship within the difference-in-differences framework.

For each app j in market a with cosine similarity $d_j \in [\theta, 1]$ to the entrant, I estimate similarity heterogeneity by interacting K B-spline basis functions with the treatment indicator:

$$y_{jt} = \sum_{k=1}^{K} \phi_k \cdot D_{jt} \cdot B_k(d_j) + \gamma_j + \eta_{a(j),t} + \varepsilon_{jt}. \tag{7}$$

D_{jt} is defined as above, $B_k(\cdot)$ are B-spline basis functions, K is the number of basis functions (degrees of freedom), γ_j are app fixed effects, and $\eta_{a(j),t}$ are market×month fixed effects. The inclusion of market×month fixed effects ensures that distance-invariant market-wide shocks are absorbed, so that $\tau(d)$ captures variation in treatment effects along the similarity gradient net of market-wide trends. This specification assumes that, absent treatment, the iOS–Android outcome gap does not vary systematically with similarity to Apple's product within a market—an assumption supported by the pre-trends documented in the event studies. This approach allows me to flexibly estimate how the treatment effect varies in an app's similarity to Apple's first-party product without imposing a specific functional form.

I select K via cross-validation, which yields $K = 3$.[16] In the analysis below, I plot the treatment effect at each similarity level,

$$\tau(d) = \sum_{k=1}^{K} \phi_k \, B_k(d). \tag{8}$$

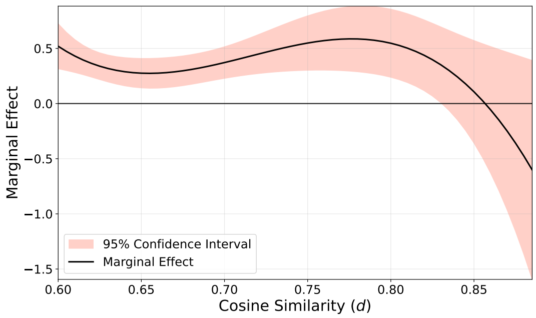
This captures the treatment effect for apps at similarity d, showing how competitive pressure varies with distance from the entrant.[17]

Figure 7 presents $\tau(d)$ for all incumbent outcomes, showing how treatment effects vary with competitive proximity to the first-party entrant. The unconditional price effect (panel (a)) is positive and statistically significant across most of the similarity range, with a non-monotonic shape that peaks among apps with moderately high similarity to Apple's entrant before reversing sharply—though imprecisely—for the most similar apps. The probability of being free (panel (c)) shows a relatively uniform decline of 5–6 percentage points that varies little with similarity, steepening at

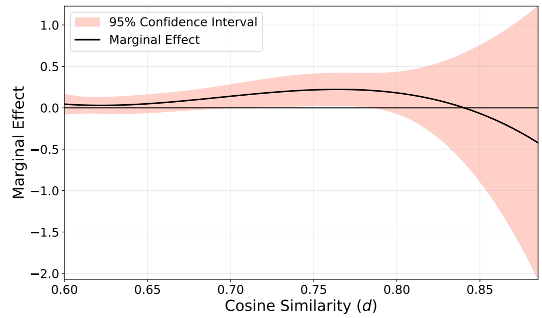
[16]Specifically, I use 10-fold cross-validation at the app level, testing values from 2 to 8 and applying the one-standard-error rule (Breiman et al., 1984; Hastie, Tibshirani, and Friedman, 2009): among all candidate values, I choose the most parsimonious K whose mean cross-validation error falls within one standard error of the minimum. This favors simpler models when additional complexity yields negligible predictive improvement, guarding against overfitting. In practice, the error surface is nearly flat across all candidate values.

[17]I evaluate and plot $\tau(d)$ on a grid $d \in [\theta, 1]$ together with pointwise 95% confidence intervals (standard errors are clustered at the app level).

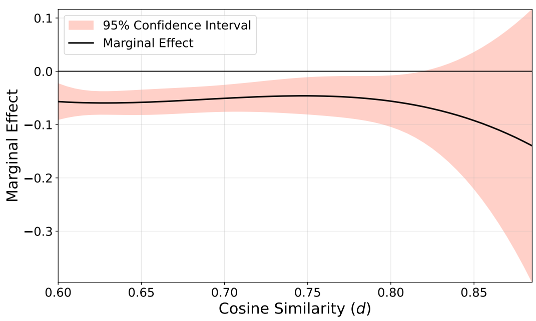
Figure 7: Incumbent Effects by Distance from Entrant



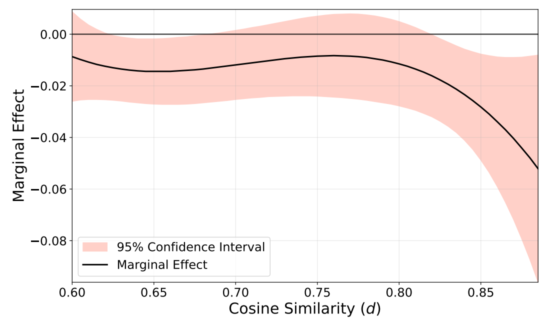
(a) Price (All Apps)



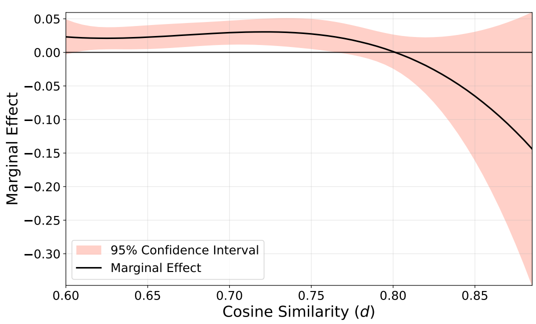
(b) Price (Paid Apps Only)



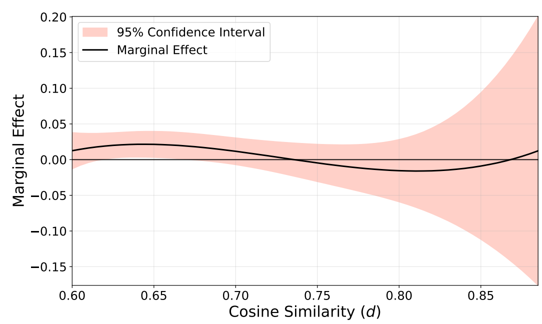
(c) Free



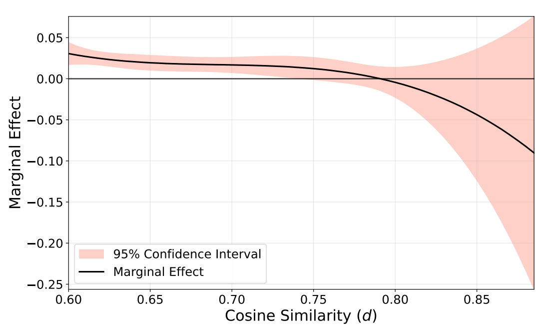
(d) In-App Purchases



(e) Rating



(f) Update Frequency



(g) Rating Count Growth

Note: Treatment effects from the incumbent distance model, Equation (7). For the IAP outcome, 5 markets are excluded due to insufficient Play Store in-app purchase data in the pre-period. Increasing cosine similarity indicates greater similarity to Apple's product, located at $d = 1$. Standard errors are clustered at the app level. Shaded areas represent 95% confidence intervals.

the highest similarity levels though with wider confidence intervals, suggesting this monetization shift is broadly distributed rather than concentrated among the closest competitors. Conditional on being paid (panel (b)), the price effect is small and generally statistically insignificant, confirming that the overall price increase reflects apps switching from free to paid rather than paid apps raising prices. The IAP effect (panel (d)) is modestly negative throughout, becoming significantly larger for the most similar apps, consistent with the closest competitors reducing in-app purchase offerings in response to entry.

The quality panels reveal a distinct pattern. Ratings (panel (e)) are significantly positive in the mid-similarity range but turn sharply negative for the apps most similar to Apple's entrant, though this extreme-similarity estimate is imprecise. This is consistent with the closest competitors being most harmed by first-party entry, while more peripheral apps may benefit from complementarities with Apple's entrant, or because they may offer a sufficient differentiated, possibly premium, version of the product. Update frequency (panel (f)) shows effects near zero and generally insignificant across the similarity range, consistent with the small, homogeneous cross-market effect documented above. Finally, rating count growth (panel (g)) is positive and significant for moderately similar apps ($d < 0.75$), declining through zero and turning negative for the most similar competitors. This suggests the positive demand shocks associated with platform-owner entry tend to accrue to apps that maintain sufficient differentiation from Apple's product.

The pooled distance gradients in Figure 7 show that apps closer to Apple's offering tend to experience more extreme effects. However, these pooled curves average over substantial cross-market variation. Figure 8 decomposes these into market-specific distance curves, estimated from a single pooled regression that interacts the B-spline basis functions in Equation (7) with market indicators, allowing each market its own distance gradient $\tau_a(d)$.

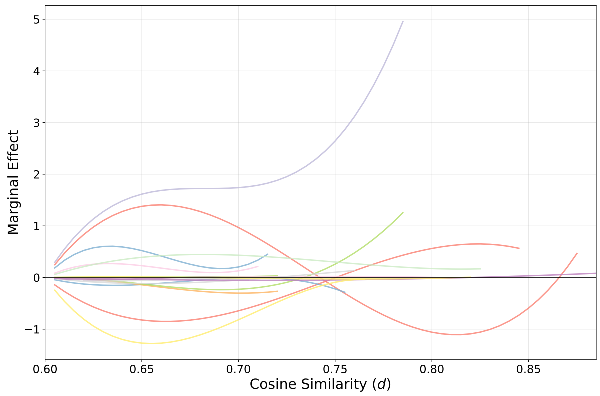
Because markets differ in their distribution of apps' cosine similarities to Apple's product, each market's curve is displayed only within its observed similarity range, avoiding extrapolation beyond the data support. Some outcomes, such as price, present remarkable variation among affected markets: some markets exhibit substantial increases, while others shift sign as cosine similarity increases. Others, such as Rating Count Growth, show a tighter distribution of curves.

Taken together, the distance gradients for incumbents and the similarity heterogeneity for entering cohorts documented in Section 3.2 paint a consistent picture: Apple's entry reshapes the competitive landscape non-uniformly, with both the direction and magnitude of effects depending on a firm's proximity to the entrant and on market-specific conditions.

