Long Range Planning 58 (2025) 102540

Contents lists available at ScienceDirect

Long Range Planning

journal homepage: www.elsevier.com/locate/lrp

Scalability in incumbent firms: The case of Nvidia

Ferran Vendrell-Herrero [a,b,*], Yancy Vaillant [c], Oscar F. Bustinza [d]

[a] International Management, Strategy Group, University of Edinburgh Business School, University of Edinburgh, 29 Buccleuch Place, Edinburgh, EH8 9JS, UK
[b] Strategic Business Development, University of Vaasa, Vaasa, FI-65101 Vaasa, Pohjanmaa, Finland
[c] Department of Strategy, Entrepreneurship & Innovation, TBS Education, 1 Place Alphonse Jourdain, Toulouse, 31068 Cedex 7, France
[d] Faculty of Economics and Business, Department of Management I, Universidad de Granada, Campus Cartuja, C.P. 18071, Granada, Spain

ARTICLE INFO

Keywords:
Scalability
Firm growth
Process model
Nvidia
Single case study

ABSTRACT

Scalability refers to the organizational capabilities required to facilitate a smoother and faster scaling process. Although it is usually associated with new ventures, this study explores how established firms can also create conditions conducive to scalability. We address this question by applying an inductive, narrative-based approach to a longitudinal, single-case study of Nvidia Corporation, a company founded in 1993 that since 2006 has undergone a profound transformation driven by the AI revolution. This case study draws on digital archives, including objective accounting information on Nvidia and its direct competitors, extensive company reports, pedagogical case studies, corporate biographies, and 464 min of recorded documentaries and interviews featuring the company's CEO. We use these sources to develop a multi-phase theoretical model outlining how established organizations can foster scalability. The model encompasses value recognition driven by systemic industry transitions, organizational adaptability, strategic renewal, and scalability, thus offering a structured framework for understanding how incumbent firms can cultivate the necessary conditions for successful scaling.

1. Introduction

The capability of some firms to attain exponential growth, what is known as 'scalability', has attracted increasing attention among academics, managers, and policymakers (Piaskowska et al., 2021; Reuber et al., 2021; Jansen et al., 2023). Scalability is now recognised as a defining trait of successful entrepreneurship (Sleuwaegen and Ramboer, 2020; Motley et al., 2023; Bohan et al., 2024), with ample research having been published on scalability at start-up (Davidsson and Henrekson, 2002; Mason and Brown, 2013; Rawhouser et al., 2022), scalable business models (Stampfl et al., 2013; Zhang et al., 2015; Sanasi, 2023), and how scaling firms configure their activities (Piaskowska et al., 2021; Reuber et al., 2021).

However, the issue has been approached far less from the organizational and strategic perspective of mature firms, other than isolated analyses from Penrosean growth theory (Penrose, 1959, 2009; Barney, 1991), the evolutionary lens (Nelson and Winter 1982; Dosi and Nelson, 1994) and the attention-based view (Ocasio, 2011; Joseph and Wilson, 2018). Although a growing number of mature firms have achieved hypergrowth, often driven by the network effects of digitalisation within a globalized world (Autor et al., 2020; Jansen et al., 2023), the aforesaid studies have tended to address their growth and performance. A clear framework explaining how

* Corresponding author. International Management, Strategy Group, University of Edinburgh Business School, University of Edinburgh, 29 Buccleuch Place, Edinburgh. EH8 9JS, UK.

E-mail addresses: fvendrel@ed.ac.uk (F. Vendrell-Herrero), y.vaillant@tbs-education.es (Y. Vaillant), oscarfb@ugr.es (O.F. Bustinza).

https://doi.org/10.1016/j.lrp.2025.102540
Received 2 July 2024; Received in revised form 23 May 2025; Accepted 2 June 2025
Available online 2 June 2025

incumbent firms can develop the necessary scalability capabilities remains elusive (Giustiziero et al., 2023; Tidhar et al., 2024).

Ideally, start-ups identify an opportunity within a value system and strike while the iron is hot by equipping themselves with the appropriate capabilities and business model for scalability (Snihur and Zott, 2020; Mihailova, 2023). In contrast, and as is the case with many other organizational and strategic processes (Penrose, 1959, 2009), incumbent firms that were established before the advent of the opportunities brought on by digital disruption bear little resemblance to start-ups in terms of their scalability pathways (Giustiziero et al., 2023; Tidhar et al., 2024). They are rarely endowed with the necessary capabilities for such scalability (Piaskowska et al., 2021), and tend to be constrained by their infrastructural, managerial, and strategic legacies. Jansen et al. (2023) classify incumbent firms that achieve exponential growth as superstars, rather than scaleups, due to their rarity and the significant challenges they face compared to start-ups. We therefore need to ask how incumbent firms can develop the right conditions to achieve scalability and, hence, capitalize on high-growth opportunities.

The study presented in this paper investigates how this might be done. Our insights are derived from an in-depth, longitudinal case study (Welch et al., 2020, 2022), employing a narrative-based methodology (Golden-Biddle and Locke, 2006; Rowlinson et al., 2014). The focal case is Nvidia Corporation—a prime example of an incumbent that has successfully achieved scalability. Utilizing extensive digital archives (Nix and Decker, 2023), our analysis incorporates financial and accounting data on Nvidia and its direct competitors (Intel, AMD, and Broadcom),[1] comprehensive company reports, pedagogical case studies covering different stages of the company's evolution, two recent corporate biographies, 273 min of interviews with CEO Jensen Huang, and 191 min of documentary footage. The insights from Nvidia's experience are then integrated with existing theoretical literature (Jansen et al., 2023; Bohan et al., 2024; Coviello et al., 2024) to develop a novel, multi-phase process model that outlines the steps that incumbent firms need to take in order to establish scalability conditions. These encompass value recognition, organizational adaptability, and strategic renewal—each representing distinct, progressive phases on the path toward scalability.

This model has major potential to influence thinking and practice in terms of organizational regeneration and technological adoption. We particularly observe how incumbent companies need to mimic the approach taken by start-ups, while simultaneously drawing on their accumulated resources and experience.

This study responds to a number of calls, particularly the appeal by Coviello et al. (2024) for better managerial insight on how to achieve scalability. The study also fills some of the gaps identified by Shepherd and Patzelt (2022) and Jansen et al. (2023) regarding micro-level research into the behaviours and capabilities needed to facilitate organisational scalability. The study also aligns with the calls by McKelvie and Wiklund (2010) for more research seeking to explain how scale is achieved (growth) rather than simply quantifying the resulting growth, by Stampfl et al. (2013) for further research into the environmental conditions for strategic scalability in an increasingly digital economy, and by Piaskowska et al. (2021) for studies on scalability in incumbent firms. What these calls clearly suggest is that the generalization of scalability from its usual context of entrepreneurship (Shepherd and Patzelt, 2022) to that of established organizations is an important issue that warrants in-depth analysis, which is precisely what this paper endeavours to provide.

This paper is organized as follows (see Golden-Biddle and Locke, 2006). First, we introduce our research question. We then outline the methodology employed in the case study of Nvidia, followed by a detailed description of the company's background and key organizational milestones in its achievement of scalability. We then present the theoretical model derived from this narrative-based analysis. The paper concludes with a discussion of its theoretical and practical implications.

2. Background and research question

Digitalisation and machine learning are profoundly transforming business and industry by making scalability accessible to many industries and business models that were previously considered unsuitable for scale-up (Jansen et al., 2023; Bohan et al., 2024). Scalability can be defined as the capabilities that determine the extent to which a business may achieve its desired value creation and capture targets as user/customer numbers increase and their needs change (Stampfl et al., 2013). For firms, scalability is no longer merely an economic issue but has become a major strategic imperative (Jacobides et al., 2018). Together with supply-side economies of scale and demand-side returns-to-scale, scaling-up also generates several value-adding attributes that companies can leverage for competitive advantage (Van Alstyne et al., 2016; Kuratko et al., 2020; Jansen et al., 2023).

A distinction should be made between scalability, scaling, and scale-up. The concepts are all related, but carry different meaning and measure. This study focuses on the former, **Scalability,** which is defined by Coviello et al. (2024, pp. 2) as "an organizational capability developed by managing and achieving coherence among a firm's technological architecture, organizational architecture, and business model." This is different from **Scaling,** which the same authors define as **"**an organizational process whereby managers transform the internal organization and leverage digital resources to rapidly expand a firm's outputs without a corresponding ex-ante increase in inputs." Finally, **Scale-up** is "a phase of organizational development where a firm is actively engaged in the scaling process" (Coviello et al., 2024, pp.15). Hence, scalability is a capability, scaling is a process, and scale-up is a phase of organizational development. Because empirical studies of scalability and its achievement are still rare, Coviello et al. (2024) call for greater research in the area, a gap that our own study seeks to fill.

