# Scale, Concentration, and Entry Timing in the Shopify App Ecosystem

Fabrizio Assabese
Judge.me LTD
London, UK
fabrizio@judge.me

Giuseppe Destefanis
University College London
London, UK
g.destefanis@ucl.ac.uk

## Abstract

The Shopify marketplace hosts more than 24,000 third-party applications used by 2.7 million merchant stores generating $706 billion in annual sales. Despite its economic scale, little empirical evidence exists on its structure and adoption drivers. Using Store Leads data, we analysed 4,213 active applications and 55 functional categories. Concentration varied: over half of categories were low-concentration, while 20% were highly concentrated. Larger categories consistently showed lower concentration, challenging assumptions of winner-take-all dynamics. Growth analysis revealed early mover decline rather than lasting advantage. In 88% of categories, later entrants grew faster, while established apps lost installations despite long market presence. This pattern held across diverse functions over more than a decade, pointing to merchant shifts towards newer applications. Opportunities for new apps therefore remain open, provided category selection is favourable, with competitive categories offering better prospects than monopolised ones.

## Keywords

Shopify, Apps, Marketplace

**ACM Reference Format:**
Fabrizio Assabese and Giuseppe Destefanis. Rio de Janiero, April 2026. Scale, Concentration, and Entry Timing in the Shopify App Ecosystem. In *Proceedings of International Workshop on Software-intensive Business (IWSiB'26)*. ACM, New York, NY, USA, 8 pages.

## 1 Introduction

Shopify[1] is a platform that enables businesses to create and operate online stores. It provides merchants with the technical foundation to sell products, handle payments, and manage orders, with subscription plans ranging from small independent shops to large international brands. A distinctive aspect of Shopify's model is that many advanced functions are not built into the core product but supplied through an external marketplace of applications. These applications, developed by third-party companies, extend store capabilities with features such as product reviews, marketing automation, logistics, subscription management, and integration with social media or advertising platforms.

This marketplace now comprises more than 24,000 applications maintained by thousands of independent software-producing organizations. For merchants, it is central to how storefronts are designed and how they compete in online retail. For developers, it offers access to a global market of potential customers. Several applications that began as small projects have grown into significant companies, showing that the marketplace operates both as infrastructure for commerce and as a channel for software entrepreneurship. Developers considering whether to build for Shopify need to know which categories are saturated, whether early entrants retain growth advantages, and how competition evolves. Merchants must evaluate the risks of relying on dominant providers in essential categories whilst assessing whether newer alternatives provide genuine improvements. For Shopify itself, adoption patterns across applications affect openness, competition, and long-term sustainability. These questions can now be studied systematically using Store Leads[2] data. Store Leads scrapes Shopify storefronts on a weekly basis, providing datasets with information on both domains (including categories, estimated revenue, and installed applications) and applications (including installs, creation date, ratings, and vendor details). These sources make it possible to analyse adoption dynamics, market concentration, and entry timing at scale. In this study, we use them to investigate how the Shopify marketplace operates as an ecosystem and what this reveals about opportunities and constraints for software-producing organizations.

We therefore investigate the following research questions:
**RQ1:** What are the dynamics and scale of the Shopify app ecosystem? *Rationale:* Establishing the size of the ecosystem, its temporal evolution, the distribution of applications across categories, geographic concentration, and adoption patterns provides the context for subsequent analyses. Understanding ecosystem growth and merchant deployment patterns frames the competitive environment that developers enter and clarifies whether building for Shopify represents a viable business direction.
**RQ2:** How concentrated is the app ecosystem within specific categories? *Rationale:* Categories vary widely in scale and importance. Some attract intense developer investment and competition, while others are dominated by only a few providers. Assessing whether adoption is concentrated or distributed is essential for evaluating competitive dynamics and identifying where opportunities for new applications remain.
**RQ3:** Do early entrants maintain growth advantages over time? *Rationale:* Although cumulative installation counts favour older applications through longer exposure, recent growth patterns reveal whether early movers sustain momentum or lose ground to newer entrants. By analysing both absolute and proportional growth, we test whether entry timing confers lasting advantage or whether markets remain contestable as categories mature.
This paper is structured as follows. Section 2 reviews related work

[1] https://www.shopify.com/

[2] https://storeleads.app

*IWSiB'26, 9th International Workshop on Software-intensive Business (IWSiB 2026)*
Rio de Janiero, April 2026.

on platform ecosystems, complementor dynamics, and market structure. Section 3 describes the Store Leads data and our analytical sample. Section 4 presents the methods used to address each research question. Section 5 reports findings on ecosystem scale (RQ1), concentration within categories (RQ2), and early mover growth dynamics (RQ3). Section 6 discusses threats to validity. Section 7 concludes our study.

## 2 Related Works

Platform ecosystems depend on complementors who develop applications on top of the platform infrastructure [3]. Early empirical work on mobile app stores documented rapid complementor proliferation and innovation patterns [2], establishing foundational questions about openness, governance, and value appropriation in two-sided markets. Subsequent research has examined complementor dynamics through three primary lenses: platform owner entry effects, governance mechanisms, and market structure.