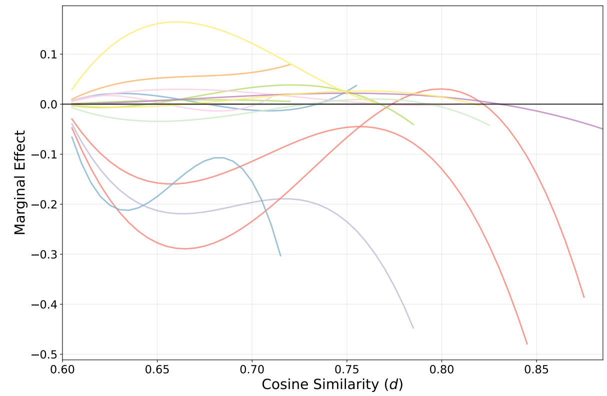
4 Concluding Discussion

This paper provides systematic evidence on how platform-owner entry affects third-party developers across multiple markets. Studying 23 instances of Apple entering submarkets in its App Store—14 through standalone app releases and 9 through integrated OS features—I find that platform-owner entry deters new third-party entrants by 23% and reshapes incumbent monetization strategies, but

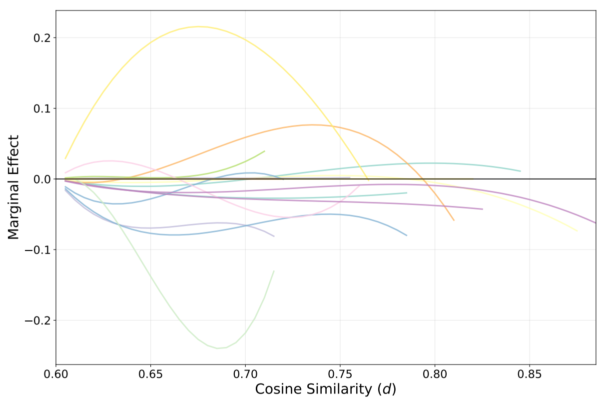
Figure 8: Incumbent Distance Gradients Across All Markets



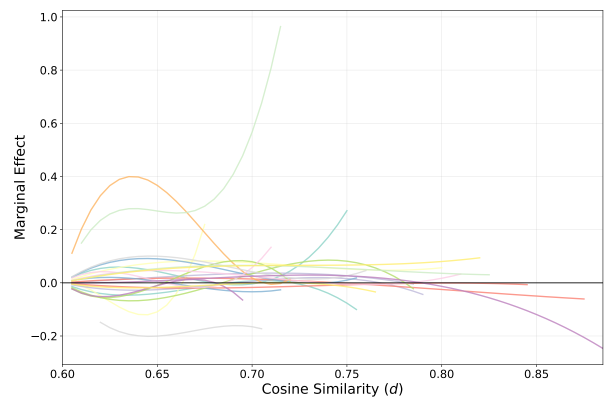
(a) Price



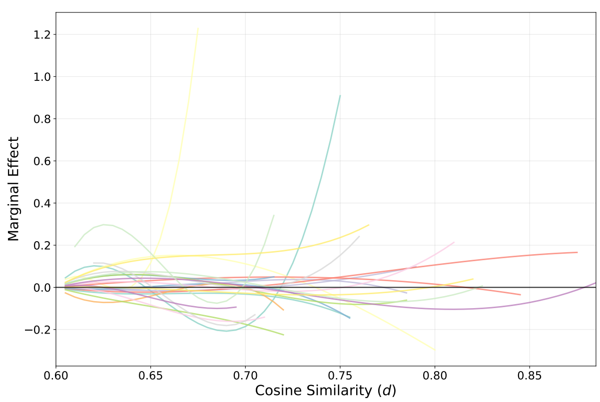
(b) Free



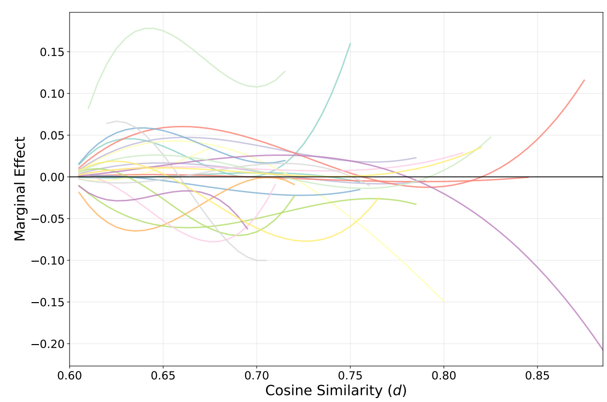
(c) In-App Purchases



(d) Avg. Rating



(e) Update



(f) Rating Count Growth

Note: Market-specific treatment effects from the incumbent distance model, Equation (7). Analysis is at the app-month level. For the IAP outcome, 5 markets are excluded due to insufficient Play Store in-app purchase data in the pre-period. Increasing cosine similarity indicates greater similarity to Apple's product, located at $d = 1$. Standard errors are clustered at the app level. Each line represents one market's treatment effects across distance. Curves are restricted to each market's observed similarity range to avoid extrapolation beyond the data.

that these averages mask wide variation across markets and outcomes. The evidence reveals that platform entry is neither uniformly beneficial nor uniformly harmful, but rather generates widely varying effects that depend systematically on market context, competitive proximity, and entry implementation.

Cross-market heterogeneity is pervasive: meta-analytic diagnostics reveal high I^2 values across all three margins of analysis—entry/exit dynamics, cohort composition, and incumbent responses—indicating that average effects often poorly represent any individual market. Perhaps most surprisingly, I find that many markets experience no statistically significant effects across several outcomes, with quality metrics particularly unaffected. This prevalence of null effects contradicts popular claims of either universal harm from foreclosure or overwhelming benefits from technology spillovers.

Within markets, competitive proximity shapes effects through distance gradients that themselves vary across markets. Entry type matters fundamentally: integrated OS features generate pricing effects on the unconditional price and free/paid margins that are roughly an order of magnitude larger than those from standalone apps. New entrant positioning also reveals that Apple's entry reshapes the product space differently across markets, with entrants clustering toward or differentiating away from Apple's position depending on market context.

Several important limitations qualify these findings and their policy implications. Most fundamentally, the reduced-form approach employed here has both advantages and disadvantages. A key advantage is that it readily accommodates the wide range of (potentially conflicting) mechanisms playing out at once when a platform-owner enters its digital marketplace. This approach is not restricted to modeling just one or two of the relevant factors, instead providing insight as to the net competitive effects of entry. However, this approach is also limited in that it cannot directly assess consumer welfare. Additionally, the combination and variation of effects documented here have ambiguous welfare implications. For example, apps changing from free to paid pricing may also reduce reliance on or otherwise change advertising intensity, data collection practices, or on- or off-platform subscription models that affect consumer welfare in ways this analysis cannot capture.

Despite these limitations, these findings open important avenues for future research. The stark heterogeneity I document—across markets, within markets by competitive distance, and between analysis margins—calls for deeper investigation into what determines these differential effects. The heterogeneous similarity results for new entrants documented in Section 3.2—where new entrants cluster toward or differentiate away from Apple's position depending on market context—suggest rich strategic interactions that other approaches could help disentangle. Understanding what drives the opposing exit responses across markets, or why similarity effects reverse sign, could help predict which markets are most vulnerable to platform competition and inform more targeted regulatory approaches. The systematic differences between standalone and integrated entry also warrant further study, particularly as platforms increasingly blur the lines between operating system features and marketplace applications.

The longer-run dynamics of platform competition remain largely unexplored. The nine-month post-entry window considered here may miss important evolutionary patterns where initial defensive responses give way to accommodation, innovation, or exit. Technology spillovers from platform entry might take years to fully materialize, particularly when new APIs and frameworks require developers to learn new skills and rebuild applications. Understanding these dynamic patterns could reveal whether the immediate competitive harm from platform entry is offset by longer-run innovation benefits, or whether initial advantages compound over time into durable market dominance.

These findings provide crucial empirical evidence for the ongoing global debate over platform regulation and antitrust enforcement. The EU's Digital Markets Act restricts gatekeepers from self-preferencing (EUR-Lex, 2022), while the previously proposed American Innovation and Choice Online Act would have similarly constrained platforms' ability to leverage nonpublic data against their business users (Congress.gov, 2022). The evidence reveals that calls for blanket restrictions (see, for example, Warren (2019)) fundamentally misunderstand the nature of platform competition. The prevalence of null effects demonstrates that categorical bans would prevent entry that causes little to no harm to the vast majority of apps in affected markets. More troublingly, the concentration of effects among close competitors means that broad statutory language could restrict entry that affects only a narrow segment of apps while leaving the vast majority unaffected.

The differences between standalone and integrated entry provide actionable guidance for antitrust enforcement and regulatory design. Integrated OS features present fundamentally different competitive concerns than standalone apps that users can choose to ignore. This distinction maps directly onto antitrust theory: integrated features resemble tying-like arrangements or technological bundling that can increase foreclosure risk, while standalone apps are more likely to compete on merit in the marketplace.[18] Antitrust authorities should therefore apply different levels of scrutiny based on entry mode, with integrated features triggering heightened review similar to merger analysis in adjacent markets. The unavoidability of integrated features creates the kind of foreclosure concerns that antitrust law is designed to address, while standalone apps' market-based adoption provides the competitive constraints that normally obviate regulatory intervention. The heterogeneity across markets further challenges the premise underlying many current regulatory proposals—that platform entry is categorically harmful to innovation and competition. Rather than categorical entry prohibitions, the findings support regulatory scrutiny calibrated to entry mode and competitive proximity—with integrated features warranting heightened attention given their substantially larger effects, and market definition in platform cases accounting for the concentration of effects among close competitors.

Ultimately, as policymakers worldwide grapple with platform power, evidence-based approaches that recognize the fundamental heterogeneity in platform competition effects will be essential for crafting interventions that preserve innovation incentives while preventing competitive harm. The complex, multi-dimensional nature of platform competition revealed in this study underscores that simple narratives fail to capture the nuanced reality of digital markets.

[18] Pre-installed standalone apps, however, may warrant greater scrutiny.

References

Abadie, Alberto, Susan Athey, Guido W. Imbens, and Jeffrey M. Wooldridge. 2023. "When Should You Adjust Standard Errors for Clustering?" *The Quarterly Journal of Economics* 138 (1):1–35.

Anderson, Simon and Özlem Bedre Defolie. 2024. "Hybrid platform model: monopolistic competition and a dominant firm." *The RAND Journal of Economics* 55 (4):684–718. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/1756-2171.12478.