While scalability underpins scale-up, it also facilitates a smoother and faster scaling process (Coviello et al., 2024, pp.16). This is achieved when a firm can grow efficiently without facing severe resource constraints (Barney, 1991), and when it is able to

[1] There are other chip companies in the smartphone industry, such as Qualcomm and MediaTek, but we limit our comparative analysis to the computer industry.

dynamically reconfigure, integrate, and adapt its resources and capabilities in response to identified growth opportunities (Teece et al., 1997). Theoretically, scalability, as a condition for exponential growth, requires an organizational and strategic balance between internal capability development and external expansion needs (Penrose, 1959). Scalability, therefore, enables the scaling process. Somaya and You (2024) view scalability as a firm's ability to adequately align opportunity, organisation, and strategy for scale-up. However, it is often much easier for new entrants to adopt scalable structures and business models from inception than it is for incumbent firms to transform their strategic and organisational foundations in pursuit of a scale-up trajectory (Menz et al., 2021). These firms are likely to face more severe constraints when it comes to dynamically adapting their capabilities to align with scalability demands.

Start-ups typically identify an opportunity within a value system and subsequently adopt the right scalable business model (Mihailova, 2023). Incumbents, however, are often constrained by their infrastructural, managerial, and strategic legacies. They must embark on complex and often disruptive transformations as they develop the adequate preconditions for scalability.

Entrepreneurial scalability is directly linked to the identification of opportunities and the capacity to generate value (Mason and Brown, 2013). However, for incumbent firms, the path is not so direct. The research question addressed in this study is therefore: 'How can established organizations cultivate the capabilities required for scalability?' In search of answers to this question, the following sections develop a theoretical model based on the experience of Nvidia Corporation, an incumbent firm that has undergone key transformations in order to achieve the right conditions for scalability.

3. Methodology

3.1. Research design and case selection

This study employs a single case study methodology to provide a nuanced exploration of Nvidia Corporation's journey towards scalability. This is an exemplary case due to Nvidia's significant scalability and influence within the technology industry. Founded in 1993 and renowned for its graphics processing units (GPUs), the company has identified and leveraged groundbreaking technological opportunities, and particularly the transformative power of artificial intelligence (AI). By 2024, this strategic focus had catapulted Nvidia to become the world's most valuable semiconductor firm.

Nvidia's extreme nature aligns with the suggestion by (Siggelkow, 2007) that exceptional cases like these are extremely useful for uncovering mechanisms and outcomes that are not so readily apparent in non-extreme cases (Welch et al., 2020, 2022). Exceptional cases often exacerbate underlying structures and phenomena, thus bringing complexity to the forefront and providing fresh insight. This approach aligns with the historical case study tradition of emphasizing the interpretation of events and decisions in their temporal contexts (Kipping and Üsdiken, 2014; Vaara and Lamberg, 2016).

Given the exploratory nature of this study and its goal of understanding pathways to scalability within incumbent firms, an inductive, longitudinal qualitative approach is particularly suitable. A single case study facilitates in-depth investigation of transformative technological trends and their implications (Yin, 2017). By asking "how," "when," and "why", this approach unravels the multifaceted phenomena that contribute to scalability conditions and capabilities within established firms (Langley, 1999; Langley et al., 2013). The adoption of a temporal process perspective allows us to meticulously organize events and business transformations into distinct temporal spaces, thus enhancing the validity and transparency of our findings (Decker, 2022).

3.2. Data collection

This research is grounded in the extensive use of secondary sources and digital archives, a methodology that offers distinct advantages for constructing historical narratives and analysing strategic transformations over time (Nix and Decker, 2023). The richness and diversity of digital archives provide access to authentic, objective records that enable a nuanced exploration of strategic and technological shifts within Nvidia.

Detailed information about our data corpus, including links to secondary sources, is presented in Table 1. Each source is referenced in the narrative using a Source ID code.

Sources taken directly from the company include Nvidia's filings with the U.S. Securities and Exchange Commission (REP), while the external sources include audited accounting reports obtained from Bureau Van Dijk via the Orbis database (FA2) and historical data on firm capitalization (FA1). By leveraging financial and accounting data, our analysis offers insights into the firm's evolution over time.[2] This is further enriched by the use of comparative benchmarks from key competitors in the semiconductor industry (AMD, Intel, and Broadcom) spanning the past 25 years. Additional data is sourced from independent journalists and researchers. Regarding the former, we analysed three documentaries totalling 191 min of viewing time (DO1, DO2, DO3) and four interviews and keynote addresses featuring CEO Jensen Huang, amounting to 273 min (IK1, IK2, IK3, IK4). From an academic standpoint, we reviewed a series of pedagogical case studies published in 2005 (CS1), 2014 (CS2), and 2019 (CS3), which capture the company at critical junctures. We also examined recent cases from Darden (CS4), Harvard (CS5) and INSEAD (CS6), as well as newly released corporate biographies (BO1, BO2). By triangulating these diverse sources, this study adopts a multidimensional approach to historical analysis, thereby enhancing its robustness and depth.

[2] In most instances, the Orbis data fully aligns with company reports. In any cases of discrepancy, we prioritize Orbis data, as it is considered more objective due to being audited by Bureau van Dijk, and to maintain consistency (all comparisons with other firms are conducted using Orbis).

Table 1
List of digital archives and sources.

SOURCE TYPE	SOURCE ID	DIGITAL SOURCE	DURATION	PERIOD COVERED
Financial & accounting	FA1	Companiesmarketcap.com	N/R	1999–2024
	FA2	Orbis (Bureau Van Dijk)	N/R	1999–2024
Reports	REP	Nvidia Annual Report	N/R	2006–2025
Case studies	CS1	IBS-CDC (2005)	N/R	1993–2005
	CS2	USC Marshall (2014)	N/R	1993–2014
	CS3	IMD (2019)	N/R	1993–2019
	CS4	Darden (2022)	N/R	1993–2022
	CS5	Harvard (2024)	N/R	1993–2024
	CS6	Insead (2025)	N/R	1993–2024
Books	BO1	The thinking machine by Stephen Witt (2025)	N/R	1993–2024
	BO2	The Nvidia Way by Tae Kim (2025)	N/R	1993–2024
Documentaries	DO1	CNBC (2023)	18 min	2006–2023
	DO2	CBS, 60 Minutes (2024)	43 min	2006–2024
	DO3	Aquired.fm (2022)	130 min	2006–2022
Interviews & Keynotes	IK1	NYTimes, Dealbook Summit (2024)	44 min	2012–2024
	IK2	CNBC Interview (2024)	34 min	2017–2024
	IK3	Stanford Talk (2024)	53 min	2017–2024
	IK4	Huang GTC Keynote (2025)	142 min	2006–2025

Financial & Reports
Financial & Reports
FA1: Companies Market Cap
FA2: Bureau van Dijk
REP: SEC EDGAR Filings
Case Studies (Available via The Case Centre)
CS1: 305-134-1
CS2: SCG-505
CS3: IMD-7-2036
CS4: UVA-S-0374
CS5: 9-725-360
CS6: 325-0005-1 & 1B

Books
BO1: *The Thinking Machine* – Penguin UK
BO2: *The NVIDIA Way* – W.W. Norton
Documentaries
DO1: CNBC – NVIDIA's Growth from Gaming to AI
DO2: CBS 60 Minutes – April 28, 2024
DO3: Acquired Podcast – NVIDIA (2006–2022)
Interviews & Keynotes
IK1: YouTube
IK2: YouTube
IK3: YouTube
IK4: NVIDIA GTC 2025

Note: The sources used in this study are categorized as follows.

3.3. *Theorizing from the case study*

The theorization process in this case study follows a narrative-based methodological approach (Rowlinson et al., 2014), synthesizing empirical evidence into a structured, chronological account of the target organization's evolution. This approach distinguishes between endogenous strategic choices and exogenous market conditions that collectively shaped organizational change over time (Argyres et al., 2020). To ensure clarity and rigour in presenting this narrative, the study adopts a Tell-Show-Tell framework, as proposed by Golden-Biddle and Locke (2006). The structure is as follows: first, the background and research question are introduced in Section 2 (Tell), followed by the methodology and case narrative in Sections 3 and 4 (Show), and concluding with the presentation of the theoretical model in Section 5 (Tell).

By employing a periodization framework, this study dissects Nvidia's history into distinct phases, each marked by strategic inflection points that contributed to its expansion and consolidation within the technology sector. Such an approach not only provides analytical clarity but also identifies key drivers of scalability, offering theoretical insights of relevance to broader discussions on firm growth and transformation in technology-intensive industries (Karsten, 2014). While Nvidia's experience is context-specific, the underlying strategic patterns identified in its journey offer generalizable implications for established firms navigating technological and market shifts.

Our periodization framework divides Nvidia's evolution into five phases (see Fig. 1), which are defined by key technological shifts that marked its evolution. Phase 1 (1993–2006) corresponds to its start-up years, focused on the design of high-performance chips. Our analysis begins in Phase 2 (2006–2012) with the launch of CUDA, which enabled the use of Nvidia's gaming GPUs for accelerated general-purpose computing. Phase 3 (2012–2017) marked Nvidia's shift toward AI in parallel with the rise of deep neural networks, eventually leading to development of the first AI supercomputer, DGX-1. Phase 4 (2017–2021) brought the company's transition to platformization, exemplified by the strategic expansion of the Nvidia DGX-Platform. Finally, Phase 5 (2022–2025) is characterized by exponential growth, with Nvidia cementing its leadership in AI and high-performance computing.