A body of work analyses how platform owner entry into complementor spaces affects competition and innovation. Studying Android, Wen and Zhu [15] found that threats of Google entry reduced complementor innovation whilst raising prices, whereas Foerderer et al. [5] documented attention spillovers from Google Photos entry that increased innovation in affected categories. Similar dynamics emerge in retail marketplaces, where Amazon's entry into third-party product categories reduces small-seller growth [16] and triggers offline disintermediation strategies [7]. Platform governance through awards [6], endorsements [1], and venture capital signals [14] shapes complementor entry decisions and resource allocation. Whilst these studies establish how platforms influence complementor behaviour, they do not systematically measure adoption distributions or concentration across functional categories.

Research on market structure in platform ecosystems remains limited. One longitudinal study of mobile app usage documented Pareto-like distributions and intra-category competitive elimination [10], but systematic category-level concentration measurement using formal metrics across revenue or install bases is scarce. Quantitative ecosystem evolution studies exist for open-source platforms like R [11], but comparable large-scale analyses of commercial B2B marketplaces remain absent. Studies of multihoming in video games [4] and cross-platform responses to entry [9] examine developer strategies but not aggregate market structure.

Enterprise and B2B platform ecosystems have received less empirical attention than consumer app stores. Prior work has examined complementor entry decisions in enterprise software platforms [8], venture capital signaling effects in Salesforce [14], and innovation responses to antitrust intervention in Microsoft ecosystems [12]. However, large-scale quantitative studies of B2B marketplace structure and post-entry growth dynamics, comparable to iOS/Android research, remain absent.

Our study provides a platform-wide empirical analysis of the Shopify ecosystem, a major B2B e-commerce platform serving 2.7 million merchant stores. We contribute systematic category-level concentration measurement across 55 functional categories using HHI and top-5 market share metrics, document substantial heterogeneity in competitive structure (challenging universal winner-take-all assumptions), and establish that late entrants systematically outperform early movers in recent growth across 88% of categories, with early movers experiencing net merchant churn whilst late movers capture growth. We focus on supply-side market structure in a B2B context, providing quantitative evidence on concentration dynamics and entry timing effects in an underexamined but economically significant platform ecosystem.

## 3 Dataset

We use data from Store Leads[3], a commercial provider that performs weekly crawls of Shopify storefronts and the Shopify app store. Our dataset, collected on 28 September 2025, consists of two files: an application-level export containing 24,826 Shopify apps and a domain-level export containing 2,701,805 Shopify stores.

The application file includes: unique identifier, name, creation date, current status (active or removed), functional categories (colon-separated labels such as product reviews, marketing, SEO), pricing (minimum and maximum monthly subscription in USD), adoption metrics (total installs, installs in the last 30 and 90 days), and quality signals (average rating, total reviews, recent review counts). Vendor fields cover name, email, website, and registered address. The domain file includes merchant name, store creation date, geographic location (city, country code), subscription plan, estimated yearly sales in USD (derived from traffic models), average product price, installed app count, and a colon-separated list of installed app names.

Install counts represent the cumulative number of stores that installed each application, as detected by Store Leads through storefront scraping. Merchants typically remove inactive application code from their storefronts to maintain site performance, suggesting that detected installations predominantly represent actively used apps rather than abandoned integrations. Store Leads tracks both cumulative installations and recent changes (installations in the last 30 and 90 days), with the latter capturing net changes including removals. Our analyses use cumulative installations to measure total market penetration (RQ1, RQ2) and recent installation metrics to assess growth dynamics (RQ3).

Dates were provided in string format and converted to date objects. Application categories and store tags were standardised through case-insensitive matching.

Our analytical sample consists of active applications (status = Active) with at least one installation. This criterion excludes abandoned projects, pre-launch listings, and removed apps whilst retaining niche applications serving small merchant groups. Of the 24,826 applications, 16,698 were active, and 4,213 of these had at least one installation. This filtered sample forms the basis of all subsequent analyses.

## 4 Methodology

We analyse the Shopify app ecosystem in three stages aligned with our research questions. First, we establish ecosystem dynamics and scale using descriptive statistics (RQ1). Second, we measure market concentration within functional categories using the Herfindahl–Hirschman Index and top N market shares (RQ2). Third, we test whether early entrants retain growth advantages by relating application creation timing to recent growth metrics rather than to

[3] https://storeleads.app

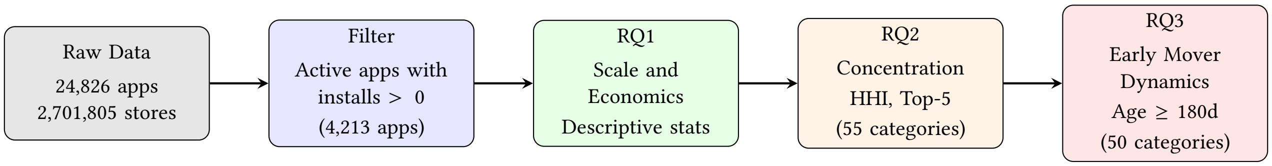


**Figure 1: Methodological workflow showing sequential analytical stages from data collection to early mover growth dynamics analysis.**

cumulative share (RQ3). Figure 1 summarises the workflow.

***Market share definition.*** For RQ2 and for descriptive context in RQ1, we compute each application's market share within its primary category as

$$s_{i,c} = \frac{installs_i}{\sum_{j=1}^{n_c} installs_j},$$

where $installs_i$ is the cumulative installation count of application $i$ in category $c$ with $n_c$ applications.