AppMonsta. 2016. "App Store Data and API." URL https://appmonsta.com. Accessed: 2024.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-In-Differences Estimates?" *The Quarterly Journal of Economics* 119 (1):249–275.

Boudreau, Kevin J. and Andrei Hagiu. 2011. "Platform Rules: Multi-Sided Platforms as Regulators." In *Platforms, Markets, and Innovation*, edited by Annabelle Gawer. Edward Elgar Publishing, Inc.

Breiman, Leo, Jerome Friedman, Charles J. Stone, and R. A. Olshen. 1984. *Classification and Regression Trees*. Belmont, CA: Wadsworth.

Cheyre, Cristobal, Benjamin T. Leyden, Sagar Baviskar, and Alessandro Acquisti. 2025. "Did Apple's App Tracking Transparency Framework Harm the App Ecosystem?" *Working Paper* .

Congress.gov. 2022. "Text – S. 2992 – 117th Congress (2021–2022): American Innovation and Choice Online Act." https://www.congress.gov/bill/117th-congress/senate-bill/2992/text. Accessed 23 June 2025.

Crémer, Jacques, Yves-Alexandre de Montjoye, and Heike Schweitzer. 2019. "Competition policy for the digital era." URL https://data.europa.eu/doi/10.2763/407537.

De Chant, Tim. 2022. "Antitrust bill that bars Big Tech self-preferencing advances in Senate." *Ars Technica* URL https://arstechnica.com/tech-policy/2022/01/antitrust-bill-that-bars-big-tech-self-preferencing-advances-in-senate/.

DerSimonian, Rebecca and Nan Laird. 1986. "Meta-Analysis in Clinical Trials." *Controlled Clinical Trials* 7 (3):177–188.

Ershov, Daniel. 2024. "Variety-Based Congestion in Online Markets: Evidence from Mobile Apps." *American Economic Journal: Microeconomics* 16 (2):180–203. URL https://www.aeaweb.org/articles?id=10.1257/mic.20200347.

EUR-Lex. 2022. "Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act)." http://data.europa.eu/eli/reg/2022/1925/oj. Official Journal L 265/1, 12 October 2022. Accessed 23 June 2025.

Foerderer, Jens, Thomas Kude, Sunil Mithas, and Armin Heinzl. 2018. "Does Platform Owner's Entry Crowd Out Innovation? Evidence from Google Photos." *Information Systems Research* 29 (2):444–460.

Gugler, Klaus, Florian Szücs, and Ulrich Wohak. 2024. "Using Natural Language Processing to Delineate Digital Markets." *Stanford Computational Antitrust* IV:33–52.

Hagiu, Andrei, Tat-How Teh, and Julian Wright. 2022. "Should platforms be allowed to sell on their own marketplaces?" *The RAND Journal of Economics* 53 (2):297–327. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/1756-2171.12408.

Halckenhaeusser, André, Jens Forderer, Armin Heinzl, and Ola Henfridsson. 2025. "Strategic Drivers of Core Expansion on Software Platforms: Evidence from Apple iOS." *Information Systems Research* URL https://doi.org/10.1287/isre.2023.0347. Articles in Advance.

Han, Sukjin, Eric Schulman, Kristen Grauman, and Santhosh Ramakrishnan. 2024. "Shapes as Product Differentiation: Neural Network Embedding in the Analysis of Markets for Fonts." Working paper.

Hansen, Bruce E. 2025. "Standard Errors for Difference-in-Difference Regression." *Journal of Applied Econometrics* 40 (3):291–309. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/jae.3110.

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* New York: Springer, 2nd ed.

Higgins, Julian P. T. and Simon G. Thompson. 2002. "Quantifying Heterogeneity in a Meta-Analysis." *Statistics in Medicine* 21 (11):1539–1558.

Higgins, Julian P. T., Simon G. Thompson, Jonathan J. Deeks, and Douglas G. Altman. 2003. "Measuring Inconsistency in Meta-Analyses." *BMJ* 327 (7414):557–560.

Hoberg, Gerard and Gordon Phillips. 2016. "Text-Based Network Industries and Endogenous Product Differentiation." *Journal of Political Economy* 124:1423–1465.

———. 2025. "Scope, Scale, and Concentration: The 21st-Century Firm." *The Journal of Finance* 80 (1):415–466.

Jabr, Wael, Dominik Gutt, Jurgen Neumann, and Dennis Kundisch. 2020. "The Fate of the App: Economic Implications of Updating under Reputation Resetting." *Working Paper* .

Johnson, Garrett, Tesary Lin, James C. Cooper, and Liang Zhong. 2025. "COPPAcalypse? The YouTube Settlement's Impact on Kids Content." Working paper, SSRN. Available at SSRN: https://ssrn.com/abstract=4430334 or http://dx.doi.org/10.2139/ssrn.4430334.

Kesler, Reinhold. 2023. "The Impact of Apple's App Tracking Transparency on App Monetization." *Working Paper* .

Lam, H. Tai. 2023. "Platform Search Design and Market Power." Working Paper.

Lee, Kwok-Hao and Leon Musolff. 2025. "Two-Sided Markets Shaped By Platform-Guided Search." Working Paper.

Leyden, Benjamin T. 2023. "There's an App for That: Understanding Product Updating Under Digitization." *Working Paper* .

———. 2025. "Platform design and innovation incentives: Evidence from the product rating system on Apple's App Store." *International Journal of Industrial Organization* 99:103133. URL https://www.sciencedirect.com/science/article/pii/S0167718724000882.

Li, Ding and Hsin-Tien Tsai. 2025. "Mobile Apps and Targeted Advertising: Competitive Effects of Data Sharing." *Journal of the European Economic Association* 24 (1):167–198. URL https://doi.org/10.1093/jeea/jvaf025.

MacKinnon, James G., Morten Ørregaard Nielsen, and Matthew D. Webb. 2023. "Cluster-Robust Inference: A Guide to Empirical Practice." *Journal of Econometrics* 232 (2):272–299.

Madsen, Erik and Nikhil Vellodi. 2025. "Insider Imitation." *Journal of Political Economy* 133 (2):652–709. URL https://doi.org/10.1086/732888.

Raval, Devesh. 2023. "Steering in One Click: Platform Self-Preferencing in the Amazon Buy Box." Working Paper.

Reimers, Nils and Iryna Gurevych. 2019. "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks." In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics. URL https://arxiv.org/abs/1908.10084.

Scott Morton, Fiona, Pascal Bouvier, Ariel Ezrachi, Bruno Jullien, Roberta Katz, Gene Kimmelman, A Douglas Melamed, and Jamie Morgenstern. 2019. "Market Structure and Antitrust Subcommittee Report." In *Stigler Committee on Digital Platforms: Final Report*, edited by Luigi Zingales, Guy Rolnik, and Filippo Maria Lancieri. George J. Stigler Center for the Study of the Economy and the State, The University of Chicago Booth School of Business, 23–138. URL https://research.chicagobooth.edu/-/media/research/stigler/pdfs/digital-platforms---committee-report---stigler-center.pdf.

Sun, Liyang and Sarah Abraham. 2021. "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects." *Journal of Econometrics* 225 (2):175–199. URL https://www.sciencedirect.com/science/article/pii/S030440762030378X. Themed Issue: Treatment Effect 1.

Teng, Xuan. 2025. "Self-preferencing, Quality Provision, and Welfare in Mobile Application Markets." Working Paper. URL https://ssrn.com/abstract=4259639.

Warren, Elizabeth. 2019. "Here's how we can break up Big Tech." *Medium* URL https://medium.com/@teamwarren/heres-how-we-can-break-up-big-tech-9ad9e0da324c.

Wen, Wen and Feng Zhu. 2019. "Threat of platform-owner entry and complementor responses: Evidence from the mobile app market." *Strategic Management Journal* 40 (9):1336–1367.

Wooldridge, Jeffrey M. 2021. "Two-Way Fixed Effects, the Two-Way Mundlak Regression, and Difference-in-Differences Estimators." Tech. rep. URL https://ssrn.com/abstract=3906345. SSRN working paper, version August 16, 2021.

Zhu, Feng and Qihong Liu. 2018. "Competing with complementors: An empirical look at Amazon.com." *Strategic Management Journal* 39 (10):2618–2642.

Supplemental Appendix

Leyden (2026): Sherlocking: The Effects of Platform-Owner Entry on the Competitive Behavior of Third-Party Firms

A Data Appendix

This appendix presents detailed market-level summary statistics for the three analysis dimensions in the main text. Table 2 in the main text reports aggregate pre-treatment and post-treatment means; the tables below decompose these by market to illustrate the heterogeneity in market characteristics.

A.1 Incumbent App Summary Statistics

Tables 8 and 9 present app-month level summary statistics for incumbent apps, broken down by market for the pre-treatment and post-treatment periods respectively. These statistics correspond to Panel A of Table 2.

	Price	Price-Paid	Free	IAP	Avg. Rating	Update	Δ Log Ratings	N
Animoji	0.17	1.59	0.89	0.34	3.70	0.13	0.04	846
Apple Music 13.0	1.10	2.99	0.63	0.34	4.00	0.27	0.04	1071
Apple Research	1.58	5.92	0.73	0.23	4.07	0.26	0.05	270
Blood Oxygen	0.84	2.19	0.62	0.31	3.37	0.23	0.08	117
Breathe	0.00		1.00	0.22	4.14	0.07	0.04	261
Camera 11.0	0.00		1.00	0.20	3.91	0.25	0.06	1737
Clock 10.0	0.00		1.00	0.26	4.17	0.18	0.04	2637
Development 11.0	0.00		1.00	0.32	3.78	0.24	0.03	1917
Ecg	2.98	4.20	0.29	0.28	3.76	0.10	0.04	495
Find Items	1.43	4.07	0.65	0.53	3.99	0.29	0.06	432
Health 10.0	0.00		1.00	0.21	3.87	0.19	0.03	252
Health 11.3	0.29	3.39	0.91	0.09	3.56	0.41	0.07	1647
Homekit 10.0	0.00		1.00	0.27	3.58	0.40	0.09	1314
Ibooks 11.4	1.19	3.92	0.70	0.26	3.64	0.30	0.06	522
Magnifier	1.32	2.87	0.54	0.32	4.12	0.12	0.02	531
Measure	3.11	3.92	0.21	0.24	3.92	0.18	0.05	720
Notes 11.0	0.00		1.00	0.32	4.04	0.24	0.04	2988
Radio	1.23	2.86	0.57	0.29	3.90	0.16	0.06	63
Reality Composer	2.67	4.85	0.45	0.37	3.95	0.24	0.05	747
Schoolwork	0.00		1.00	0.33	4.01	0.33	0.06	1143
Sleep	1.56	3.55	0.56	0.48	4.27	0.29	0.05	1566
Translate	1.77	4.70	0.62	0.59	4.25	0.24	0.04	4311
Walkie Talkie	1.54	3.44	0.55	0.33	3.70	0.26	0.05	675
All Markets	0.81	4.01	0.80	0.35	3.98	0.25	0.05	26262