The theorization process followed an iterative approach, integrating empirical data with theoretical constructs through continuous team discussions (Eisenhardt and Graebner, 2007; Ridder et al., 2014). Rather than relying on coding, our analysis involved a comprehensive review of all available secondary sources, including documents and video materials. Through repeated deliberation, we sought to identify the key strategic decisions that enabled the company to scale. This process was guided by three research vectors: technological shifts, scalability, and decision-making. Iterative engagement allowed us to discern emergent patterns and refine

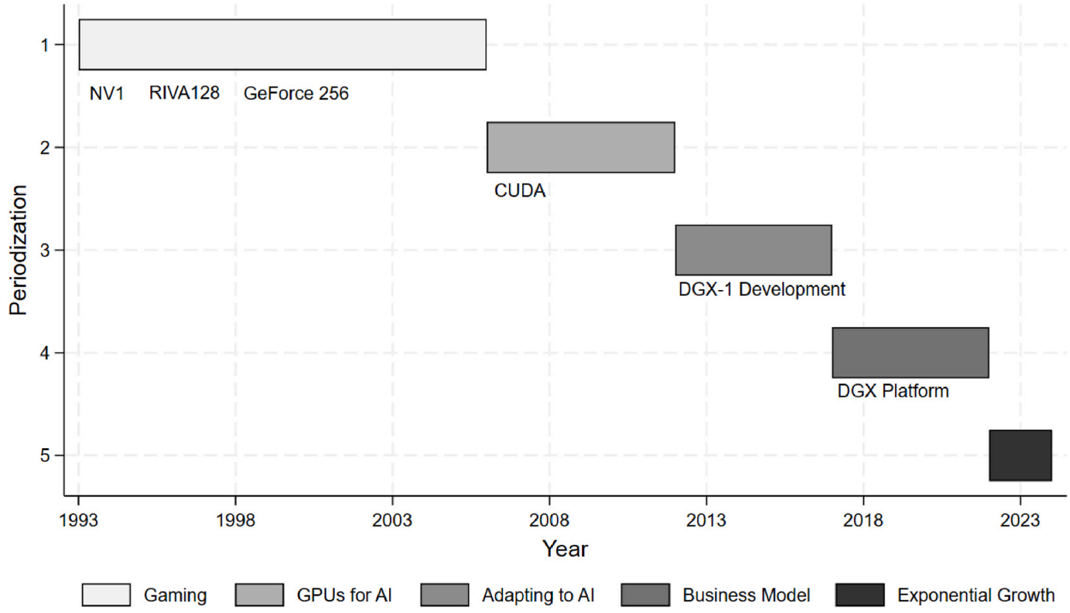

Fig. 1. Case timeline.

theoretical propositions, ensuring rigour and validity. Our analysis culminated in a process-based model outlining the sequential phases of scalability within incumbent firms.

To provide contextual grounding, the narrative begins with an overview of the company, followed by an examination of key indicators of size and value, such as employment, revenue, productivity and market capitalization. These metrics are analysed in relation to the company's primary competitors, AMD, Intel and Broadcom, to establish a comparative benchmark. The narrative then systematically traces the company's evolution across periods, highlighting the key factors that drove each transition.

4. Case narrative

4.1. Case background

Founded in 1993 by Jensen Huang, Chris Malachowsky, and Curtis Priem in Santa Clara, California, Nvidia Corporation was established with the mission of accelerating computing. One of its first applications was in graphics processing technology, initially targeting the gaming and multimedia sectors with the launch of NV1 in 1995 and Riva 128 in 1997. A pivotal moment was the release of GeForce 256 in 1999, the world's first graphics processing unit (GPU), which consolidated Nvidia's dominant position in the graphics computing industry (CS1).

By 2005, the company had emerged as a key player in the semiconductor industry, supplying state-of-the-art graphics processors for applications in gaming, professional visualization, and data centres (CS2). The introduction of the GeForce and Quadro series further reinforced Nvidia's leadership in GPU technology, driving innovation in digital graphics. By this time, Nvidia had expanded globally, with operations in North America, Europe, and Asia, serving both consumer and corporate markets (CS3, CS4, BO1).

Despite maintaining a relatively stable business trajectory for several decades, Nvidia experienced an unprecedented surge in growth from 2022 onwards, fuelled by the rapid expansion of the AI market and the increasing demand for its specialized AI chips (CS5). This transformation underscores Nvidia's role as an incumbent scale-up firm, leveraging its existing market position and technological expertise to capitalize on emerging opportunities in the industry (BO2, CS6).

Nvidia's key competitors have been Intel, AMD and, since 2010, Broadcom, which has arguably followed a more standard scaling process since its start-up stage. Assuming that Nvidia, Intel, AMD, and Broadcom together account for the entire semiconductor industry,[3] the market has undergone significant transformation over time (FA2). Intel, which once dominated with a market share of over 90 %, has been in steady decline, dropping to 28.2 % in the latest data. In contrast, Nvidia has experienced remarkable growth from 0.49 % to 32.36 %, overtaking both AMD and Broadcom. AMD's share has fluctuated but remains at around 12 %, while Broadcom is in a strong position at 27.39 %. This shift reflects a more competitive industry landscape, with several players now holding substantial market shares, reducing Intel's traditional dominance.

[3] We focus on computer-based semiconductors only. Total market size was calculated annually by aggregating the revenues of the four firms, with each company's market share derived as a percentage of this total. While this approach excludes other potential market participants, it is deemed suitable for illustrating the declining market share of Intel and the rising shares of Nvidia and Broadcom.

Fig. 2 showcases Nvidia's evolution since 1999 in terms of value (market capitalization and productivity) and size (revenue and workforce), highlighted by a solid, thick black line. For comparative purposes, the graph also shows the evolution of its main competitors in dotted/dashed thin lines. As the growth rates observed during this period are exponential, we present the graph on a logarithmic scale.

In terms of workforce (see Panel D –FA2), Intel is much more diversified and hence remains the largest company, rising from 64,000 employees in 1999 to 125,000 in 2024. The other three companies have also expanded their workforces, with Nvidia experiencing the most significant relative increase—from approximately 200 employees in 1999 to 26,000 in 2024—a scale now comparable to AMD (25,000) and Broadcom (20,000). Despite doubling its workforce since 2021, Nvidia's employee base is still five times smaller than Intel's. However, in terms of annual sales (see Panel C –FA2), Nvidia has now surpassed Intel, generating $61 billion compared to Intel's $53 billion and Broadcom's $51 billion. Notably, just in the last year Nvidia's revenue grew by 125 %, a clear illustration of its remarkable expansion.

This expansion has naturally led to a substantial increase in its overall value. As Nvidia's revenue has grown at a faster pace than its workforce, labour productivity has significantly improved over the past decade (see Panel B – FA2). This has played a crucial role in the company's increasing market capitalization, as shown in Panel A (FA1). Interestingly, Nvidia's growth trajectory closely resembles that of Broadcom, an archetypal scale-up company, while the more mature Intel and AMD have demonstrated relatively stable and moderate growth.

Our analysis identifies 2006 as the pivotal year in Nvidia's growth trajectory. We shall now examine in detail the four distinct stages of the company's development from then to the present, so as to subsequently analyse the evolution of scalability at Nvidia.

4.2. Phase 1: value recognition (2006–2012)

In 2002, NVIDIA experienced a sharp decline in market value and was embroiled in insider-trading allegations (CS1). The firm responded by realigning its growth strategy with prevailing industry trends. Anticipating rising demand for high-performance parallel computing, it pursued new architectural approaches to advance the technology accordingly (CS2). In 2006, work began on Nvidia's CUDA parallel computing platform, driven by the vision that computing could be significantly accelerated and that this would one day be critical for advanced computation. CUDA, an application programming interface (API), meant that software could use certain types of graphics processing units (GPUs) for accelerated general-purpose processing (CS2). In 2012, AlexNet, a convolutional neural network (CNN) architecture powered by CUDA-enabled GPUs was the very first instance of image recognition with deep machine learning capabilities (BO1). According to Nvidia CEO, Jensen Huang, this was AI's "*Big Bang moment*" when "*machines learned how to learn. And they learned it … on Nvidia*" (DO3).

CUDA's multi-core parallel GPU systems proved far more efficient than general-purpose central processing unit (CPUs) for handling large blocks of data, particularly for algorithms in situations where this is done in parallel, such as deep learning. Having demonstrated the machine learning capabilities of AlexNet, Nvidia realized that AI would not only revolutionise computing, but many other areas of society and that its own GPUs could potentially play an essential role in that revolution (DO1, BO2). This breakthrough prompted Nvidia "*to go all in*" on the development of a GPU-based computer specifically designed for AI (Huang, IK4).

Huang describes how "*luck founded on vision*" led to the delivery in 2016 of the first AI supercomputer to OpenAI's Elon Musk (DO2). Nvidia had anticipated the significance of this opportunity long before the market did, leading to an "*incredible technology risk that would enable its future scale-out capacity*" (Huang, IK4). Investors initially dismissed Nvidia's AI compatible GPU development as a "*useless investment*", and took ten years (2006–2016) to appreciate its real value (Bryan Catanzaro, VP of Applied Deep Learning Research, DO2).

4.3. Phase 2: adaptation (2012–2017)

In 2012, Nvidia "*recognised the opportunity and change necessary*", and took a gamble on where technology would be ten years later (Huang, IK1). They subsequently reoriented their efforts towards generating a new AI-specialized chip, the DGX, which took five years to build (Huang, IK1).