***RQ1: Dynamics and scale.*** We report application counts by operational status and market participation, derive a single primary category per app from the first label in `app_store_categories`, and describe category size distributions. We characterise the merchant side using estimated yearly sales, installed app counts per store, creation dates, geographic codes, and pricing models classified as freemium when `min_price` equals $0 and paid otherwise. We identify 231 primary categories and, for later stages, the subset with sufficient competition.

***RQ2: Concentration within categories.*** We restrict to categories with at least 20 active applications that have installs greater than zero. This focuses on established markets and yields more stable concentration estimates. For each category, we compute

$$\text{HHI}_c = \sum_{i=1}^{n_c} s_{i,c}^2,$$

and classify concentration using: HHI below 0.15 low, HHI from 0.15 to below 0.25 moderate, HHI at least 0.25 high [13]. As a complement, we report the cumulative share of the top five applications. We examine how concentration relates to category size and total installations.

***RQ3: Early mover growth dynamics.*** We test whether earlier entry is associated with stronger recent growth. We use the app creation date as a proxy for entry timing and analyse only applications aged at least 180 days to avoid launch volatility. Growth is decomposed into two metrics per app: velocity, the number of installs in the last 90 days, and rate, the ratio of installs in the last 90 days to cumulative installs. Within each category we rank apps by creation date, then compute Spearman correlations between creation rank and each growth metric. Negative correlations for velocity indicate earlier apps adding more installs in absolute terms. Negative correlations for rate indicate earlier apps growing faster proportionally. Positive signs indicate advantages for later entrants. Categories are labelled as Strong FMA when both correlations are negative, FMA Eroding when velocity is negative and rate is positive, Late Mover Advantage when both are positive, and Mixed or none when signs conflict or effects are near zero. Statistical significance is assessed at $p < 0.05$ with Benjamini–Hochberg false discovery rate control across categories.

***Acceleration analysis.*** To detect shifts in momentum, we compare 30 day installs with the 90 day average divided by three. Apps with higher 30 day value are marked as accelerating. We define early movers as the first quartile by creation date and late movers as the fourth quartile. Within each category we test differences in acceleration between these groups using the Mann–Whitney U test with false discovery rate control and report Cohen's $d$ as an effect size.

***Scenario projections.*** To contextualise observed gaps, we project median early and late mover baselines under three simple scenarios for 20 quarters and report 24 month outcomes as a practical horizon. The scenarios are exponential growth with constant rate, linear growth with constant absolute velocity, and a decay model with rates declining linearly to zero over the horizon. We also report time to parity, defined as the number of quarters for late movers to reach 50% of early mover installs under each scenario. These are illustrative calculations under stated assumptions rather than forecasts.

***Inclusion criteria.*** All inferential analyses use mature applications aged at least 180 days and categories with at least 20 active apps with installs greater than zero. Early and late mover groups are defined as the first and fourth quartiles by creation date within each category.

## 5 Results

### 5.1 RQ1: Dynamics and Scale of the Shopify App Ecosystem

*Ecosystem Scale.* The dataset contained 24,826 applications and 2,701,805 Shopify stores. Of the total applications, 16,698 (67.3%) were marked as Active, indicating availability in the Shopify App Store at the time of data collection. The remaining 8,128 applications (32.7%) had been removed or delisted. However, the majority of active applications showed no recorded installations. Only 4,213 active applications (17.0% of the total) had recorded at least one installation, whilst 12,485 active applications (74.8% of active apps) showed zero installs. This pattern likely reflects both genuine difficulty in achieving market entry and measurement limitations inherent in web scraping methodologies, which cannot detect installations when Store Leads fails to identify the application script in merchant storefronts.

Applications were distributed across 231 primary functional categories. We identified 55 categories containing 100 or more applications and 126 categories containing 20 or more applications. The largest category, analytics, contained 655 applications, followed by shipping (591), discounts (561), chat (500), and dropshipping (447). The median category size was 27 applications, with the 25th percentile at 3 applications and the 75th percentile at 97 applications. Approximately 25% of categories contained fewer than 10 applications.

**Economic Significance** Stores collectively generated an estimated $706.3 billion in annual sales. These figures derive from traffic models rather than reported financials and represent merchant e-commerce activity, not app subscription revenue (which cannot be estimated due to unavailable pricing tier adoption data). Revenue distribution was highly skewed: the median store generated $6,000 annually whilst the mean was $261,435. At the 25th percentile, stores generated $600 annually, whilst at the 99th percentile stores reached $4.0 million. Maximum estimated revenues exceeded $10 billion for individual domains.

**Application Adoption Patterns** Stores varied substantially in application deployment. The median store installed 3 applications (mean: 4.0). A substantial portion (515,206, 19.1%) installed zero applications. Most stores (1,457,466, 53.9%) used 1 to 5 applications. Stores with 6 to 10 applications represented 18.0% (486,776), whilst those with 11 to 20 applications accounted for 8.1% (220,071). Only 22,286 stores (0.8%) installed more than 20 applications (maximum: 141). Install distribution across applications showed extreme concentration. The median install count was zero and the 75th percentile was 2 installs. Store Leads detects installations through storefront scripts, potentially undercounting private integrations, direct partnerships, or applications with limited client-side presence. Zero-install rates varied substantially by category, exceeding 95% in technical infrastructure categories (payment providers, ERP, accounting) whilst remaining lower in consumer-facing categories, suggesting backend integrations are particularly difficult to detect.