Table 8: Incumbent App Summary Statistics by Market (Pre-Treatment)

	Price	Price-Paid	Free	IAP	Avg. Rating	Update	Δ Log Ratings	N
Animoji	0.85	1.61	0.47	0.37	3.68	0.11	0.02	846
Apple Music 13.0	1.10	3.05	0.64	0.36	3.99	0.25	0.04	1071
Apple Research	1.56	5.87	0.73	0.23	4.10	0.31	0.02	270
Blood Oxygen	0.84	2.19	0.62	0.38	3.45	0.21	0.05	117
Breathe	0.00		1.00	0.22	4.09	0.08	0.02	261
Camera 11.0	1.84	4.12	0.55	0.23	3.92	0.24	0.04	1737
Clock 10.0	0.00		1.00	0.29	4.16	0.17	0.02	2637
Development 11.0	4.12	5.95	0.31	0.32	3.83	0.24	0.03	1917
Ecg	2.95	4.16	0.29	0.31	3.76	0.12	0.03	495
Find Items	1.13	3.62	0.69	0.58	3.99	0.24	0.02	432
Health 10.0	0.00		1.00	0.26	3.88	0.19	0.01	252
Health 11.3	0.49	3.58	0.86	0.09	3.57	0.39	0.06	1647
Homekit 10.0	0.00		1.00	0.30	3.52	0.41	0.04	1314
Ibooks 11.4	1.37	3.97	0.66	0.26	3.73	0.26	0.03	522
Magnifier	1.32	2.85	0.53	0.32	4.13	0.09	0.01	531
Measure	3.08	3.94	0.22	0.26	3.94	0.16	0.04	720
Notes 11.0	2.43	4.33	0.44	0.33	4.06	0.25	0.03	2988
Radio	1.26	2.93	0.57	0.29	4.40	0.16	-0.00	63
Reality Composer	2.66	4.89	0.46	0.37	3.89	0.23	0.03	747
Schoolwork	0.00		1.00	0.36	3.98	0.29	0.03	1143
Sleep	1.73	3.98	0.57	0.51	4.26	0.27	0.03	1566
Translate	1.78	4.78	0.63	0.61	4.24	0.21	0.02	4311
Walkie Talkie	1.53	3.37	0.55	0.36	3.70	0.28	0.03	675
All Markets	1.55	4.30	0.64	0.37	3.98	0.24	0.03	26262

Table 9: Incumbent App Summary Statistics by Market (Post-Treatment)

A.2 Entry and Exit Summary Statistics

Tables 10 and 11 present market-month level entry and exit statistics, broken down by market for the pre-treatment and post-treatment periods respectively. These statistics use the market dynamics sample and correspond to Panel B of Table 2.

	Log Entry	Log Exit	N
Animoji	2.06	1.54	18
Apple Music 13.0	1.18	1.42	18
Apple Research	0.76	0.60	18
Blood Oxygen	0.45	0.31	18
Breathe	1.08	0.22	18
Camera 11.0	2.21	1.52	18
Clock 10.0	2.59	1.52	18
Development 11.0	1.53	1.32	18
Ecg	0.59	0.62	18
Find Items	0.51	0.51	18
Health 10.0	0.61	0.08	18
Health 11.3	1.82	1.33	18
Homekit 10.0	1.83	0.78	18
Ibooks 11.4	0.67	0.84	18
Magnifier	0.57	0.52	18
Measure	1.29	1.43	18
Notes 11.0	2.37	2.16	18
Radio	0.14	0.23	18
Reality Composer	1.37	1.19	18
Schoolwork	1.93	0.92	18
Sleep	1.55	1.52	18
Translate	2.13	2.52	18
Walkie Talkie	1.59	1.71	18
All Markets	1.34	1.08	414

Table 10: Market-Level Entry and Exit by Market (Pre-Treatment)

A.3 Entering Cohort Summary Statistics

Tables 12 and 13 present summary statistics for the entering cohort analysis (repeated cross-section), broken down by market for the pre-treatment and post-treatment periods respectively. These statistics use the market dynamics sample and correspond to Panel C of Table 2.

	Log Entry	Log Exit	N
Animoji	1.83	1.87	18
Apple Music 13.0	0.79	1.18	18
Apple Research	0.65	0.50	18
Blood Oxygen	0.47	0.39	18
Breathe	1.11	0.63	18
Camera 11.0	1.89	1.93	18
Clock 10.0	2.65	1.98	18
Development 11.0	1.11	1.38	18
Ecg	0.54	0.48	18
Find Items	0.55	0.71	18
Health 10.0	0.57	0.50	18
Health 11.3	1.80	1.53	18
Homekit 10.0	2.03	0.97	18
Ibooks 11.4	0.42	1.04	18
Magnifier	0.31	0.67	18
Measure	1.04	1.03	18
Notes 11.0	2.10	2.16	18
Radio	0.29	0.25	18
Reality Composer	0.99	1.24	18
Schoolwork	2.05	1.02	18
Sleep	1.19	1.64	18
Translate	2.00	2.68	18
Walkie Talkie	1.18	1.63	18
All Markets	1.20	1.19	414

Table 11: Market-Level Entry and Exit by Market (Post-Treatment)

	Price	Price-Paid	Free	IAP	Avg. Rating	Log # Ratings	Similarity	N
Animoji	0.32	2.25	0.86	0.08	4.44	0.79	0.64	162
Apple Music 13.0	0.79	3.24	0.76	0.08	4.80	1.72	0.65	49
Apple Research	0.61	2.66	0.77	0.04	2.35	0.77	0.63	26
Blood Oxygen	1.30	4.24	0.69	0.15	4.73	1.73	0.66	13
Breathe	0.00		1.00	0.17	4.22	0.35	0.68	59
Camera 11.0	0.00		1.00	0.10	4.26	0.70	0.69	165
Clock 10.0	0.00		1.00	0.10	4.46	0.58	0.65	245
Development 11.0	0.00		1.00	0.17	3.99	0.31	0.65	100
Ecg	2.63	4.54	0.42	0.11	5.00	0.04	0.67	19
Find Items	1.00	2.24	0.56	0.39	4.60	1.45	0.65	18
Health 10.0	0.00		1.00	0.00	4.84	0.47	0.63	22
Health 11.3	0.09	2.19	0.96	0.03	4.39	0.39	0.65	116
Homekit 10.0	0.00		1.00	0.10	4.29	0.88	0.65	119
Ibooks 11.4	1.08	3.12	0.65	0.04	4.83	0.78	0.63	23
Magnifier	1.16	2.99	0.61	0.28	4.60	2.42	0.66	18
Measure	1.98	2.72	0.27	0.10	3.76	0.58	0.67	63
Notes 11.0	0.00		1.00	0.08	4.33	0.62	0.65	216
Radio	0.25	0.99	0.75	0.00	4.83	3.24	0.64	4
Reality Composer	1.94	4.10	0.53	0.17	4.33	1.10	0.63	59
Schoolwork	0.00		1.00	0.05	4.33	0.53	0.64	140
Sleep	0.96	2.69	0.64	0.32	4.61	1.15	0.65	78
Translate	0.69	3.30	0.79	0.55	4.24	0.96	0.68	172
Walkie Talkie	1.10	2.23	0.51	0.16	3.82	1.48	0.65	79
All Markets	0.39	2.94	0.87	0.15	4.33	0.77	0.65	1965

Table 12: Entering Cohort Summary Statistics by Market (Pre-Treatment)

	Price	Price-Paid	Free	IAP	Avg. Rating	Log # Ratings	Similarity	N
Animoji	0.66	2.19	0.70	0.16	3.94	1.89	0.68	99
Apple Music 13.0	0.37	2.74	0.87	0.20	4.47	1.40	0.66	30
Apple Research	0.78	3.59	0.78	0.00	4.00	1.47	0.64	23
Blood Oxygen	1.12	3.59	0.69	0.12	4.25	1.61	0.70	16
Breathe	0.00		1.00	0.07	5.00	0.09	0.66	50
Camera 11.0	1.41	3.00	0.53	0.10	4.29	0.85	0.67	111
Clock 10.0	0.00		1.00	0.07	4.50	0.59	0.64	257
Development 11.0	2.26	4.43	0.49	0.23	4.36	0.91	0.65	53
Ecg	2.35	4.99	0.53	0.29	3.20	0.49	0.66	17
Find Items	0.78	3.49	0.78	0.44	4.44	1.48	0.66	18
Health 10.0	0.00		1.00	0.11	4.82	0.67	0.62	16
Health 11.3	0.25	2.79	0.91	0.04	4.82	0.50	0.64	111
Homekit 10.0	0.00		1.00	0.08	4.28	1.10	0.64	125
Ibooks 11.4	0.92	2.39	0.62	0.08	4.02	1.02	0.64	13
Magnifier	1.15	2.87	0.60	0.10	3.45	1.63	0.65	10
Measure	1.54	3.57	0.57	0.16	4.22	0.87	0.67	44
Notes 11.0	1.55	3.23	0.52	0.14	4.38	0.75	0.65	159
Radio	0.56	4.49	0.88	0.38	3.93	1.07	0.65	8
Reality Composer	1.19	3.66	0.67	0.30	4.11	0.96	0.65	46
Schoolwork	0.00		1.00	0.10	4.40	0.48	0.64	171
Sleep	0.95	2.80	0.66	0.50	4.48	1.39	0.65	50
Translate	0.74	2.84	0.74	0.45	4.33	1.25	0.67	130
Walkie Talkie	0.95	2.49	0.62	0.19	3.24	1.96	0.65	47
All Markets	0.66	3.16	0.79	0.17	4.27	0.92	0.65	1604

Table 13: Entering Cohort Summary Statistics by Market (Post-Treatment)