"*Everything in Nvidia was changed towards adapting to AI*" (Huang, IK2), a shift so fully transformational (Catanzaro, IK2) that the founding directors had to "*completely relearn everything we had to know*" (Huang, IK2). Nvidia still mostly designs chips, but since it recognised the value of AI, the company is "*so much more than the chip … We figured out a decade ago that what was needed was a reinvention of the computing problem*" (Huang, IK1). These insights would ultimately affect all of Nvidia's internal work systems and methods.

In order to play a significant role in the AI revolution that it had forecast, Nvidia intended to scale its organizational capacity and technological infrastructure. It was venturing into the unknown as it endeavoured to "*reinvent and disrupt itself*" (Huang, IK2 and REP). But as Huang himself points out (IK4), before they could actually implement and achieve such ambitions, they first needed to grow their capacity to scale.

Nvidia's workforce boomed from 6029 in 2012 and 9227 in 2017 to a massive 26,196 as of 2024 (FA2). New purpose-built facilities were inaugurated, up to 11,000 expert developers were hired, and internal systems were re-engineered, all before a market for AI GPUs had even emerged (CS3, BO1).

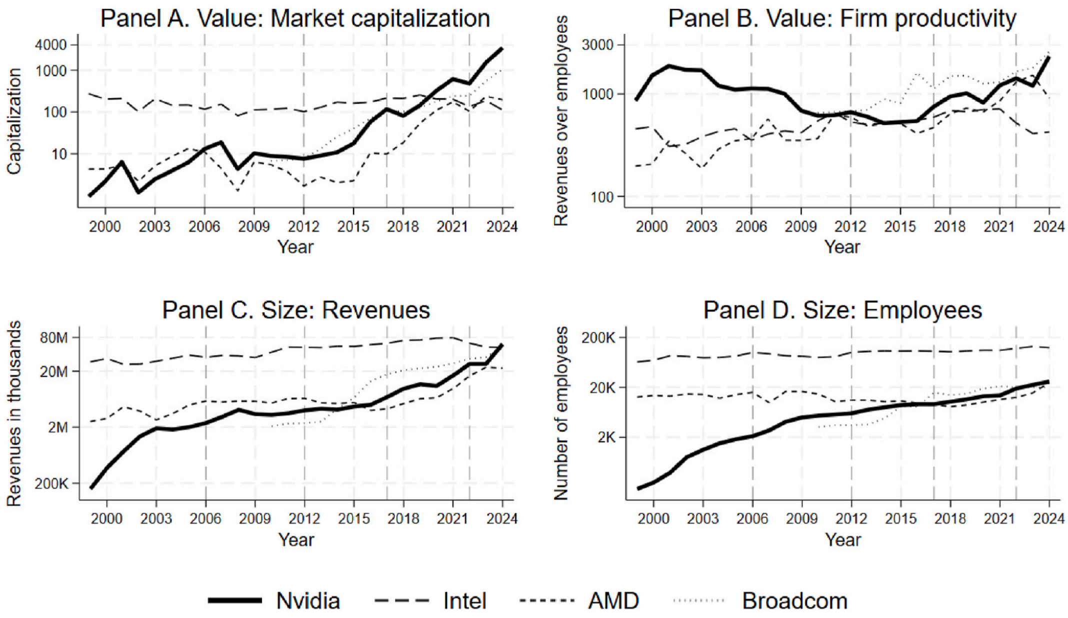

Fig. 2. Size and value growth indicators.
Note: All analysis and graphs are based on the author's elaboration utilizing FA1 and FA2. Due to the exponential growth exhibited by Nvidia, all graphs have been plotted on a logarithmic scale to enable meaningful comparisons with its competitors. In a normal scale, Nvidia's growth in metrics such as market capitalization appears disproportionately large, hindering direct comparisons with other firms.

4.4. Phase 3: strategic renewal (2017–2022)

While Nvidia's video graphics GPU business has remained stable in terms of revenue, AI GPUs have boomed, now attracting most of the company's strategic attention. Unlike its graphics strategy, Nvidia has had to reinvent the way it addresses the AI computing industry (Huang, IK1), shifting from being a mere designer and supplier of components to orchestrating the entire industry. Nvidia began to engage with software developers to get them to see the opportunities of AI-enhanced applications, and from there gradually put all the building blocks in place—the hardware, the associated development software, user-facing applications, and services to simulate the physical world via an astonishingly high-performance GPU architecture (DO3). Nvidia extended its reach from the manufacture of commodity graphics cards focused on the gaming industry to enter other markets, from enterprise data centres to scientific computing. "*We knew that we had to go beyond being simple chip designers to being key players in creating the market for AI*" (Huang, IK3).

Nvidia is now much more network. Its customers are at the same time collaborators, or even potential competitors. Through these collaborations Nvidia is not only contributing to the advancement of AI computing, but also to the development of its operating system, infrastructure and, increasingly, its applications, including agentic AI-based robotics that have the capacity to reason (Huang, IK4).

Nvidia has become a one-stop shop for AI data centre hardware. It has strategically segmented its customers (e.g. gamers, crypto and AI data centres) to maximise personalisation, value, and revenue (DO2). These customers are no longer just buying GPUs; they are buying bundled solutions, including architecture, systems, data centres, CUDA, CUDA-X, and more. By combining its products in this manner, Nvidia has been able to generate 66 % gross margins (DO3, REP). "*The important thing about our software is that it's built on top of our platform. It means that it activates all of Nvidia's hardware chips and system platforms*" (Huang, IK3). Nvidia's ultimate ambition is to bring accelerated computing to everyone, partnering with others to provide all the hardware, software, and service solutions for any computing workload to run as efficiently as possible.

4.5. Phase 4: scalability (2022–2025)

Nvidia's latest Blackwell GPUs are up to 50,000 times faster than its first G-force GPU, and this capacity for acceleration is expected to continue exponentially as Nvidia harnesses its own AI to generate reinforcement learning loops based on synthetic data generation (Huang, IK4).

Largely as a result of this, Nvidia's sales revenue rose by 126 % in 2024 (FA2), and its stock price soared up by 2200 % between 2019 and 2024 (FA1), exemplifying how incumbent organizations can successfully navigate the capability development stages required to achieve the necessary conditions for scalability.

Jensen Huang attributes this success to the fact that Nvidia "*were the only ones who got*" the potential of AI (CS5, BO1). Nvidia worked long and hard to create the necessary scalability capabilities to reach the point where they are now, at the forefront of an industry in which its competitors could still be playing catch-up for some time (REP, DO2, BO2).

5. Theoretical model

From the case narrative analysis, a theoretical model emerged that depicts the process that incumbent firms need to follow to the achieve the same kind of scalability potential as start-ups (see Table 2). This model comprises four stages that reflect how technological shifts transform decision-making within the organization: Recognition, Adaptation, Renewal and Scale-Up. There are also three transition phases (Discovery, Development, Exploitation), which we shall now explore.[4]

The Discovery Process: Based on the Nvidia case, we theorize that the transition from opportunity recognition to organizational adaptability can be described as the discovery process (Forsgren, 2016). The proactive ability of an organization to identify potentially favourable conditions in an evolving value system, such as a technological shift or a change in consumer preferences, is what (Baron, 2006) calls 'opportunity recognition'. This is considered crucial for scalability due to its direct connection to dynamic capitalization on growth potential to create value-generating ventures, whereby firms swiftly adapt their resources and processes to remain competitive (Ardichvili et al., 2003; Teece et al., 1997). Firms that fail to recognise such industry shifts often struggle to adapt, while those that do are able to quickly adjust their organizational structures in response (Penrose, 1959; Teece, 2007).

Nvidia's early investment in AI-compatible GPUs during the CUDA development era is a prime example of such strategic foresight. By onboarding new knowledge, the company was able to update its organizational practices and policies in order to develop innovative new capabilities, and consequently enhance both its internal and external adaptability (Zahra et al., 2006).

It is this ability to anticipate changes in an evolving value system and therefore identify new scale-up opportunities that endows firms with the flexibility required for more effective organizational adaptation of their processes and structures (Vaillant and Lafuente, 2019) in response.

Evolutions in an industry's value system not only affect where and how firms compete (Massa et al., 2017), but also what, where, when, why, how, and how much consumers purchase. As a result, firms must proactively align their competencies with the changing key success factors of both industry and market (Porter, 1979). This is precisely what Nvidia did, and was thus able to transition from a graphics-focused company to a fully-fledged orchestrator of the AI computing industry. As Jacobides (2005) argues, firms that recognise such changes in the value system are better equipped to adapt their own boundaries and structures to them.

The Development Process: We now examine the second transition, from organizational adaptability to strategic renewal, which can be conceptualized as the development process (Forsgren, 2016). In order for incumbent firms to embark on a scalable strategic trajectory, the mere identification of an opportunity or will to act upon it are insufficient (Sirmon et al., 2011). Entire structures and boundaries must be reconfigured (Jacobides and Billinger, 2006), which can be problematic for older firms, which often struggle to depart from long-established routines. Organisational adaptation is therefore essential before these mature firms can fully engage in transitional value systems. Nvidia's reorientation of its organizational strategies towards AI from 2012 onwards exemplifies this ability to act upon perceived opportunities by recombining the required resources and reconfiguring the firm's production structures (McKelvie and Wiklund, 2010; Penrose, 1959).

