From Real Estate to Hybrid (Real Estate + Tech) How Public Markets Reprice Data Center REITs

by

Andrew Oh

Bachelor of Science in Statistics
University College London

Submitted to the MIT Center for Real Estate in partial fulfillment of the requirements for the degree of

Master of Science in Real Estate Development

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2026

Authored by: Andrew Oh
January 16, 2026

Certified by: Professor Manish Srivastava
Visiting Professor, Massachusetts Institute of Technology
Clinical Professor, New York University
Thesis Supervisor

Accepted by: Professor Siqi Zheng
STL Champion Professor of Urban and Real Estate Sustainability
MIT Center for Real Estate + Department of Urban Studies and Planning
Faculty Director, MIT Center for Real Estate (CRE)
Director, MIT Sustainable Urbanization Lab

From Real Estate to Hybrid (Real Estate + Tech)
How Public Markets Reprice Data Center REITs

by

Andrew Oh

Submitted to the MIT Center for Real Estate on January 16, 2026 in partial fulfillment of the requirements for the degree of

Master of Science in Real Estate Development

ABSTRACT

This thesis investigates the evolving pricing identity of U.S. listed data center REITs from 2015 to 2025. Data centers occupy an ambiguous position in public markets: they are structured as real estate investment trusts yet increasingly perceived as technology infrastructure. This study asks whether public market pricing reflects this ambiguity and, if so, how the pricing relationship has changed over time.

Using weekly return data for a two-constituent data center REIT basket (Equinix and Digital Realty), an equal-weight REIT benchmark excluding data centers, and two technology equity ETFs (XLK and SMH), the analysis employs three complementary methods: 52-week rolling correlations, regime-level two-factor identity regressions with Newey-West standard errors, and rolling two-factor betas. The sample is partitioned into four regimes aligned with macroeconomic conditions: pre-COVID baseline (2015–2019), COVID shock and aftermath (2020–2021), rate-hike stress (2022–2023), and post-hike period (2024–2025).

The findings reveal a pronounced identity shift. In the baseline regime, data center REITs exhibited strong REIT alignment and near-zero technology exposure, consistent with traditional real estate pricing. During 2020–2021, REIT alignment collapsed while technology alignment emerged under the broad tech proxy, though not under semiconductors. Post-2022, a hybrid identity crystallized: both REIT and technology coefficients became positive and statistically significant across proxy specifications. Rolling estimates confirm that this hybrid structure is persistent rather than episodic.

The results contribute to understanding how public markets classify assets with hybrid characteristics. The shift from REIT-only to hybrid pricing identity suggests that investors increasingly view data center REITs through a blended real estate and technology lens, with implications for benchmarking, portfolio construction, and the organizational boundary between operating platforms and property ownership. The pricing identity of listed data center REITs has evolved from predominantly real estate to a durable hybrid of real estate and technology.

Thesis Supervisor: Professor Manish Srivastava

Title: Visiting Professor, Massachusetts Institute of Technology

Clinical Professor, New York University

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Why “data center identity” became a real estate question

Data centers have become indispensable physical infrastructure for the digital economy. As digital services expanded, so did demand for facilities that can reliably host compute, storage, and network equipment at scale. Within institutional real estate, this shift also appeared as part of a broader rotation toward “alternate” property types, as investors sought diversification beyond traditional categories [19].

In listed real estate markets, the same expansion has been framed as the rise of “new economy” real estate, where REITs provide investable exposure to the physical assets supporting internet communications and digital activity. Infrastructure and data center REITs are no longer peripheral. They now account for a meaningful share of overall listed REIT market capitalization [11].

At a more basic level, the internet is not placeless. Ownership and control of key physical nodes, including carrier hotels, colocation facilities, and large-scale data centers, shape how digital activity is organized in space and financed through capital markets [12].

This thesis starts from a simple observation: if data centers are increasingly understood as both real assets with location, leasing, and capital intensity **and** as infrastructure nodes embedded in a technology-driven system, then “what a data center is” becomes a valuation question. The core object of study is not technology itself, but **identity as revealed through public market pricing relationships**.

1.2 Two features that distinguish data centers from conventional property

Two features set data centers apart from many traditional property sectors.

First, revenue and contracting depend directly on power, redundancy, and operational services. Data center arrangements can range from retail colocation to wholesale or hyperscale structures, and they often include managed services and “remote hands” functions that resemble operating platform capabilities rather than passive space rental [14]. Rent can also be structured with explicit reference to power availability rather than floor area alone, and leases may require dedicated support areas for backup generation and related equipment [14].

Second, supply is constrained by infrastructure dependencies that are unusually binding for a real estate sector. Barriers to entry include not only high capital expenditure and specialized development expertise, but also the ability to secure network connectivity and substantial, continuous power with reliable backup. This requires coordination with utilities and telecom ecosystems [1].

Because these constraints are physical and locational, “identity” cannot be read from narrative alone. It should be visible in how public markets price risk and growth for the listed vehicles that own and operate these assets.