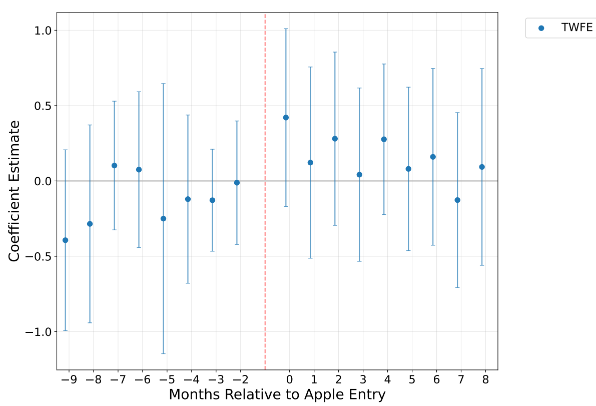
B Cohort Analysis Event Studies

This appendix presents event study estimates for the cohort composition analysis described in Section 3.1.2. I estimate:

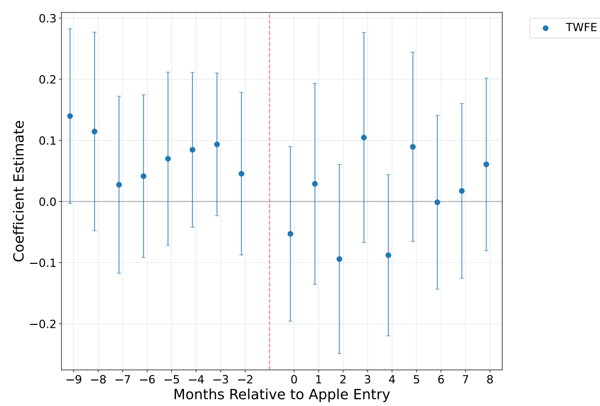
$$y_j = \sum_{k \neq -1} \gamma_k \cdot \mathbf{1}[t_j - t_a^* = k] + \sum_{k \neq -1} \beta_k \cdot \mathbf{1}[t_j - t_a^* = k] \times \text{AS}_j + \delta_{ap} + \eta_t + \varepsilon_j, \tag{9}$$

where t_j is the month app j enters, t_a^* is Apple's entry month in market a, γ_k captures common relative-time effects across platforms, δ_{ap} are market×platform fixed effects, and η_t are calendar month fixed effects. AS_j indicates whether app j is on the App Store. Coefficients β_k show how the App Store–Play Store gap in entry characteristics evolves relative to Apple's entry timing.

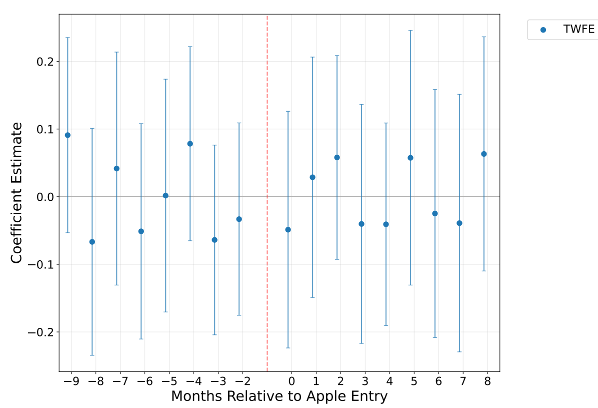
Figure 9: Cohort Event Studies: Monetization Outcomes



(a) Price (All Apps)



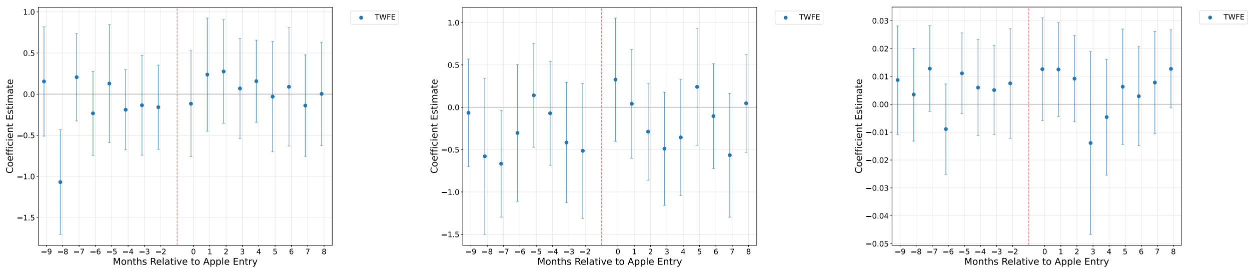
(b) Free



(c) In-App Purchases

Note: Event study estimates from cross-sectional DiD model showing how the App Store–Play Store gap in monetization characteristics evolves relative to Apple's entry timing. Each app is observed once at entry. Standard errors are clustered at the market-platform level. Error bars represent 95% confidence intervals.

Figure 10: Cohort Event Studies: Quality and Location Outcomes



(a) Rating (b) Log Rating Count (c) Similarity

Note: Event study estimates from cross-sectional DiD model showing how the App Store–Play Store gap in quality and location characteristics evolves relative to Apple's entry timing. Each app is observed once at entry. Standard errors are clustered at the market-platform level. Error bars represent 95% confidence intervals.

C Market-Level Heterogeneity

This appendix presents forest plots of market-specific treatment effects for all outcomes across the three margins of analysis. These figures complement the summary diagnostics reported in Table 6 and the selected forest plots presented in the main text (Figure 6).

D Robustness

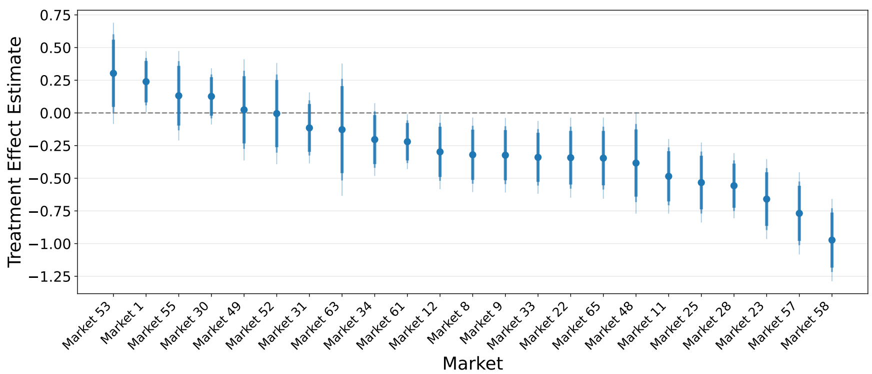
This appendix presents robustness checks for the main analysis: a comparison across DiD estimators for both incumbent and entry/exit results, an alternative fixed effects specification, and alternative clustering of standard errors.

D.1 Comparison Across DiD Estimators

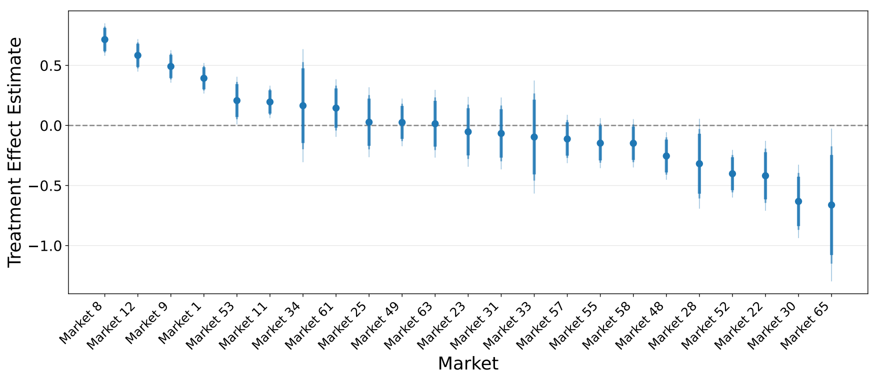
This subsection compares results across multiple difference-in-differences estimators to assess robustness. Alongside the primary TWFE estimator, I report results from the Extended Two-Way Fixed Effects (ETWFE) estimator of Wooldridge (2021) and the interaction-weighted estimator of Sun and Abraham (2021), both of which provide consistent estimates under heterogeneous treatment effects by avoiding the use of already-treated units as controls.

Incumbent Analysis Robustness. Table 14 presents estimates across all three estimators for incumbent app outcomes. Results are qualitatively similar across methods—all estimators agree on the direction and significance of effects. ETWFE and Sun-Abraham tend to produce somewhat larger point estimates for monetization outcomes. TWFE and SA use the primary app and market×month fixed effects specification. ETWFE uses app and month fixed effects. In this implementation, each market constitutes its own cohort, so the saturated market×time treatment interactions absorb market-specific time variation, making separate market×month fixed effects re-

Figure 11: Market Heterogeneity in Entry and Exit



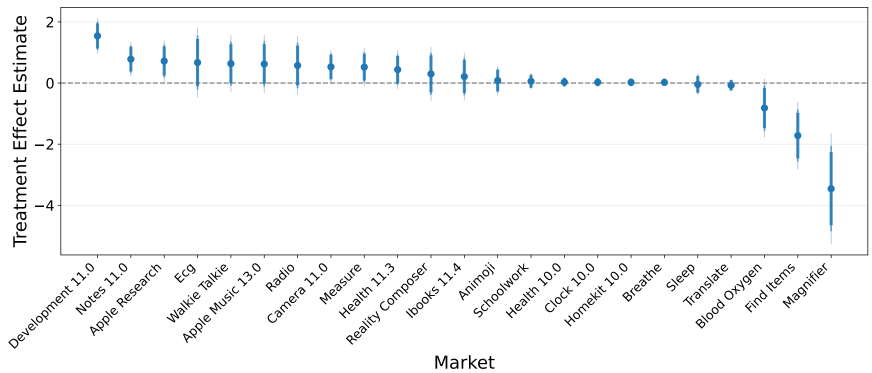
(a) Log Entry Count



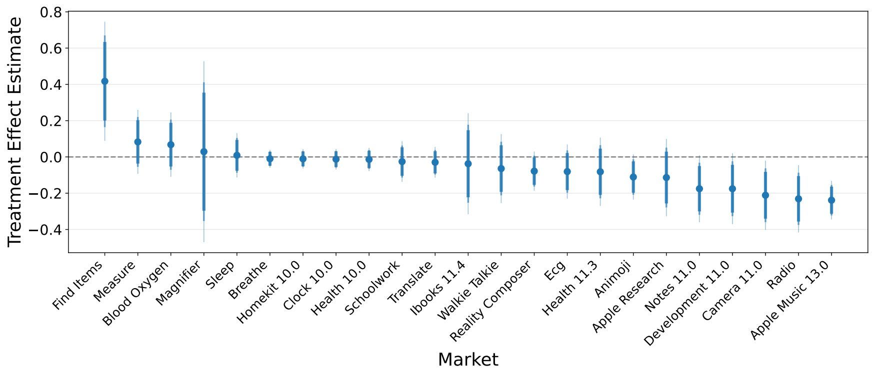
(b) Log Exit Count

Note: Market-specific treatment effects from the entry/exit model, Equation (3). Analysis uses the market dynamics sample at the market-platform-month level. Standard errors are clustered at the market-platform level. Points show estimates with 90%, 95%, and 99% confidence intervals (thick to thin lines).