Nvidia not only transformed its business model but also created the necessary structural tracks for scaling in the AI industry. As strategy is typically path-dependent and closely intertwined with existing architectures, mature firms often have to overcome bureaucratic inertia before they can engage in effective strategic change (McKelvie and Wiklund, 2010; Sirmon et al., 2011). But if they do manage to make the right adaptations, they will be able to boost their productive agility and redefine their capability boundaries, and thus be in a position to truly scale their business models (Teece et al., 1997). To be able to successfully implement a scalable business model in response to changing value opportunities within its industry, incumbent firms must first lay down the adapted structural tracks upon which its renewed strategic path will scale.

In short, incumbent firms need to address issues of organizational complexity, inertia, and undue scope before they can develop scalable business models (Carnes et al., 2017). It is not enough for a firm's technology to be scalable, the entire organisation must be able to provide the conditions for that scalability (Stampfl et al., 2013). This makes them more able to fend off commoditization and segregate markets as they transition toward emerging value systems. Firms that fail to do so risk launching scalable strategies that their existing architectures are unable to support.

The Exploitation Process: We now examine the transition from strategic renewal to scalability conditions, which has been conceptualized as the exploitation process (Forsgren, 2016). This transition underscores the importance of aligning strategic flexibility with business growth. As Stampfl et al. (2013, pp. 229) note, "*today, it is no longer a matter of big companies outperforming small start-ups; rather, those firms which grow fast and are responsive to change end up winning the game.*" Consequently, strategic flexibility emerges as a key predictor of growth for incumbent firms (Gilbert et al., 2006; Demir et al., 2017). Nvidia's transition into an industry orchestrator through the creation of a comprehensive platform for AI hardware and software integration illustrates its strategic flexibility and alignment with growth potential. Scalability, in turn, is facilitated by a firm's strategic business model (Stampfl et al., 2013; Zhang et al., 2015; Kohler, 2018) which, if effective, not only pre-orients a firm for scaling but also ensures that scalability is achieved through the integration of strategic renewal with organizational and technological adaptations.

Visnjic et al. (2022, pp. 76) observed how incumbents had to "*embed then scale out.*" Technologies, cost and revenue structures, internal configurations, and scale-adapted boundary conditions must all be aligned with clearly attuned strategic paths and business models to support sustained growth (Stampfl et al., 2013). Nvidia scaled its own AI strategy by renewing its business model and

[4] In our case study, Recognition is exogenous: Nvidia read macro-industry signals—specifically, the anticipated surge in demand for parallel computing—and invested accordingly. Recognition could, of course, be endogenous and thus more disruptive, arising from the creation of a novel market or use; this is not the situation here, as Nvidia did not forge a new industry but foresaw an evolution within the existing computing sector.

Table 2
Theoretical Model and connection to case study.

	PHASE 1	DISCOVERY	PHASE 2	DEVELOPMENT	PHASE 3	EXPLOITATION	PHASE 4
THEORETICAL PATH	**Value Recognition**	→	**Adaptation**	→	**Strategic Renewal**	→	**Scalability**
Construct Definition	Identifying and leveraging new trends and opportunities in the value system.		Making organizational structural changes to align with emerging opportunities.		Reconfiguring strategy to establish new revenue streams and business models through alliances and innovation.		Building capabilities to sustain and exploit exponential growth.
Practical Action	Analyse market shifts and emerging technologies; identify opportunities and align organizational focus.		Implement structural changes (e.g., creating new divisions, launching products) to enhance adaptability.		Form partnerships, diversify offerings, and implement innovative strategies for sustained competitive advantage.		Optimize operations, leverage strategic positioning, and scale resources to capitalize on growth opportunities.
CASE STUDY: NVIDIA	**Recognising Nvidia GPUs importance for AI revolution**	→	**Transformational change towards adapting to AI**	→	**Segmentation and business model innovation**	→	**Exponential Market Growth**
Periodization	2006–2012		2012–2017		2017–2022		2022–2024
Exemplary Actions	−2006: Work on CUDA parallel computing platform was initiated to significantly accelerate processing.		−2014: Launched CUDA Deep Neural Network Library.		Products now bundled together under the Nvidia platform and sold as solutions.		Achieved record market valuation; June 2024, became largest market-cap company in the world.
	−2012: CUDA powers the first instance of deep machine learning.		−2016: Released NVIDIA DGX platform for AI model training.		New one-stop shop strategy for AI data centre hardware.		Orchestrator of the generational AI transformation.
	−2012: Leveraged GPUs for deep neural networks, enabling scope economies.		−2017: Developed Volta series, furthering advancements in deep learning.		Expanded into wider industries, from gaming, to enterprise data centres, to scientific computing.		Essential supplier to the surge of new players using or offering AI-based applications

engaging with developers and partners in line with its growth objectives. Had it not done so, incoherences would have generated resistance to scaling (Visnjic et al., 2022).

While traditional hierarchical companies can usually accommodate a small number of minor strategic incoherences thanks to the self-correcting effects of established organisational routines and inertia (Rigby et al., 2016), that same inertia often has the opposite effect in the case of incumbent firms. Organisational routines can (and usually will) produce all kinds of 'antibodies' that hinder scale-up unless management is able to instil a compatible strategic path and successfully communicates its renewed values and principles throughout the enterprise (Rigby et al., 2016).

Zhang et al. (2015) found that scalability is largely dependent on three main dimensions of business model design: customer identification, customer engagement, and value chain linkages. Customer identification and market positioning have a significant effect on scale, mostly through network effects, whereby proper targeting can increase the size of the network related to a new product/service, ultimately raising its value, as well as the potential for learning-by-using benefits (Zhang et al., 2015). Business models that generate network effects also foster 'lock-in', whereby greater switching costs dissuade customers from migrating to competitors (Amit and Zott, 2001; Stampfl et al., 2013). Similarly, embedding products in broader value systems and technological infrastructures generates the potential for scale through technological interrelatedness (Vendrell-Herrero et al., 2022), whereby a wider range of sub-technologies become part of the firm's infrastructure, further increasing adoption and scale (Zhang et al., 2015).

With digitalisation, high-value customization becomes more scalable through co-creation and customer self-customization. By transferring part of the production process to the customer, the cost of satisfying specific or expensive needs and tastes is minimized (Zhang et al., 2015). Better customer engagement improves their understanding of the technology, resulting in a larger number of users and a greater scale potential.

Network governance, such as platforms compatible with value systems, has also been observed to promote scalability (Jacobides et al., 2021; Stampfl et al., 2013). While traditional hierarchical models tend to hinder scalability (Zhang et al., 2015), networked value chains can facilitate it by reaching new customers who can be serviced without having to make major investments in enlarging capacity. Open governance models that adopt the principles of modular design and combinatorial innovation are therefore better equipped to reduce value generation costs while engaging broader communities of users (Yoo et al., 2012).

Again, the case of Nvidia Corporation illustrates how incumbent firms pursuing scalability must first make the corresponding adjustments to their strategic trajectories. This often entails adoption of the right customer targets and engagement tactics, along with suitable value chain linkages.

Achieving Scalability Conditions: In summary, for scalability conditions to take shape within established firms, any prior strategic trajectory must be redirected towards a more growth-compatible orientation through the implementation of a scalable business model (Demir et al., 2017). However, if such a strategic renewal is to be successfully executed, firms must first demonstrate organizational adaptability through the reconfiguration of their internal structures and boundary designs, so as to exploit scale-up opportunities arising from value system reconfiguration and foster the necessary agility for renewal and scalability (Belitski et al., 2023). But before they can do any of this, incumbent firms must be able to recognise and position themselves within new value boundary conditions that offer potential scale-up opportunities (Zott and Amit, 2013; Mithani, 2023).

In short, the key capabilities for incumbent companies to achieve scalability include i) an aptitude for recognising valuable scaling opportunities, ii) a capacity for consequent organisational adaptation, and iii) competence for strategic renewal.

6. Discussion and implications

6.1. Key insights and contributions

This paper set out to explore how established organizations can develop the capabilities that underpin scalability and to use the insights to propose a theoretical model. By observing the evolution of Nvidia Corporation as a prime example of an incumbent firm that has successfully attained scalability, and comparing its experience with theorizations in the relevant literature, we were able to produce a novel multi-phase model, as shown in Table 2, that describes the individually exclusive and progressive phases involved in successful transition by incumbents into scale-up organizations. These include value recognition, organisational adaptability, and strategic renewal.

Our study contributes to academia by addressing gaps in the understanding of how scalability is achieved, as observed by McKelvie and Wiklund (2010) over a decade ago, and more recently by Shepherd and Patzelt (2022) in their remark that the 'how' question of firm growth is still poorly understood. Ongoing factors such as the digital and AI transition affecting all aspects of the economy have heightened the need to break down the mechanisms by which incumbent companies in particular can scale up and remain competitive. Our findings also respond to the calls by Coviello et al. (2024) for better managerial insight on the achievement of scalability, by Stampfl et al. (2013) for further research into the requirements for strategic scalability in the context of digital transition, by Shepherd and Patzelt (2022) concerning management's role in organisational scalability, and by Piaskowska et al. (2021) with specific regard to mature and incumbent firms.