**Temporal Trends** Application creation accelerated substantially over time, as shown in Figure 2. Among applications created since 2015, annual counts increased from 203 in 2015 to 4,106 in 2024, representing more than a twentyfold increase over nine years. The peak year for application creation was 2024, followed by 2023 (3,719 applications) and 2022 (3,102 applications). The year 2025, represented only partially in the dataset (data collected 28 September 2025), had already recorded 3,920 new applications, suggesting continued growth. Store creation followed a similar trajectory, with 76,062 new stores in 2015 rising to 466,906 in 2024.

The temporal span of the dataset extended 16.3 years for applications (earliest created 2 June 2009) and 19.3 years for stores (earliest created 2 June 2006). This longitudinal coverage enabled examination of first-mover dynamics in RQ3. The data confirmed that the Shopify ecosystem matured substantially after 2015, with both supply-side (application development) and demand-side (merchant adoption) growth accelerating in recent years.

**Geographic Distribution** Stores exhibited geographic concentration in a small number of markets. The United States accounted for 1,035,557 stores (38.3% of the total), making it the dominant market by a substantial margin. The United Kingdom ranked second with 208,724 stores (7.7%), followed by Canada (137,644 stores, 5.1%), India (130,523 stores, 4.8%), and Australia (126,226 stores, 4.7%). The top five countries collectively represented 60.7% of all stores in the dataset. Beyond these leading markets, adoption was distributed across more than 100 additional countries, though individual counts remained modest. This geographic pattern indicates that whilst Shopify operates as a global platform, the majority of merchant activity concentrates in English-speaking markets and India.

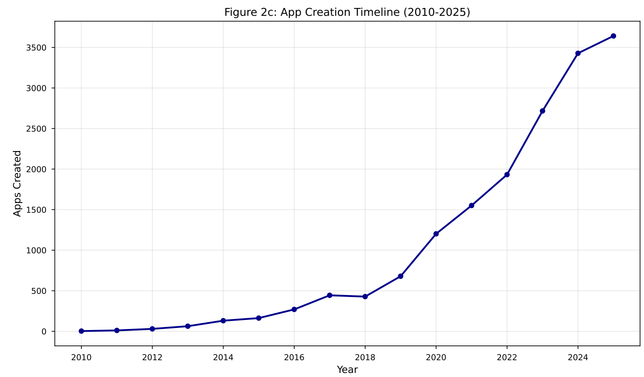


**Figure 2: Application creation timeline from 2010 to 2025, showing more than twentyfold growth since 2015.**

**Pricing Models** Among the 4,213 active applications with at least one installation, the majority employed freemium pricing strategies. A total of 2,922 applications (69.4%) offered a minimum price of $0, indicating free tiers or entirely free offerings. The remaining 1,291 applications (30.6%) required payment from the outset, with minimum prices greater than zero. This distribution suggests that application developers predominantly rely on freemium models to reduce adoption barriers, likely converting users to paid tiers after initial trial periods or through premium feature upsells.

***Summary*** The Shopify app ecosystem comprises more than 24,000 applications serving 2.7 million merchant stores with an estimated $706 billion in aggregate annual merchant revenue. Of 16,698 active applications, 74.8% recorded zero installations at the time of data collection, though this figure likely reflects both measurement limitations in web scraping methodologies and genuine market entry barriers. Among applications with recorded installations, install counts remained heavily concentrated, with the median at zero and the 75th percentile at 2 installations. A small number of applications dominated their categories whilst the majority showed limited recorded adoption. Merchants typically deployed small numbers of applications (median of 3), and geographic activity concentrated in the United States, which accounted for 38.3% of stores. The ecosystem grew rapidly from 2015 to 2024, with application creation accelerating twentyfold. Most active applications employed freemium pricing models.

## 5.2 RQ2: Concentration Within Categories

We analysed concentration patterns across 55 categories that contained at least 20 active applications with market participation. These categories collectively represented the competitive landscape of the Shopify app ecosystem, spanning functional areas from product reviews and email marketing to logistics and payments.

**Overall Concentration Patterns** Concentration varied substantially across categories, indicating that winner-take-all dynamics do not uniformly characterise the ecosystem. The mean Herfindahl-Hirschman Index across all analysed categories was 0.190, with a median of 0.144. Mean top-5 market share was 71.2%, whilst the median was 71.1%. Figure 3 presents the distribution of HHI values across all categories. The distribution exhibits right skew, with most categories clustering in the low to moderate range (HHI between 0.05 and 0.20) whilst a smaller number of categories extend into high concentration territory (HHI above 0.25). This pattern suggests that whilst leading applications typically capture the majority of adoption within their categories, a substantial portion of the market remains available to competitors in most functional areas.