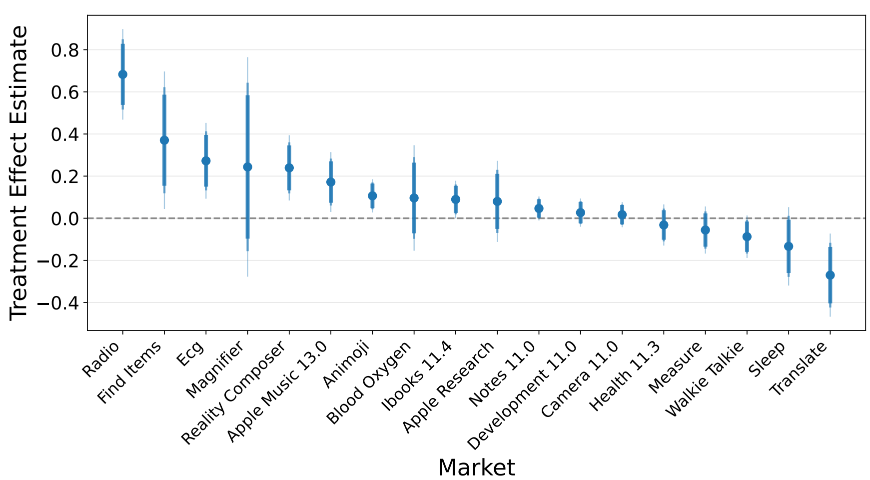
Figure 12: Market Heterogeneity in Cohort Monetization



(a) Price



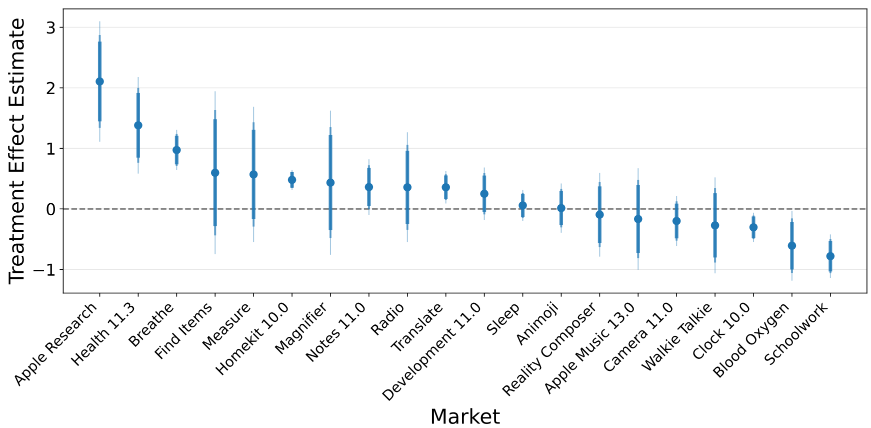
(b) Free



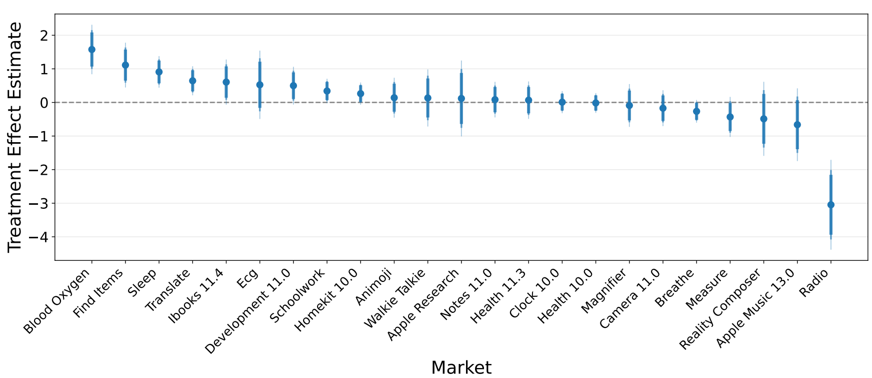
(c) In-App Purchases

Note: Market-specific treatment effects from the cohort composition model, Equation (4). Analysis uses the market dynamics sample. Each app is observed once at entry. Standard errors are clustered at the market-platform level. Points show estimates with 90%, 95%, and 99% confidence intervals (thick to thin lines).

Figure 13: Market Heterogeneity in Cohort Quality



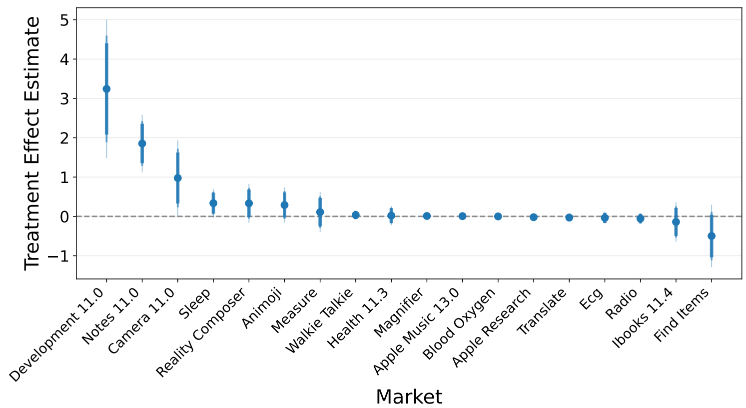
(a) Avg. Rating



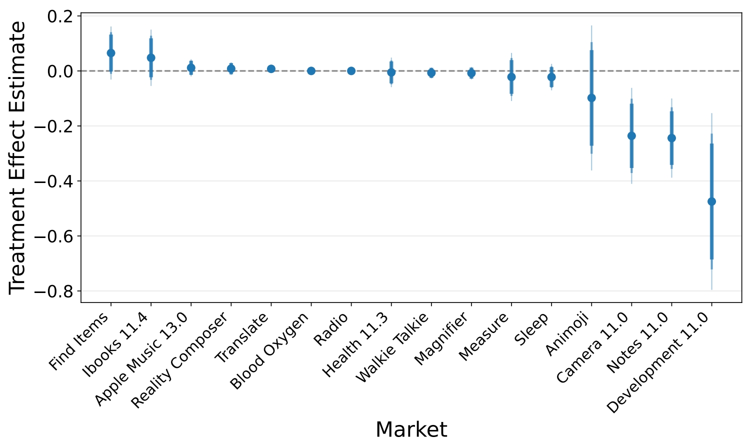
(b) Log Rating Count

Note: Market-specific treatment effects from the cohort composition model, Equation (4). Analysis uses the market dynamics sample. Each app is observed once at entry. Standard errors are clustered at the market-platform level. Points show estimates with 90%, 95%, and 99% confidence intervals (thick to thin lines).

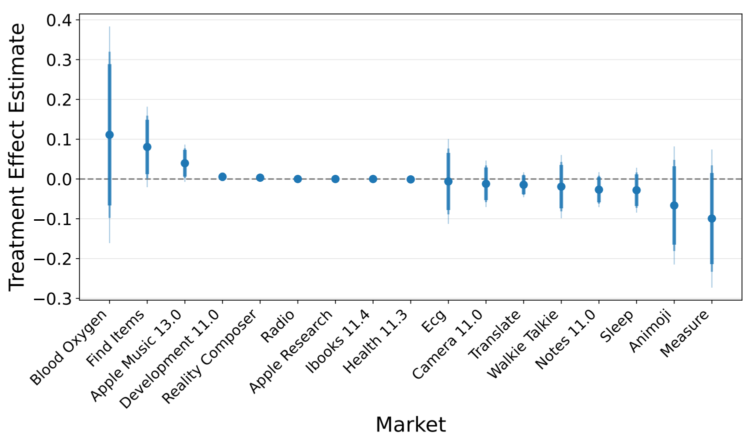
Figure 14: Incumbent Market Heterogeneity in Monetization



(a) Price (All Apps)



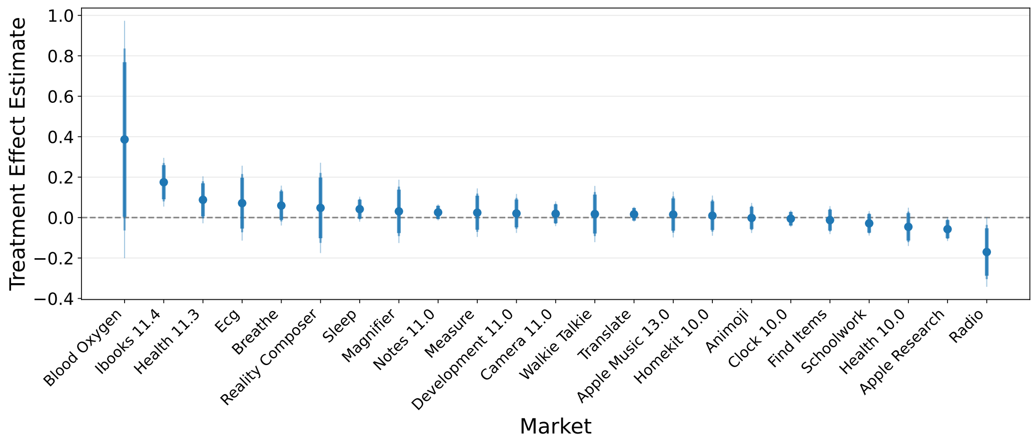
(b) Free



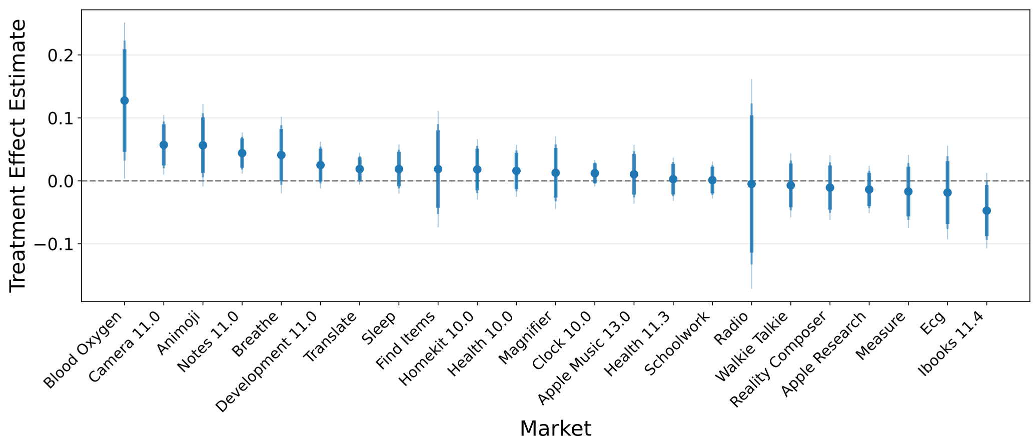
(c) In-App Purchases

Note: Market-specific treatment effects from the incumbent analysis model, Equation (5). Analysis is at the app-month level. For the IAP outcome, 5 markets are excluded due to insufficient Android in-app purchase data in the pre-period. Markets with insufficient outcome variation are omitted. Standard errors are clustered at the app level. Points show estimates with 90%, 95%, and 99% confidence intervals (thick to thin lines).