The paper also helps to extend scalability research from its traditional domain of entrepreneurship (Shepherd and Patzelt, 2022) towards the strategies of more established organizations, thus generalizing the concept and introducing fresh nuances on the organizational changes required. The disruptive nature of scalability often means that it becomes easier to start a new venture from scratch rather than adapting an existing firm. While entrepreneurial scalability tends to directly stem from the identification of opportunities and the capacity to generate value, incumbent firms require major resource adaptation, flexible capabilities, and strong alignment with external expansion needs (Penrose, 1959; Barney, 1991; Teece et al., 1997). Our study helps clarify how incumbent firms should adapt organisationally and strategically to such disruption.

Beyond these contributions, a collateral finding from our analysis of Nvidia's trajectory is the dual decision-making logic that appears to have played a key part in its scale-up success. Although Nvidia can be classed as a large firm by most metrics, its decision-making under uncertainty has often resembled that of smaller entrepreneurial start-ups. As suggested by Berends et al. (2014), simultaneous managerial causation (goal-driven decision-making typically associated with resource-rich and structurally-sound corporations in stable operating environments) and entrepreneurial effectuation (means-driven decision-making typically associated with entrepreneurial start-ups in unpredictable environments) can coexist, and when they do so, contribute to more effective innovation processes. Similarly, Matalamäki et al. (2017) found that established firms experiencing successful growth often combined instances of effectual and causal logic. Indeed, Nvidia demonstrated clear goal-driven attitudes and focus on predicting an uncertain future, which are clearly characteristics of a causation-based decision-making logic. However, the firm and its CEO also exploited emergent contingencies while building key strategic alliances, which align with effectuation (Sarasvathy, 2001). Although simultaneous, effectual logic was more prominent during the early stages of Nvidia's scalability, while causation was more dominant in the later stages, consistent with the sequence reported by Matalamäki et al. (2017).

The theoretical model presented in this paper is grounded in a longitudinal analysis of the real-world example of Nvidia's exponential growth, offering a comprehensive illustration of how an established firm can successfully develop its capabilities to achieve scalability, thereby providing valuable insights for scholars in this field. While some may critique the choice of Nvidia, particularly regarding the role of luck in exceptional business success (Denrell and Fang, 2010), this study accounts for such concerns by critically examining both internal and external factors that contributed to its scalability. While AI posed an opportunity for everyone, established firms like Intel and AMD failed to capitalize on it, while start-ups like Broadcom did. Notably, Nvidia's scalability trajectory aligns more closely with Broadcom's than with its incumbent peers, suggesting that Nvidia's growth was not merely a matter of luck but the result of deliberate strategic decisions.

6.2. Managerial implications

The article's central message for managers is that established firms can achieve exponential growth, but doing so requires rethinking their market positioning as if they were new entrants. This involves engaging with new value systems, implementing organizational adaptation mechanisms, and, critically, devising strategies aligned with emerging opportunities and the evolving nature of the firm. These factors are all essential for developing the necessary growth capabilities.

This study's findings help to clarify the journey that incumbent firms need to take in order to adapt organisationally and strategically to the major systemic and technologically-driven disruptions that are affecting most areas of the economy. Our proposed theoretical model outlines the underlining capability conditions that incumbents need to achieve on the road to effective scalability. These are presented as a series of steps, and scalability will not be achievable without negotiating each one of these in sequence. In other words, an incumbent firm cannot fulfil its scalability requirements if it does not first renew its strategy, which it will not be able to implement unless it first adapts its organisational structure and boundary architecture. But doing so would be meaningless if it is not preceded by the recognition of valuable scale-up opportunities within the firm's industry.

Overall, our findings offer major potential to influence thinking and practice in organizational regeneration and technological adoption. This is especially important in an era of digital transformation of all aspects of production, business and society in general (Menz et al., 2021). Digitalisation and AI are no longer niche areas confined to certain sectors and industries, but are destined to become mainstream, if they have not done so already. Hence, firms must reconsider the manner in which they are generating value and rethink their corporate and business strategies, redefining where and how they compete.

6.3. Limitations and future research

This paper presents a theoretical model derived from Nvidia's growth trajectory to explain how established organizations can build scalability capabilities. While the framework highlights deliberate strategic actions, it also identifies two key areas for future research.

First, the role of serendipity-based alternative explanations shaping scalability warrants further exploration. Second, the framework's conceptual nature calls for quantitative validation through operationalized constructs, large-scale datasets, or cross-industry analyses to test its robustness and generalizability. Addressing these limitations would enhance both the theoretical and practical relevance of the model for organizations facing scalability challenges.

The role of alternative explanations. While the development stages we identify—value recognition, organizational adaptability, and strategic renewal—offer a clear and actionable framework for managers seeking to build scalability capabilities, it is important to acknowledge that Nvidia's trajectory may also be explained by alternative factors. These explanations do not contradict our model but suggest that deliberate strategy alone cannot fully account for the firm's scalability; contingency, path dependence, and emergent forces also played a role.

One such explanation is the role of emergent strategy (e.g., see Mintzberg and Waters, 1985; Burgelman, 1994; Mirabeau et al., 2018). Although CUDA was introduced in 2006 with a vision of general-purpose GPU computing, the AI use case only materialized years later, catalysed externally through breakthroughs like AlexNet. The success of CUDA in AI was not the result of a tightly controlled strategic plan, but rather the product of an enabling architecture that was later appropriated by developers in unanticipated ways. Huang himself has described CUDA's use in AI as a "Big Bang" moment, reflecting how pivotal outcomes often emerge from experimentation and external discovery rather than deliberate sequencing (also see the earlier discussion in section 6.1 on NVIDIA's dual effectuation-causation decision-making logic).

Luck and chance (e.g., see De Rond and Thietart, 2007; Liu and De Rond, 2016) could also be seen as significant factors. The decision to invest heavily in AI chips occurred well before AI markets were validated. Nvidia's delivery of the first AI supercomputer to OpenAI now appears as a major strategic milestone, but at the time, it was a speculative and risky move. Jensen Huang has acknowledged that many of Nvidia's transformative moments involved "luck founded on vision," suggesting that even visionary leadership often operates in highly uncertain, ambiguous contexts.

Pre-adaptation and shadow options (e.g., see Cattani, 2006; Andriani and Cattani, 2024) are additional elements to consider. Nvidia's early dominance in gaming graphics endowed it with deep capabilities in parallel computing, developer ecosystem building, and hardware acceleration—competences that later proved critical in AI. The firm did not need to build these from scratch; instead, it repurposed and extended existing capabilities. CUDA, for example, was initially targeted at scientific computing and financial modelling, yet ultimately found a new trajectory in AI development.

Finally, starting conditions matter. Despite the mentioned constraints limiting the scalability of incumbent firms, such as their infrastructural, managerial, and strategic legacies that often give vantage to start-ups, not all of Nvidia's competencies hampered its scale-up potential. Nvidia was already a large, successful, and R&D-intensive firm with strong IP control, and sufficient capital support to pursue long-term bets like CUDA and DGX. Unlike smaller firms and those in less technology-intensive industries, Nvidia could afford to build platform capabilities over many years without immediate commercial return. This underscores the enabling role of existing scale and legitimacy in making such long-term investments viable.

Taken together, these alternative explanations point to a broader understanding of scalability—not only as a structured managerial process, but also as an outcome shaped by chance, history, and adaptive responsiveness. While our phased framework offers valuable guidance for replication, Nvidia's success also reflects a confluence of factors beyond pure deliberate action. These alternative explanations—emergence, luck, pre-adaptation, and firm-level endowments—highlight important areas for further research. Future studies could explore how these elements interact with managerial intent, and under what conditions they enable or constrain scalability in other firms and industries.

Quantitative validation. This study shares the inherent limitations of qualitative research. While its inductive and longitudinal nature enhances theoretical understanding, it does not allow for broad empirical testing across multiple firms. Future research should address this gap, perhaps using survey data to test its key constructs of recognition, adaptation, renewal, and scalability, which are not directly observable.

Future studies could also employ mediation models, for instance to examine adaptation as a mediator between recognition and renewal, and renewal as a mediator between adaptation and scalability. Additionally, path dependency analysis may offer valuable insights.

We conceptualize scalability as a continuum rather than a dichotomy, and as a capability rather than realized growth. Consequently, future research could test these relationships not only within high-growth firms but also among established companies with moderate growth, aligning more closely with our focus on incumbent firms.

CRediT authorship contribution statement

Ferran Vendrell-Herrero: Writing – review & editing, Writing – original draft, Visualization, Supervision, Methodology, Investigation, Formal analysis, Conceptualization. **Yancy Vaillant:** Writing – review & editing, Writing – original draft, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Oscar F. Bustinza:** Writing – review & editing, Writing – original draft, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used Deepl and ChatGPT to revise the manuscript and improve language and readability. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Acknowledgements

Ferran Vendrell-Herrero acknowledges support by the project PID2022-136235NB-I00 funded by MICIU/AEI/10.13039/501100011033 and by ERDF/EU. Oscar F. Bustinza acknowledges support from the Ministry of Universities of Spain within the framework of the State Program to Develop, Attract and Retain Talent, State Mobility Subprogram, of the State Plan for Scientific, Technical and Innovation Research 2021–2023 (Reference: PRX22/00176).