1.3 AI as a clarifying context

This section explains how generative AI fits into the thesis. AI is not treated as a causal variable, but as a contextual layer that sharpened investor attention on compute infrastructure.

Forecasts of potential productivity impact have been framed at a scale large enough to influence expectations about long-run demand for compute capacity [7]. Indicators of business adoption and private investment have been rising rapidly, reinforcing the salience of AI-driven demand narratives in public discourse [18].

For real assets, the operational distinction between training and inference is useful. Training is commonly associated with locations prioritizing low-cost electricity and abundant land, while inference is more naturally associated with low latency and proximity to users [10]. Related empirical work suggests that third-party data centers historically exhibit a strong urban bias and local-demand responsiveness, while major cloud providers are more sensitive to cost factors and less likely to locate in urban areas [10].

AI-linked workloads also intensify engineering constraints that translate into real estate constraints. Higher rack densities, increased power requirements, and more demanding cooling systems are frequently cited as salient differences relative to conventional facilities [9]. Historical averages of roughly 4–5 kW per rack have risen to 8–10 kW in recent years, with some facilities above 20 kW. AI data centers are described as operating above 30 kW per rack, with emerging designs requiring 100 kW per rack [9].

These observations matter because they provide an economically grounded reason why the market may update what it believes it is buying when it buys "data center" exposure. The thesis remains disciplined about interpretation: it studies identity through pricing relationships and uses sector context to interpret those relationships.

1.4 Scope and the public-market lens

This section explains why the thesis focuses on public markets.

The public market offers one large practical advantage: **pricing is continuous and observable.** Data center assets are increasingly discussed as critical infrastructure supporting modern daily life and business operations, including hyperscale demand and cloud-driven consumption [20].

At the same time, the data center investment ecosystem spans both public and private capital. Private markets finance a large share of capacity expansion and platform formation, while listed markets package a narrower slice into investable securities that trade inside broader equity and REIT ecosystems.

At the sector level, most cloud and data infrastructure investment comes from private capital. A World Bank assessment reports that roughly 96 percent of financing comes from private sector investors, underscoring how little of overall capacity expansion is visible in public markets [27].

Within the listed real estate universe, by contrast, data center exposure is packaged into a narrow REIT sector concentrated in a small number of names. As of June 2024, data center REITs represented 9.5 percent of the FTSE Nareit All Equity Index by market capitalization[1], while the FTSE Nareit Equity Data Centers sector consisted of only two constituents [21]. This concentration makes the listed "data center REIT sector" a useful signal for public market pricing, but also an inherently partial lens on the broader data center ecosystem.

This thesis uses public REITs as a lens not because public markets contain the whole industry, but because they provide a consistent window into how investors classify and price

[1] June 2024 figures represent the most recent sector-level breakdown available from Nareit at the time of data collection.

the category over time. The empirical focus is defined around pricing relationships that can be measured at high frequency across multiple regimes within 2015–2025.

1.5 Research questions and hypotheses

This section presents the research questions and hypotheses that guide the empirical analysis. All are centered on the pricing identity of a U.S. listed data center REIT basket relative to a REIT benchmark (excluding data centers) and technology equity ETFs.

RQ1 (Regime-level identity shift): How does the public-market pricing identity of the U.S. listed data center REIT basket evolve across macro regimes (2015–2019; 2020–2021; 2022–2023; 2024–2025)[2], measured by co-movement and factor loadings relative to (i) an equal-weight REIT benchmark excluding data centers and (ii) technology ETFs?

RQ2 (Time-varying identity confirmation): Do time-varying (rolling) factor loadings confirm that the data center REIT basket's identity is not constant through time, and does the evolution appear consistent across technology ETFs (XLK and SMH)[3]

H1 (Identity is regime-dependent): The pricing identity of the data center REIT basket is not stable across regimes. Both co-movement measures and factor exposures vary materially by regime.

H2 (Tech-alignment is observable and robust): Conditional on a REIT benchmark factor, the estimated technology exposure of the data center REIT basket is economically meaningful and exhibits consistent patterns across alternative technology proxies. Time-varying (rolling) estimates corroborate the regime-level evidence.

A note on terminology: Throughout the thesis, "tech alignment" means pricing sensitivity to a technology equity proxy **conditional on** a REIT benchmark factor. "REIT alignment" means pricing sensitivity to the REIT benchmark factor **conditional on** the technology proxy. These are factor sensitivities estimated over moving time windows (rolling factor loadings). The two technology ETFs used in this study (XLK and SMH) are referred to interchangeably as "technology proxies" or "tech proxies" throughout the thesis, reflecting their role as empirical proxies for broader technology sector exposure rather than as the object of study itself.

[2] The term "regime" refers to discrete time periods defined by distinct macroeconomic conditions. Detailed regime definitions are provided in Chapter 3, Section 3.5.

[3] Neither XLK nor SMH includes data center REITs, nor have they historically. Equinix converted to REIT status effective January 1, 2015, prior to its S&P 500 inclusion in March 2015, and has been classified under Real Estate since then. Digital Realty has operated as a REIT since its founding in 2004. XLK tracks the S&P 500 Technology Sector, which classifies both companies under Real Estate. SMH tracks the 25 largest U.S.-listed semiconductor companies, excluding REITs entirely.