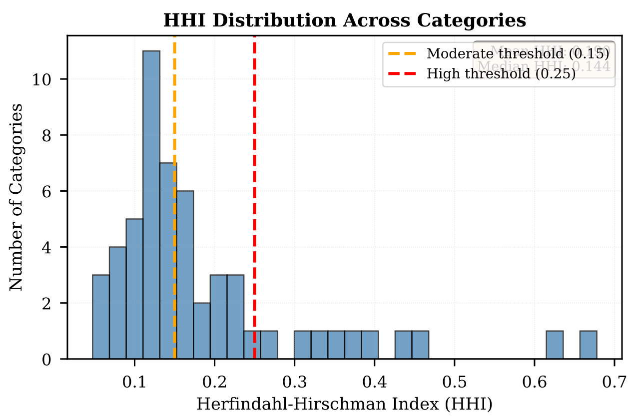


**Figure 3: HHI distribution across 55 categories. Dashed lines show thresholds for moderate (0.15) and high (0.25) concentration.**

Based on conventional HHI thresholds [13], we identified 11 categories (20.0%) exhibiting high concentration (HHI ≥ 0.25), 16 categories (29.1%) exhibiting moderate concentration (0.15 ≤ HHI < 0.25), and 28 categories (50.9%) exhibiting low concentration (HHI < 0.15). The majority of categories therefore operated in competitive or moderately competitive regimes rather than monopolistic structures. This finding indicates that whilst dominant applications exist in many categories, the ecosystem maintains sufficient diversity that new entrants face varied competitive conditions depending on their chosen functional area.

**Highly Concentrated Categories** The 11 highly concentrated categories exhibited strong winner-take-all dynamics. The most concentrated category, social proof, recorded an HHI of 0.678, with the leading application capturing 81.9% market share and the top five applications collectively holding 93.6%. Other highly concentrated categories included design elements (HHI = 0.634), chat (HHI = 0.467), wishlists (HHI = 0.428), and search and filters (HHI = 0.393). These categories exhibited clear market dominance, with single applications controlling majority positions.

Among the top 20 categories ranked by total installation volume, three exceeded the high concentration threshold. Social proof exhibited the highest concentration (HHI = 0.678), with Instafeed dominating with 81.9% market share. Chat also qualified as highly concentrated (HHI = 0.467), with Shopify Inbox holding 67.8% market share. Product reviews (HHI = 0.253) also exceeded the threshold, with Judge.me Product Reviews App capturing 47.7% of installations. The remaining 17 categories among the top 20 operated in low or moderate concentration regimes, indicating that the largest and most economically significant functional areas generally maintained competitive structures.

**Competitive Categories** At the opposite end of the distribution, the least concentrated categories exhibited fragmented market structures with no dominant providers. The category with the lowest concentration, discounts, recorded an HHI of 0.047, with the leading application (Discount Ninja) capturing 11.4% market share and the top five collectively holding 42.9%. Other highly competitive categories included ads (HHI = 0.052), upsell and cross-sell (HHI = 0.059), image gallery (HHI = 0.076), and product variants (HHI = 0.080). These categories featured dozens or hundreds of applications competing for adoption, with no single provider establishing clear dominance.

Among the top 20 categories by installation volume, several major functional areas remained highly competitive despite their economic importance. The upsell and cross-sell category, with 149 active applications and 237,012 total installations, exhibited an HHI of only 0.059, with the leading application (Selleasy) capturing 16.1% market share. Similarly, the discounts category (HHI = 0.047) maintained highly competitive structures despite substantial total adoption. Email marketing (HHI = 0.191) and page builder (HHI = 0.210) operated in moderate concentration regimes, indicating viable competition among multiple providers.

*Concentration and Category Characteristics.* We examined whether concentration varied systematically with category size, measured by the number of active applications. The mean HHI decreased with category size. Smaller categories exhibited higher concentration, whilst larger categories sustained more competitors at viable market shares. This inverse relationship between category size and concentration contradicts the expectation that larger markets would exhibit stronger winner-take-all effects due to network externalities or economies of scale. Instead, larger categories sustained more competitors at viable market shares, suggesting that functional diversity, merchant heterogeneity, or niche differentiation opportunities expand more rapidly than concentration forces in growing categories.

We also examined whether concentration correlated with total category adoption, measured by aggregate installations across all applications in a category. No clear systematic relationship emerged. Categories with fewer than 100,000 total installations exhibited HHI values ranging from 0.05 to 0.68, whilst categories with more than 1 million installations spanned HHI values from 0.05 to 0.47. High adoption categories included both highly concentrated examples (social proof with 281,351 installations and HHI = 0.678) and competitive examples (upsell and cross-sell with 237,012 installations and HHI = 0.059). This pattern indicates that market size alone does not determine competitive structure.

**Market Share Distribution Within Categories** Examining the distribution of market shares within individual categories revealed additional structure beyond aggregate HHI values. In highly concentrated categories, the leading application typically captured between 50% and 82% market share, whilst the top five applications collectively held 80% to 95%. In moderately concentrated categories, the leader held 20% to 50% share, and the top five captured 60% to 80%. In low concentration categories, even the leading application rarely exceeded 20% share, and the top five seldom surpassed 55%.

The relationship between top-1 market share and top-5 market share varied across categories. In winner-take-all categories, a single dominant application often held market share approaching that of the entire top five combined, indicating weak secondary competition. In more fragmented categories, the top five applications held substantially more combined share than the leader alone, indicating that multiple competitors operated at comparable scale. This pattern suggests that concentration manifests through different mechanisms: some categories exhibit single-firm dominance, whilst others feature oligopolistic structures with several large competitors and a fragmented tail.