Figure 15: Incumbent Market Heterogeneity in Quality



(a) Rating



(b) Rating Count Growth

Note: Market-specific treatment effects from the incumbent analysis model, Equation (5). Analysis is at the app-month level. Markets with insufficient outcome variation are omitted. Standard errors are clustered at the app level. Points show estimates with 90%, 95%, and 99% confidence intervals (thick to thin lines).

Table 14: DiD Point Estimates

Outcome	(1) TWFE	(2) ETWFE	(3) SA
Δ Log Rating Count	0.0191***	0.0106***	0.0093**
	(0.0038)	(0.0028)	(0.0040)
Free App	-0.0559***	-0.1022***	-0.0725***
	(0.0097)	(0.0070)	(0.0114)
In-App Purchases	-0.0125**	-0.0071***	-0.0080*
	(0.0061)	(0.0025)	(0.0046)
Price	0.3687***	0.6882***	0.4765***
	(0.0513)	(0.0437)	(0.0609)
Price (Paid Apps)	0.0787	—	0.0504
	(0.0493)	—	(0.0447)
Avg. Rating	0.0227***	0.0161***	0.0082
	(0.0072)	(0.0045)	(0.0054)
Update	0.0143*	0.0008	0.0169
	(0.0084)	(0.0104)	(0.0152)

Notes: Analysis is at the app-month level. Columns show treatment effect estimates from TWFE (two-way fixed effects), ETWFE (extended two-way fixed effects; Wooldridge (2021)), and SA (interaction-weighted estimator; Sun and Abraham (2021)). For the IAP outcome, 5 markets are excluded due to insufficient Android in-app purchase data in the pre-period. Standard errors clustered at the app level are in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

dundant. The ETWFE estimate for Price (Paid Apps) is excluded due to insufficient within-market variation in paid app prices, which prevents reliable estimation of market-specific treatment effects.

Entry/Exit Analysis Robustness. Table 15 presents market-level entry and exit results across estimators using the same market dynamics sample as the main text (Table 3). Entry effects are negative across all estimators, consistent with the entry deterrence documented in the main text. Exit effects are small and imprecisely estimated, consistent with the null average effect reported in the main text. Results are qualitatively similar across methods—all estimators agree on the direction and significance of effects. ETWFE and Sun-Abraham tend to produce somewhat larger point estimates for monetization outcomes.

D.2 Alternative Fixed Effects Specification

The main text uses app and market×month fixed effects ($\gamma_j + \eta_{a(j),t}$) to absorb market-specific time shocks. This subsection presents results using the simpler app and month fixed effects ($\gamma_j + \eta_t$) as a robustness check. Results are directionally consistent with the main specification, but point

Table 15: DiD Point Estimates

Outcome	(1) TWFE	(2) ETWFE	(3) SA
Log Entry Count	-0.2653***	-0.1158	-0.1158
	(0.0714)	(0.4063)	(0.0985)
Log Exit Count	-0.0209	0.1491	0.1491
	(0.0746)	(0.4056)	(0.1070)

Notes: Analysis is at the platform-market-month level. Columns show treatment effect estimates from TWFE (two-way fixed effects), ETWFE (extended two-way fixed effects; Wooldridge (2021)), and SA (interaction-weighted estimator; Sun and Abraham (2021)). Standard errors clustered at the platform-market level are in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

estimates tend to be larger in magnitude and more precisely estimated. This is consistent with market×month fixed effects being the more conservative specification, as they absorb market-specific time variation that the simpler month fixed effects leave in the error term.

Monetization Outcomes

Table 16: Incumbent App Monetization Outcomes

	(1) Price	(2) Price (Paid Apps)	(3) Free App	(4) In-App Purchases
ATT	0.5474***	0.0588*	-0.0797***	-0.0102**
	(0.0470)	(0.0350)	(0.0081)	(0.0045)
Baseline Mean	0.8069	4.014	0.7990	0.3626
N	52,524	14,741	52,524	41,310

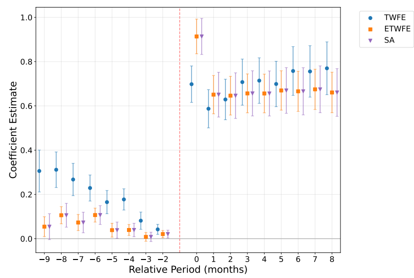
Notes: Estimates from the incumbent analysis model at the app-month level. For the IAP outcome, 5 markets are excluded due to insufficient Android in-app purchase data in the pre-period. Standard errors clustered at the app level are in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

Quality Outcomes

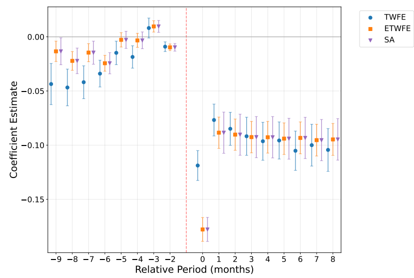
D.3 Clustering Robustness

All incumbent app-level regressions cluster standard errors at the app level. This choice addresses within-app serial correlation (Bertrand, Duflo, and Mullainathan, 2004) and provides reliable asymptotic inference given the large number of clusters (thousands of apps). Because the analysis covers the full population of apps on both platforms—rather than a sample drawn from a larger population—the design-based motivation for coarser clustering is weaker than in sampled settings (Abadie et al., 2023).

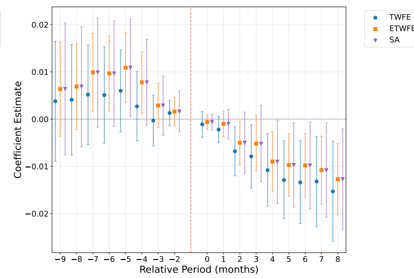
Figure 16: Monetization Dynamics (App + Month FE)



(a) Price (All Apps)



(b) Free



(c) In-App Purchases

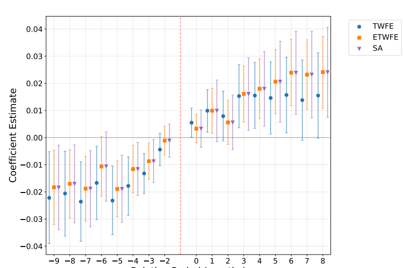
Note: Event study estimates using app and month fixed effects. TWFE, ETWFE (Wooldridge (2021)), and SA (Sun and Abraham (2021)) are shown. Analysis is at the app-month level. For the IAP outcome, 5 markets are excluded due to insufficient Play Store in-app purchase data in the pre-period. Standard errors are clustered at the app level. Error bars represent 95% confidence intervals.

Table 17: Incumbent App Quality Outcomes

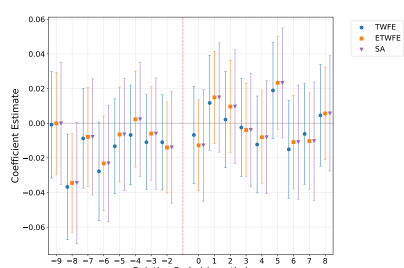
	(1) Avg. Rating	(2) Update	(3) Δ Log Rating Count
ATT	0.0251***	0.0118	0.0163***
	(0.0059)	(0.0073)	(0.0034)
Baseline Mean	3.978	0.2515	0.0468
N	50,099	52,524	52,524

Notes: Estimates from the incumbent analysis model at the app-month level. Standard errors clustered at the app level are in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

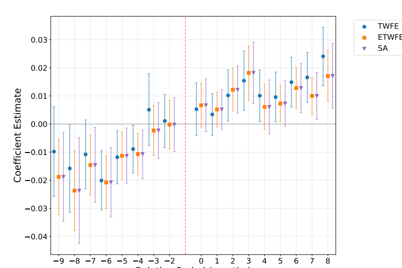
Figure 17: Quality Dynamics (App + Month FE)



(a) Rating



(b) Update Frequency



(c) Rating Count Growth

Note: Event study estimates using app and month fixed effects. TWFE, ETWFE (Wooldridge (2021)), and SA (Sun and Abraham (2021)) are shown. Analysis is at the app-month level. Standard errors are clustered at the app level. Error bars represent 95% confidence intervals.

Nevertheless, since treatment is assigned at the market level (with 23 affected markets, yielding approximately 46 market-platform cells), one may be concerned about within-market residual correlation among apps. Standard market-level cluster-robust variance estimators (CRV1) are known to be downward biased with so few clusters (Hansen, 2025; MacKinnon, Nielsen, and Webb, 2023). Instead, I report CRV3J (cluster jackknife) standard errors clustered at the market-platform level, which are never downward biased by construction.