Data availability

This research relies on freely accessible digital archives, with a source list (Table 1) providing direct links.

References

Amit, R., Zott, C., 2001. Value creation in e-business. Strateg. Manag. J. 22, 493–520. https://doi.org/10.1002/smj.187.

Andriani, P., Cattani, G., 2024. Toward an evolutionary theory of shadow options: emergent functions of Corning's glass-based innovations. Organ. Sci. 35, 2249–2271.
Ardichvili, A., Cardozo, R., Ray, S., 2003. A theory of entrepreneurial opportunity identification and development. J. Bus. Ventur. 18, 105–123.
Argyres, N.S., De Massis, A., Foss, N.J., Frattini, F., Jones, G., Silverman, B.S., 2020. History-informed strategy research: the promise of history and historical research methods in advancing strategy scholarship. Strateg. Manag. J. 41, 343–368. https://doi.org/10.1002/smj.3118.
Autor, D., Dorn, D., Katz, L.F., Patterson, C., Van Reenen, J., 2020. The fall of the labor share and the rise of superstar firms. Q. J. Econ. 135, 645–709.
Barney, J., 1991. The resource-based model of the firm: origins, implications, and prospects. J. Manag. 17, 97–98. https://doi.org/10.1177/014920639101700107.
Baron, R.A., 2006. Opportunity recognition as pattern recognition: how entrepreneurs "connect the dots" to identify new business opportunities. Acad. Manag. Perspect. 20, 104–119. https://doi.org/10.5465/amp.2006.19873412.
Belitski, M., Martin, J., Stettler, T., Wales, W., 2023. Organizational scaling: the role of knowledge spillovers in driving multinational enterprise persistent rapid growth. J. World Bus. 58, 101461.
Berends, H., Jelinek, M., Reymen, I., Stultiëns, R., 2014. Product innovation processes in small firms: combining entrepreneurial effectuation and managerial causation. J. Prod. Innovat. Manag. 31, 616–635. https://doi.org/10.1111/jpim.12117.
Bohan, S., Tippmann, E., Levie, J., Igoe, J., Bowers, B., 2024. What is scaling? J. Bus. Ventur. 39, 106355.
Burgelman, R.A., 1994. Fading memories: a process theory of strategic business exit in dynamic environments. Adm. Sci. Q. 24–56.
Carnes, C.M., Chirico, F., Hitt, M.A., Huh, D.W., Pisano, V., 2017. Resource orchestration for innovation: structuring and bundling resources in growth- and maturity-stage firms. Long. Range Plan. 50, 472–486. https://doi.org/10.1016/j.lrp.2016.07.003.
Cattani, G., 2006. Technological pre-adaptation, speciation, and emergence of new technologies: how Corning invented and developed fiber optics. Ind. Corp. Change 15, 285–318.
Coviello, N., Autio, E., Nambisan, S., Patzelt, H., Thomas, L.D., 2024. Organizational scaling, scalability, and scale-up: definitional harmonization and a research agenda. J. Bus. Ventur. 39, 106419.
Davidsson, P., Henrekson, M., 2002. Determinants of the prevalance of start-ups and high-growth firms. Small Bus. Econ. 19, 81–104. https://doi.org/10.1023/A:1016264116508.
De Rond, M., Thietart, R., 2007. Choice, chance, and inevitability in strategy. Strateg. Manag. J. 28, 535–551. https://doi.org/10.1002/smj.602.
Decker, S., 2022. Introducing the eventful temporality of historical research into international business. J. World Bus. 57, 101380.
Demir, R., Wennberg, K., McKelvie, A., 2017. The strategic management of high-growth firms: a review and theoretical conceptualization. Long. Range Plan. 50, 431–456.
Denrell, J., Fang, C., 2010. Predicting the next big thing: success as a signal of poor judgment. Manag. Sci. 56, 1653–1667. https://doi.org/10.1287/mnsc.1100.1220.
Dosi, G., Nelson, R.R., 1994. An introduction to evolutionary theories in economics. J. Evol. Econ. 4, 153–172. https://doi.org/10.1007/BF01236366.
Eisenhardt, K.M., Graebner, M.E., 2007. Theory building from cases: opportunities and challenges. Acad. Manag. J. 50, 25–32. https://doi.org/10.5465/amj.2007.24160888.
Forsgren, M., 2016. A note on the revisited Uppsala internationalization process model–the implications of business networks and entrepreneurship. J. Int. Bus. Stud. 47, 1135–1144.
Gilbert, B.A., McDougall, P.P., Audretsch, D.B., 2006. New venture growth: a review and extension. J. Manag. 32, 926–950. https://doi.org/10.1177/0149206306293860.
Giustiziero, G., Kretschmer, T., Somaya, D., Wu, B., 2023. Hyperspecialization and hyperscaling: a resource-based theory of the digital firm. Strateg. Manag. J. 44, 1391–1424. https://doi.org/10.1002/smj.3365.
Golden-Biddle, K., Locke, K., 2006. Composing Qualitative Research. Sage Publications.
Jacobides, M.G., 2005. Industry change through vertical disintegration: how and why markets emerged in mortgage banking. Acad. Manag. J. 48. https://doi.org/10.5465/AMJ.2005.17407912.
Jacobides, M.G., Billinger, S., 2006. Designing the boundaries of the firm: from "make, buy, or ally" to the dynamic benefits of vertical architecture. Organ. Sci. 17, 171–312. https://doi.org/10.1287/orsc.1050.0167.
Jacobides, M.G., Brusoni, S., Candelon, F., 2021. The evolutionary dynamics of the artificial intelligence ecosystem. Strategy Science 6 (4), 412–435.
Jacobides, M.G., Cennamo, C., Gawer, A., 2018. Towards a theory of ecosystems. Strateg. Manag. J. 39, 2255–2276. https://doi.org/10.1002/smj.2904.
Jansen, J.J.P., Heavey, C., Mom, T.J.M., Simsek, Z., Zahra, S.A., 2023. Scaling-up: building, leading and sustaining rapid growth over time. Journal of Management Studies 60, 581–604. https://doi.org/10.1111/joms.12910.
Joseph, J., Wilson, A.J., 2018. The growth of the firm: an attention-based view. Strateg. Manag. J. 39, 1779–1800. https://doi.org/10.1002/smj.2715.
Karsten, L., 2014. Time as a periodization of management practices. Manag. Organ. Hist. 9, 414–432. https://doi.org/10.1080/17449359.2014.980270.
Kipping, M., Üsdiken, B., 2014. History in organization and management theory: more than meets the eye. Acad. Manag. Ann. 8, 535–588. https://doi.org/10.5465/19416520.2014.911579.
Kohler, T., 2018. How to scale crowdsourcing platforms. Calif. Manag. Rev. 60, 98–121. https://doi.org/10.1177/0008125617738261.
Kuratko, D.F., Holt, H.L., Neubert, E., 2020. Blitzscaling: the good, the bad, and the ugly. Bus. Horiz. 63, 109–119.
Langley, A., 1999. Strategies for theorizing from process data. Acad. Manag. Rev. 24, 691. https://doi.org/10.2307/259349.
Langley, A., Smallman, C., Tsoukas, H., Van De Ven, A.H., 2013. Process studies of change in organization and management: unveiling temporality, activity, and flow. Acad. Manag. J. 56, 1–13. https://doi.org/10.5465/amj.2013.4001.
Liu, C., De Rond, M., 2016. Good night, and good luck: perspectives on luck in management scholarship. Acad. Manag. Ann. 10, 409–451.
Mason, C., Brown, R., 2013. Creating good public policy to support high-growth firms. Small Bus. Econ. 40, 211–225. https://doi.org/10.1007/s11187-011-9369-9.
Massa, L., Tucci, C.L., Afuah, A., 2017. A critical assessment of business model research. Acad. Manag. Ann. 11, 3–104. https://doi.org/10.5465/annals.2014.0072.
Matalamäki, M., Vuorinen, T., Varamäki, E., Sorama, K., 2017. Business growth in established companies; roles of effectuation and causation. J. Enterprising Cult. 25, 123–148. https://doi.org/10.1142/S0218495817500054.
McKelvie, A., Wiklund, J., 2010. Advancing firm growth research: a focus on growth mode instead of growth rate. Enterpren. Theor. Pract. 34, 261–288. https://doi.org/10.1111/j.1540-6520.2010.00375.x.
Menz, M., Kunisch, S., Birkinshaw, J., Collis, D.J., Foss, N.J., Hoskisson, R.E., Prescott, J.E., 2021. Corporate strategy and the theory of the firm in the digital age. Journal of Management Studies 58, 1695–1720. https://doi.org/10.1111/joms.12760.
Mihailova, I., 2023. Business model adaptation for realized international scaling of born-digitals. J. World Bus. 58, 101418.
Mintzberg, H., Waters, J.A., 1985. Of strategies, deliberate and emergent. Strateg. Manag. J. 6, 257–272.
Mirabeau, L., Maguire, S., Hardy, C., 2018. Bridging practice and process research to study transient manifestations of strategy. Strateg. Manag. J. 39, 582–605. https://doi.org/10.1002/smj.2732.
Mithani, M.A., 2023. Scaling digital and non-digital business models in foreign markets: the case of financial advice industry in the United States. J. World Bus. 58, 101457.
Motley, D.C., Eesley, C.E., Koo, W., 2023. Born into chaos: how founding conditions shape whether ventures survive or thrive when experiencing environmental change. Strateg. Entrep. J. 17, 408–444. https://doi.org/10.1002/sej.1461.
Nelson, R.R., Winter, S.G., 1982. The Schumpeterian tradeoff revisited. Am. Econ. Rev. 72, 114–132.
Nix, A., Decker, S., 2023. Using digital sources: the future of business history? Bus. Hist. 65, 1048–1071. https://doi.org/10.1080/00076791.2021.1909572.
Ocasio, W., 2011. Attention to attention. Organ. Sci. 22, 1286–1296. https://doi.org/10.1287/orsc.1100.0602.
Penrose, E.T., 2009. Excerpt from the theory of the growth of the firm. In: Knowledge and Strategy. Routledge, London, pp. 63–67.
Penrose, E.T., 1959. The Theory of the Growth of the Firm. Oxford University Press.
Piaskowska, D., Tippmann, E., Monaghan, S., 2021. Scale-up modes: profiling activity configurations in scaling strategies. Long. Range Plan. 54, 102101.
Porter, M.E., 1979. The structure within industries and companies' performance. Rev. Econ. Stat. 61. https://doi.org/10.2307/1924589, 214–214.