***Summary.*** Concentration within Shopify app categories varies substantially depending on functional area. Whilst 20% of categories exhibit winner-take-all dynamics with dominant providers capturing majority market share, more than half of all categories operate in competitive regimes where the top five applications collectively hold less than 60% of installs. Larger categories, measured by number of competing applications, systematically exhibit lower concentration, suggesting that market growth enables rather than forecloses competition. The most economically significant categories span the full range of concentration levels, with some (social proof, chat, product reviews) dominated by incumbents and others (upsell and cross-sell, discounts, email marketing) maintaining diverse competitive fields. These patterns indicate that opportunities for new applications depend critically on category selection, with competitive functional areas offering substantially greater prospects for market entry than monopolised categories.

## 5.3 RQ3: Early Mover Growth Dynamics

With RQ3 we aimed at understanding whether apps entering categories earlier maintain growth advantages over time. Whilst cumulative install counts mechanically favour older applications through longer market exposure, recent growth patterns reveal whether early movers sustain competitive momentum or face erosion from newer entrants. We analysed 3,456 applications across 50 categories, restricting analysis to apps with at least 180 days of market tenure to ensure growth patterns had stabilised beyond initial launch volatility.

We decomposed growth into two dimensions: velocity (absolute growth measured by installations in the last 90 days) and rate (proportional growth calculated as recent installations divided by total installations). This separation distinguishes visibility advantages from momentum effects. We calculated Spearman rank correlations between app creation timing and both growth dimensions, applying FDR correction (Benjamini-Hochberg procedure, $q < 0.05$) to control for multiple comparisons across categories.

*Growth Patterns Favour Late Entrants.* Contrary to expectations of persistent first mover advantages, late entrants exhibited stronger growth across both dimensions in the majority of categories. Of 50 categories, 45 (90.0%) showed positive velocity correlations, indicating that later-created apps achieved higher absolute growth than earlier-created apps. Similarly, 48 categories (96.0%) exhibited positive rate correlations, indicating faster proportional growth amongst late entrants. Nine velocity correlations and eight rate correlations reached statistical significance after FDR correction.

The pattern proved consistent across major functional areas. Email marketing, the largest category by installation volume, showed velocity correlation of $\rho = 0.261$ and rate correlation of $\rho = 0.203$. Product reviews exhibited $\rho = 0.186$ for velocity and $\rho = 0.066$ for rate. Chat applications showed $\rho = 0.203$ for both dimensions. Only five categories showed negative velocity correlations, indicating early mover growth advantages, and only two showed negative rate correlations.

**Early Mover Obsolescence, Not Late Mover Advantage.** Examination of median growth values revealed that the observed pattern reflects early mover decline rather than exceptional late mover performance. Figure 4 presents median velocity for early movers (first quartile by creation date) versus late movers (fourth quartile) across all 50 categories. The concentration of 37 categories (74%) in the upper-left quadrant demonstrates systematic early mover obsolescence: early movers experienced net merchant churn (negative velocity) whilst late movers captured growth (positive velocity). Of 50 categories, 45 exhibited negative median velocity for early movers, indicating net churn where merchant uninstallations exceeded new installations, whilst 42 late movers exhibited positive median velocity.

In product reviews, early movers recorded median velocity of −12.5 installations per 90 days whilst late movers recorded +2 installations. Email marketing showed −1.5 for early movers and +1.5 for late movers. Order tracking exhibited the starkest contrast, with early movers at −87 and late movers at +1.5. This pattern indicates app lifecycle dynamics where early movers experience merchant abandonment whilst late movers capture growth. Store Leads data track net changes in visible installations, detecting both additions and removals as merchants modify their storefronts. Negative growth values therefore represent genuine churn rather than measurement artefacts. The internal consistency of the data supports this interpretation: velocity and rate signs agreed in 97% of measurements, confirming that both metrics capture the same underlying dynamics.

We classified categories into competitive dynamics based on correlation patterns. Only one category (2.0%) exhibited strong first mover advantage (negative correlations for both velocity and rate), visible as the red point in Figure 4. Four categories (8.0%, orange points) showed the eroding first mover advantage pattern (negative velocity correlation, positive rate correlation), where early movers maintained absolute growth advantages despite late movers achieving faster proportional growth. The remaining 44 categories (88.0%, green points) favoured late entrants on both dimensions, whilst one category showed no clear pattern.

Whilst the systematic pattern favours late entrants, individual outcomes vary within categories. Product reviews illustrates this heterogeneity: the category exhibits high concentration with a dominant early mover sustaining strong growth, whilst other early