Table 18: Clustering Robustness: App-Level vs. Market-Platform CRV3J Standard Errors

	ATT	App-Level SE	CRV3J SE
Price	0.369	(0.051)***	(0.324)
Price (Paid Apps)	0.079	(0.049)	(0.108)
Free	-0.056	(0.0097)***	(0.048)
In-App Purchases	-0.013	(0.0061)**	(0.0090)
Rating	0.023	(0.0072)***	(0.011)**
Update	0.014	(0.0084)*	(0.011)
Δ Log Rating Count	0.019	(0.0038)***	(0.0074)***

Notes: Column "ATT" reports the average treatment effect on the treated from the TWFE specification (Equation (5)). Column "App-Level SE" reports standard errors clustered at the app level (primary specification). Column "CRV3J SE" reports cluster-jackknife standard errors (CRV3J) clustered at the market-platform level, which guard against downward bias with few clusters (MacKinnon, Nielsen, and Webb, 2023; Hansen, 2025). Significance stars on each SE column use the respective inference. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 18 presents the comparison. For each outcome, I report the ATT estimate alongside both the primary app-level standard errors and the CRV3J standard errors. The CRV3J standard errors are larger for most outcomes, as expected given that only 46 market-platform clusters underlie the variance estimate. For quality outcomes (rating and rating count growth), which display broadly distributed effects across markets, significance is maintained under CRV3J inference. For monetization outcomes such as price and free, the CRV3J standard errors are substantially larger and these average effects are no longer statistically significant at conventional levels. This reflects both the small number of market-platform clusters and the concentration of pricing effects in a subset of markets—consistent with the market heterogeneity analysis in Section 3.2.

E Single-Homing Sample Analysis

A potential concern with the cross-platform identification strategy is that multi-homing apps—those available on both iOS and Android—may transmit treatment effects across platforms. If an App Store developer adjusts its strategy in response to Apple's entry, the Play Store version of the same app may also change, contaminating the control group. This appendix restricts the incumbent sample to single-homing apps (those available on only one platform) and shows that the main findings are not driven by such spillovers.

E.1 Cross-Platform Record Linkage

Since no common identifier links apps across the App Store and Play Store, I identify multi-homing apps through probabilistic record linkage using the `recordlinkage` Python library. I generate candidate pairs via a Sorted Neighbourhood indexer on app names (window size 21), then compute Jaro-Winkler string similarity on three fields: app name, developer name, and developer URL. A pair is classified as a match if at least two of the three fields exceed a similarity threshold of 0.85. Apps with no cross-platform match are classified as single-homing.

E.2 Sample Characteristics

Table 19 presents summary statistics for the single-homing sample. Restricting to single-homing apps reduces the incumbent sample from 26,262 to 15,264 unique apps per period (a 42% reduction). Panels B and C are unchanged, as market-level entry/exit and cohort composition analyses do not condition on individual app homing status.

E.3 Incumbent App Results

Table 20 compares TWFE ATT estimates from the full sample and the single-homing sample. All seven outcomes have the same sign in both samples, and five of seven retain statistical significance despite the smaller sample. Monetization effects are somewhat larger in the single-homing sample—the price increase rises from \$0.37 to \$0.41, and the decline in the share of free apps from 5.6 to 6.3 percentage points—consistent with spillovers modestly attenuating the full-sample estimates. Rating remains significant at the 10% level in the single-homing sample, with a slightly smaller point estimate (0.017 vs. 0.023). Update loses significance due to both a smaller point estimate (0.005 vs. 0.014) and wider standard errors from the reduced sample. The qualitative conclusions are unchanged.

Figures 18 and 19 present event studies for the single-homing sample. Pre-trends are flat across outcomes, and the post-entry dynamics are consistent with the full-sample event studies in the main text.

E.4 Cross-Market Heterogeneity

Table 21 reports market-level heterogeneity diagnostics for the single-homing sample. The I^2 statistics are comparable to the full-sample results in Section 3.2. Sign agreement is somewhat lower for monetization outcomes, reflecting noisier market-level estimates from smaller per-market samples.

Table 19: Summary Statistics

	Pre-Treatment	Post-Treatment	Change
Panel A: Incumbent Apps			
Price	0.832	1.725	+0.893
Price (Paid Apps)	3.805	4.140	+0.335
Free	0.781	0.583	-0.198
In-App Purchases	0.357	0.377	+0.020
Avg. Rating	3.990	3.990	+0.000
Update	0.214	0.202	-0.012
Δ Log # Ratings	0.047	0.028	-0.018
N	15,264	15,264	
Panel B: Market-Level Entry and Exit			
Log Entry	1.340	1.198	-0.142
Log Exit	1.079	1.191	+0.113
N	414	414	
Panel C: Entering Cohorts			
Price	0.392	0.658	+0.266
Price (Paid Apps)	2.938	3.160	+0.221
Free	0.867	0.792	-0.075
In-App Purchases	0.153	0.173	+0.019
Avg. Rating	4.329	4.269	-0.060
Log # Ratings	0.767	0.918	+0.152
Similarity	0.655	0.653	-0.002
N	1,965	1,604	

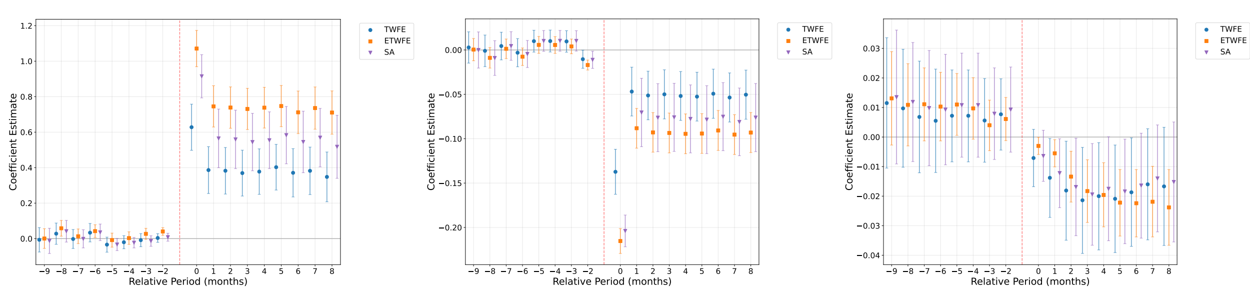
Notes: Panel A reports app-month level statistics for incumbent apps observed in an 18-month balanced panel around each entry event. Panel B reports market-month level log entry and exit counts. Panel C reports statistics for apps observed once upon entry (repeated cross-section). Pre-Treatment refers to the 9 months before Apple's entry; Post-Treatment refers to the 9 months after.

Table 20: Incumbent Treatment Effects: All Apps vs. Single-Homing Apps

Outcome	All Apps (1)	Single-Homing (2)
Monetization		
Price	0.369***	0.406***
	(0.051)	(0.067)
Free App	-0.056***	-0.063***
	(0.010)	(0.014)
In-App Purchases	-0.013**	-0.024**
	(0.006)	(0.011)
Price (Paid Apps)	0.079	0.069
	(0.049)	(0.081)
Quality		
Avg. Rating	0.023***	0.017*
	(0.007)	(0.010)
Update	0.014*	0.005
	(0.008)	(0.012)
Δ Log Rating Count	0.019***	0.026***
	(0.004)	(0.006)
Observations	52,524	30,528

Notes: This table compares TWFE ATT estimates using all apps (column 1) versus only single-homing apps (column 2). Single-homing apps are those available on only one platform, excluding apps identified as operating on both iOS and Android via record linkage. All specifications include app and market×month fixed effects with standard errors clustered at the app level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

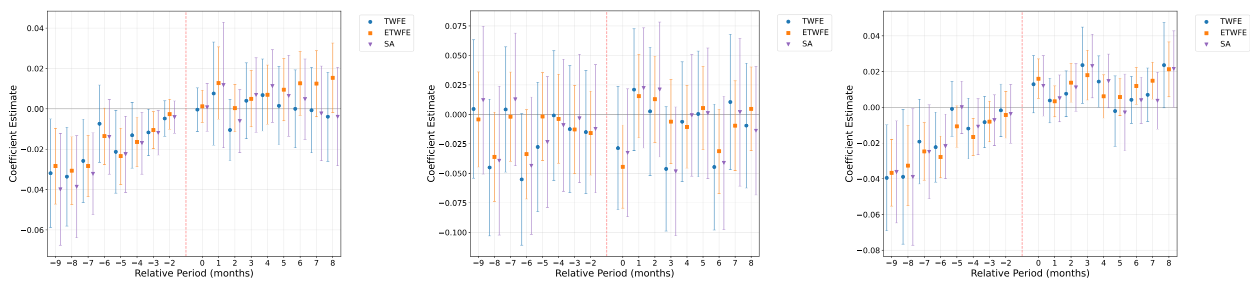
Figure 18: Single-Homing Sample: Monetization Dynamics



(a) Price (All Apps) (b) Free (c) In-App Purchases

Note: Event study estimates for monetization outcomes in the single-homing sample. TWFE, ETWFE (Wooldridge (2021)), and SA (Sun and Abraham (2021)) are shown. All specifications use app and market×month fixed effects (except ETWFE, which uses app and month fixed effects). For the IAP outcome, 5 markets are excluded due to insufficient Play Store in-app purchase data in the pre-period. Standard errors are clustered at the app level. Error bars represent 95% confidence intervals.

Figure 19: Single-Homing Sample: Quality Dynamics



(a) Rating (b) Update (c) Δ Log Rating Count

Note: Event study estimates for quality outcomes in the single-homing sample. TWFE, ETWFE (Wooldridge (2021)), and SA (Sun and Abraham (2021)) are shown. All specifications use app and market×month fixed effects (except ETWFE, which uses app and month fixed effects). Standard errors are clustered at the app level. Error bars represent 95% confidence intervals.

Table 21: Cross-Market Heterogeneity in Incumbent Treatment Effects

Outcome	Est. (SE)	Sign Agree.	% Null	I^2 (Q p)	$\sqrt{\tau^2}$
Monetization					
Price	0.406	56%	75%	85%	0.110
	(0.067)			(<0.001)	
Free	-0.063	50%	79%	72%	0.024
	(0.014)			(<0.001)	
In-App Purchases	-0.024	54%	100%	50%	0.009
	(0.011)			(0.020)	
Price (Paid Apps)	0.069	64%	100%	0%	0.000
	(0.081)			(0.591)	
Quality					
Rating	0.017	67%	76%	66%	0.042
	(0.010)			(<0.001)	
Update Frequency	0.005	57%	90%	22%	0.025
	(0.012)			(0.180)	
Δ Log Rating Count	0.026	76%	76%	55%	0.026
	(0.006)			(0.001)	

Notes: This table reports heterogeneity diagnostics applied to market-level ATTs. Est. is the TWFE estimate. Sign Agreement is the share of markets whose point estimate matches the sign of the overall estimate. % Null is the share of markets where the 95% confidence interval includes zero. I^2 measures the proportion of total variation due to between-market heterogeneity rather than sampling error; Q p-value tests the null of homogeneous effects. $\sqrt{\tau^2}$ is the DerSimonian–Laird estimate of the between-market standard deviation of true effects.