Rawhouser, H., Webb, J.W., Rodrigues, J., Waldron, T.L., Kumaraswamy, A., Amankwah-Amoah, J., Grady, A., 2022. Scaling, blockchain technology, and entrepreneurial opportunities in developing countries. J. Bus. Ventur. Insights 18, e00325.
Reuber, A.R., Tippmann, E., Monaghan, S., 2021. Global scaling as a logic of multinationalization. J. Int. Bus. Stud. 52, 1031–1046. https://doi.org/10.1057/s41267-021-00417-2.
Ridder, H., Hoon, C., McCandless Baluch, A., 2014. Entering a dialogue: positioning case study findings towards theory. Br. J. Manag. 25, 373–387. https://doi.org/10.1111/1467-8551.12000.
Rigby, D.K., Sutherland, J., Takeuchi, H., 2016. Embracing agile. Harv. Bus. Rev. MAY 2016.
Rowlinson, M., Hassard, J., Decker, S., 2014. Research strategies for organizational history: a dialogue between historical theory and organization theory. Acad. Manag. Rev. 39, 250–274. https://doi.org/10.5465/amr.2012.0203.
Sanasi, S., 2023. Entrepreneurial experimentation in business model dynamics: current understanding and future opportunities. Int. Entrep. Manag. J. 19, 805–836. https://doi.org/10.1007/s11365-023-00836-7.
Sarasvathy, S.D., 2001. Causation and effectuation: toward a theoretical shift from economic inevitability to entrepreneurial contingency. Acad. Manag. Rev. 26, 243. https://doi.org/10.2307/259121.
Shepherd, D.A., Patzelt, H., 2022. A call for research on the scaling of organizations and the scaling of social impact. Entrep. Theory Pract. 46, 255–268. https://doi.org/10.1177/1042258720950599.
Siggelkow, N., 2007. Persuasion with case studies. Acad. Manag. J. 50, 20–24. https://doi.org/10.5465/amj.2007.24160882.
Sirmon, D.G., Hitt, M.A., Ireland, R.D., Gilbert, B.A., 2011. Resource orchestration to create competitive advantage: breadth, depth, and life cycle effects. J. Manag. 37. https://doi.org/10.1177/0149206310385695.
Sleuwaegen, L., Ramboer, S., 2020. Regional competitiveness and high growth firms in the EU: the creativity premium. Appl. Econ. 52. https://doi.org/10.1080/00036846.2019.1686454.
Snihur, Y., Zott, C., 2020. The genesis and metamorphosis of novelty imprints: how business model innovation emerges in young ventures. Acad. Manag. J. 63, 554–583. https://doi.org/10.5465/amj.2017.0706.
Somaya, D., You, J., 2024. Scalability, venture capital availability, and unicorns: evidence from the valuation and timing of IPOs. J. Bus. Ventur. 39, 106345.
Stampfl, G., Prügl, R., Osterloh, V., 2013. An explorative model of business model scalability. Int. J. Prod. Dev. 18. https://doi.org/10.1504/IJPD.2013.055014.
Teece, D.J., 2007. Explicating dynamic capabilities: the nature and microfoundations of (sustainable) enterprise performance. Strateg. Manag. J. 28, 1319–1350. https://doi.org/10.1002/smj.640.
Teece, D.J., Pisano, G., Shuen, A., 1997. Dynamic capabilities and strategic management. Strateg. Manag. J. 18, 509–533. https://doi.org/10.1002/(SICI)1097-0266(199708)18:7<509::AID-SMJ882>3.0.CO;2-Z.
Tidhar, R., Hallen, B.L., Eisenhardt, K.M., 2024. Measure twice, cut once: unit profitability, scalability, and the exceptional growth of new firms. Organization Science orsc 2021, 15970. https://doi.org/10.1287/orsc.2021.15970.
Vaara, E., Lamberg, J.-A., 2016. Taking historical embeddedness seriously: three historical approaches to advance strategy process and practice research. Acad. Manag. Rev. 41, 633–657. https://doi.org/10.5465/amr.2014.0172.
Vaillant, Y., Lafuente, E., 2019. The increased international propensity of serial entrepreneurs demonstrating ambidextrous strategic agility: a precursor to international marketing agility. Int. Mark. Rev. 36, 239–259. https://doi.org/10.1108/IMR-01-2018-0015.
Van Alstyne, M.W., Parker, G.G., Choudary, S.P., 2016. Pipelines, platforms, and the new rules of strategy. Harv. Bus. Rev. 94, 54–62.
Vendrell-Herrero, F., Vaillant, Y., Bustinza, O.F., Lafuente, E., 2022. Product lifespan: the missing link in servitization. Prod. Plann. Control 33, 1372–1388.
Visnjic, I., Birkinshaw, J., Linz, C., 2022. When gradual change beats radical transformation. MIT Sloan Manag. Rev. 63, 74–78.
Welch, C., Paavilainen-Mäntymäki, E., Piekkari, R., Plakoyiannaki, E., 2022. Reconciling theory and context: how the case study can set a new agenda for international business research. J. Int. Bus. Stud. 53, 4–26.
Welch, C., Piekkari, R., Plakoyiannaki, E., Paavilainen-Mäntymäki, E., 2020. Theorising from case studies: towards a pluralist future for international business research. In: Eden, L., Nielsen, B.B., Verbeke, A. (Eds.), Research Methods in International Business, JIBS Special Collections. Springer International Publishing, Cham, pp. 171–220. https://doi.org/10.1007/978-3-030-22113-3_9.
Yin, R.K., 2017. Case Study Research and Applications: Design and Methods. Sage publications.
Yoo, Y., Boland, R.J., Lyytinen, K., Majchrzak, A., 2012. Organizing for innovation in the digitized world. Organ. Sci. 23. https://doi.org/10.1287/orsc.1120.0771.
Zahra, S.A., Sapienza, H.J., Davidsson, P., 2006. Entrepreneurship and Dynamic Capabilities: a review, model and research agenda. Journal of Management Studies 43, 917–955. https://doi.org/10.1111/j.1467-6486.2006.00616.x.
Zhang, J.J., Lichtenstein, Y., Gander, J., 2015. Designing scalable digital business models. Adv. Strat. Manag. 33, 241–277. https://doi.org/10.1108/S0742-332220150000033006.
Zott, C., Amit, R., 2013. The business model: a theoretically anchored robust construct for strategic analysis. Strateg. Organ. 11, 403–411. https://doi.org/10.1177/1476127013510466.

Ferran Vendrell-Herrero (PhD) is an Associate Professor at the University of Edinburgh Business School (UK) and Visiting Professor at the University of Vaasa (Finland). His research focuses on the intersection between innovation, strategy, and international business. He has published in world-leading journals such as the Journal of International Business Studies, Journal of World Business, Journal of Product Innovation Management, International Journal of Operations and Production Management, and Long Range Planning, among others.

Yancy Vaillant (PhD, HDR) is a Full Professor in the Department of Strategy, Entrepreneurship, and Innovation at Toulouse Business School (TBS Education). From a strategic management lens, Prof. Vaillant's research focuses on innovation and digital servitization. He is also interested in studying the relevance of digital platforms for business value-creation processes. His research has been published in Regional Studies, Small Business Economics, International Journal of Production Economics, Technovation, among others. He has collaborated with the European, Spanish or Catalan administrations conducting studies in areas related to innovation systems and innovation-based development programs.

Oscar F. Bustinza (PhD) is a Professor of Management at the University of Granada. His work aims to analyse the drivers of firm boundary choices, demand chain management, and service innovation based on data-driven analysis. His research has been published in the Journal of Supply Chain Management, International Journal of Operations & Production Management, Journal of Business Research, International Journal of Production Economics, and British Journal of Management, among other outlets. He has been the Principal Investigator on various funded research projects examining the servitization of business.