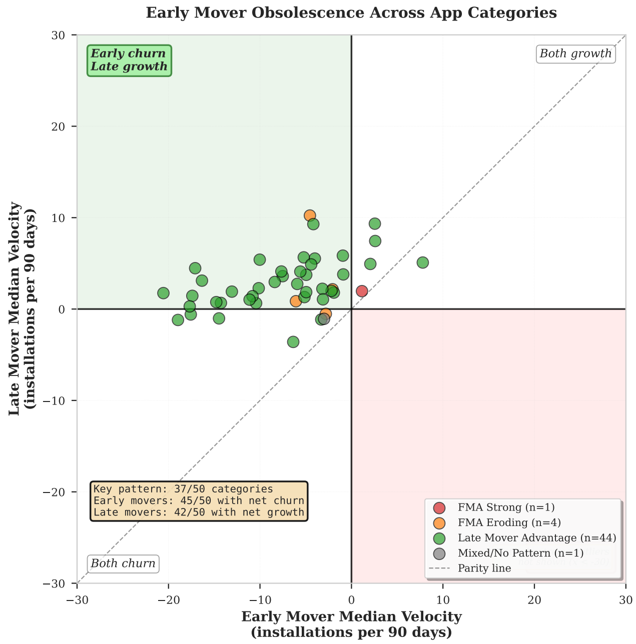


**Figure 4: Early mover obsolescence across 50 categories. Points compare median 90-day installs of early (Q1) vs late (Q4) movers. Upper-left (37, 74%) shows churn among early movers. Colours: red (Strong FMA), orange (Eroding), green (Late Advantage), grey (Mixed).**

entrants experience typical obsolescence patterns. This demonstrates that early entry combined with sustained quality and continuous innovation can maintain competitive positions, though the category-level pattern of late mover advantage persists.

**No Evidence of Momentum Shifts.** To test whether late movers exhibited accelerating growth that might eventually reverse early mover advantages, we compared recent momentum (30-day installations) against longer-term averages (90-day installations divided by three). Apps with 30-day growth exceeding their 90-day average were classified as accelerating. Within each category, we compared acceleration rates between early movers and late movers using Mann-Whitney U tests with FDR correction.

Of 50 categories, 21 (42.0%) showed higher acceleration amongst late movers, whilst 29 showed higher acceleration amongst early movers. However, no category reached statistical significance after FDR correction. Mean Cohen's $d$ was $-0.080$, indicating a negligible effect size. The proportion of apps exhibiting acceleration was 36.7% for early movers and 38.3% for late movers, a difference of 1.6 percentage points. These results provide no evidence of directional momentum shifts that would alter competitive dynamics over time.

***Illustrative Catch-Up Scenarios.*** To contextualise the magnitude of growth differences, we projected forward under three mathematical scenarios, emphasising that these represent illustrative bounds rather than predictions. We calculated baseline installation counts and growth metrics for early and late movers in each category, then projected forward over 20 quarters (five years) under exponential growth (constant rates), linear growth (constant absolute additions), and decay (rates declining linearly to zero). These scenarios assume growth patterns persist without market saturation, competitive responses, quality changes, or platform algorithm shifts.

Under the realistic decay scenario, which assumes growth rates gradually decline to zero over the projection horizon, only five categories provided sufficient positive growth data for projection. Among these, one category (20.0%) showed late movers reaching 50% of early mover size within two years, whilst four (80.0%) showed late movers failing to reach this threshold within five years. Mean ratio after 24 months was 0.58, indicating late movers reached 57.8% of early mover installation volumes under these assumptions. Median time to parity was indeterminate due to insufficient categories achieving the threshold.

However, the limited applicability of these projections (five categories rather than 50) reflects the prevalence of negative growth amongst early movers, which violates the mathematical assumptions required for forward projection. The projections therefore provide limited insight beyond confirming that observed growth differentials, if sustained, would require extended periods for convergence.

***Summary.*** Recent growth patterns systematically favour late entrants over early movers across 88% of analysed categories. This pattern reflects early mover obsolescence, with established apps experiencing net merchant churn whilst newer alternatives capture growth. The dynamics suggest app markets remain contestable despite installed base advantages, as merchants migrate towards newer solutions over time. No evidence emerged of accelerating late mover momentum that would further alter competitive dynamics. These findings indicate that entry timing confers advantages through recency rather than incumbency, contradicting traditional first mover advantage theory in platform app ecosystems.

## 6 Threats to Validity

*Revenue measurement.* Store revenue estimates represent merchant e-commerce sales rather than app subscription revenue. The $706 billion figure contextualises merchant base scale but not developer revenue.

*Construct validity.* Install counts combine active and abandoned usage, do not capture monetisation, and mechanically advantage older applications. We mitigated this by restricting to active apps with installs, using relative metrics (market share within categories), and modelling temporal effects through velocity and rate decomposition in RQ3. Store Leads tracks net changes; negative recent values indicate genuine churn. Creation date proxies entry timing, conflating age with launch sequence, but remains the best available temporal measure.

*Internal validity.* We required categories with at least 20 active apps and analysed only applications aged 180+ days in RQ3, avoiding small-market noise and launch volatility. Early and late movers (first and fourth quartiles by creation date) maximise contrast whilst retaining sample size. Survivorship bias would bias against observing early mover decline; the consistent pattern across categories suggests genuine dynamics rather than artefacts.

*External validity.* Results derive from Shopify (September 2025), which targets SMEs, uses subscription pricing, and curates its store. RQ1 scale metrics likely transfer to other B2B platforms. RQ2 concentration patterns may vary by governance and maturity, though the inverse category size-concentration relationship challenges universal winner-take-all assumptions. RQ3 obsolescence patterns may extend to ecosystems with feasible switching but not high lock-in platforms.

*Temporal validity.* Growth metrics rely on a single 90-day window (July-September 2025), precluding temporal stability assessment. Patterns could reflect period-specific conditions. The systematic nature across 50 categories suggests robustness, though longitudinal data would confirm generalizability.

*Data quality.* Weekly scraping misses rapid changes; cumulative installs cannot decrease despite uninstalls, though recent metrics detect churn. Incomplete linking affects demand-side analysis but not concentration or growth measures. Preprocessing reduced inconsistencies. Revenue estimates use traffic models, introducing absolute error but preserving relative patterns.

## 7 Conclusion

This study analysed the Shopify app ecosystem using data from 24,826 applications and 2.7 million merchant stores. We examined ecosystem scale, competitive structure across categories, and the role of entry timing in growth dynamics. Of the 24,826 applications, 4,213 achieved market entry (at least one installation). Concentration varied across the 55 categories analysed: over half were low-concentration, while 20% were highly concentrated. Larger categories consistently showed lower concentration, contradicting expectations of winner-take-all outcomes. Entry timing revealed systematic early mover decline rather than sustained advantage. In 88% of categories, later entrants achieved higher growth, while many early movers experienced net merchant churn. These patterns suggest that markets remain contestable, with recency outweighing incumbency.

*Theoretical Implications.* These findings contribute to platform ecosystem theory in three ways. First, the inverse relationship between category size and concentration challenges universal winner-take-all assumptions, demonstrating that market growth can enable competitive diversity through niche differentiation rather than consolidation. Second, the systematic pattern of early mover obsolescence extends first mover advantage theory by showing that timing advantages prove contingent rather than persistent, requiring sustained quality and innovation to maintain positions. Third, the heterogeneity across functional categories indicates that ecosystem dynamics depend on category characteristics rather than following universal laws, supporting context-specific rather than deterministic platform theories.

*Practical Implications.* Developers face heterogeneous conditions depending on category structure. Competitive categories offer better prospects, and late movers can succeed even in mature markets through feature improvements and design innovation. For platform operators, the inverse link between category size and concentration shows that ecosystem growth can expand competitive opportunities, though within-category leaders still benefit from installed base effects. For merchants, the prevalence of early mover obsolescence indicates the value of periodically re-evaluating app portfolios rather than relying on incumbents.

## References

[1] Shiva Agarwal, Cameron D Miller, and Martin Ganco. 2023. Growing platforms within platforms: How platforms manage the adoption of complementor products in the presence of network effects? *Strategic Management Journal* 44, 8 (2023), 1879–1910.
[2] Kevin J Boudreau. 2012. Let a thousand flowers bloom? An early look at large numbers of software app developers and patterns of innovation. *Organization science* 23, 5 (2012), 1409–1427.
[3] Marco Ceccagnoli, Chris Forman, Peng Huang, and DJ Wu. 2012. Cocreation of value in a platform ecosystem! The case of enterprise software. *MIS quarterly* (2012), 263–290.
[4] Carmelo Cennamo and Juan Santaló. 2019. Generativity tension and value creation in platform ecosystems. *Organization science* 30, 3 (2019), 617–641.
[5] Jens Foerderer, 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 (2018), 444–460.
[6] Jens Foerderer, Nele Lueker, and Armin Heinzl. 2021. And the winner is...? The desirable and undesirable effects of platform awards. *Information Systems Research* 32, 4 (2021), 1155–1172.
[7] Shu He, Jing Peng, Jianbin Li, and Liping Xu. 2020. Impact of platform owner's entry on third-party stores. *Information Systems Research* 31, 4 (2020), 1467–1484.
[8] Peng Huang, Marco Ceccagnoli, Chris Forman, and DJ Wu. 2009. When do ISVs join a platform ecosystem? Evidence from the enterprise software industry. *ICIS 2009 Proceedings* (2009), 161.
[9] Aldona Kapacinskaite and Ahmadreza Mostajabi. 2024. Competing with the platform: Complementor positioning and cross-platform response to entry. *Strategic Management Journal* 45, 12 (2024), 2577–2607.
[10] Tong Li, Yali Fan, Yong Li, Sasu Tarkoma, and Pan Hui. 2021. Understanding the long-term evolution of mobile app usage. *IEEE Transactions on Mobile Computing* 22, 2 (2021), 1213–1230.
[11] Konstantinos Plakidas, Srdjan Stevanetic, Daniel Schall, Tudor B Ionescu, and Uwe Zdun. 2016. How do software ecosystems evolve? A quantitative assessment of the R ecosystem.. In *Proceedings of the 20th international systems and software product line conference*. 89–98.
[12] Sruthi Thatchenkery and Riitta Katila. 2023. Innovation and profitability following antitrust intervention against a dominant platform: The wild, wild west? *Strategic Management Journal* 44, 4 (2023), 943–976.
[13] U.S. Department of Justice and Federal Trade Commission. 2010. *Horizontal Merger Guidelines*. Technical Report. U.S. Department of Justice. https://www.justice.gov/atr/horizontal-merger-guidelines-08192010 Issued August 19, 2010.
[14] Joey Van Angeren and Arvind Karunakaran. 2023. Anchored inferential learning: Platform-specific uncertainty, venture capital investments by the platform owner, and the impact on complementors. *Organization Science* 34, 3 (2023), 1027–1050.
[15] 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 (2019), 1336–1367.
[16] Feng Zhu and Qihong Liu. 2018. Competing with complementors: An empirical look at Amazon. com. *Strategic management journal* 39, 10 (2018), 2618–2642.