1.6 Thesis structure and the role of OpCo/PropCo

This section outlines how the thesis is organized and explains where OpCo/PropCo fits.

OpCo/PropCo as an interpretive lens. A forward-looking implication of "identity change" is organizational. As operating complexity, customer services, and platform economics grow in importance, separation between real estate ownership and operating platforms becomes a plausible industry direction. Strategic alternatives in sector discourse have included PropCo and/or OpCo structures [15].

In this thesis, OpCo/PropCo is not treated as a hypothesis to be econometrically proven. It is used as an interpretive and forward-looking lens. Chapter 2 frames why operating platforms and service layers are central to how the asset is evolving. Chapter 6 revisits OpCo/PropCo as a structured way to think about where the sector could go next.

Thesis structure (6 chapters):

- **Chapter 1 (Introduction):** Defines the valuation puzzle as "data center identity change" and states RQs/Hs.
- **Chapter 2 (Background):** Establishes the conceptual building blocks, including business models, demand heterogeneity, infrastructure constraints, and why hybrid characteristics can shift public-market classification.
- **Chapter 3 (Data & Method):** Documents data construction and index definitions, rolling co-movement measures, regime partitions, and the identity regression framework used to estimate REIT and technology exposures.
- **Chapter 4 (Results):** Presents regime summaries and time-varying evidence, and interprets results strictly as pricing relationships rather than causal mechanisms.
- **Chapter 5 (Discussion):** Interprets what identity change implies for valuation narratives, capital allocation, and the public-versus-private boundary.
- **Chapter 6 (Conclusion):** Develops a forward view of the sector, including the OpCo/PropCo pathway and the role of power and grid integration constraints.

Chapter 2

Background and Conceptual Framework

2.1 The cloud is physical: data centers as locational and leaseable assets

This section establishes a foundational point: the cloud is not placeless. Core internet functions rely on specific buildings that host servers and enable interconnection, and those buildings are governed through real estate ownership and contracts [12].

This physical grounding matters because it places data centers inside a familiar real estate logic. They are capital-intensive sites with specialized engineering systems, security, and resilience features. At the same time, they also sit inside a network logic. Interconnection and exchange functions can turn a building into an operational node that shapes how data is stored and transmitted [12].

A concrete illustration: One Summer Street in Boston has lower floors serving as traditional retail while upper floors function as a critical internet exchange and server environment supported by security systems and backup power [12]. This example shows the asset's double character. It circulates financially like commercial real estate while remaining anchored physically by the equipment and network connections it hosts.

Historical context. Over the past three decades, the sector has moved from dial-up era computing toward cloud computing and hyperscale demand, with the physical footprint evolving alongside technology's expanding role in daily life [20]. Each wave reshaped what the "building" was economically. In some periods, emphasis tilted toward location and connectivity density; in others, toward scale delivery, cost efficiency, and the ability to secure large power loads [10].

Within the listed universe, data centers are a relatively recent addition to the formal REIT sector map. FTSE Russell's "new economy" work describes data centers as part of a newer group of REIT sectors providing exposure to digital infrastructure, highlighting that data center and infrastructure REITs together account for a meaningful share of listed REIT market capitalization [11]. Institutional real estate analysis in the early 2020s began treating data centers explicitly as part of a broader rotation into "alternate" property sectors, alongside categories such as healthcare, self-storage, and student housing [19].

Data centers as a leaseable real asset. Data center cash flows are formed through contracts that blend space provision with operational commitments. Contracts can be structured around rack space and power-driven units, and electricity can represent a major operating cost that is sometimes passed through to customers [1].

This changes the negotiation focus relative to conventional commercial leases. Uptime expectations and performance remedies become central because downtime can create losses for tenants that differ in magnitude and character from typical office disruptions. Service-level agreements and service credits are therefore integral to how risk is allocated between operator and customer [14].

Lease structure is also tied to the facility's operating envelope. Tenants' access paths to redundant power and cooling, resilience expectations, and the allocation of responsibility for fit-out versus core infrastructure shape contract terms and cash flow stability [20].

2.2 Business models and facility types

This section describes the different types of data center facilities and business models.

Even without AI as a narrative layer, the sector contains distinct facility and business model archetypes. A common split is between hyperscale owner-occupiers and third-party providers, and within third-party models between retail and wholesale approaches [20].

Another widely used segmentation distinguishes enterprise, colocation, hyperscale, edge, and carrier-neutral interconnection facility types, which map to different customer needs and infrastructure characteristics [1]. In that framing, colocation operators provide shared infrastructure for multiple customers, hyperscale facilities can be oriented toward large dedicated capacity needs, and carrier-neutral assets emphasize connectivity and interconnection capabilities.

Contracts and economics vary by type. Retail colocation contracts can be priced per rack or per kW, and customers often pay for incremental power. Electricity costs may be passed through, while contract lengths differ by segment. Some segments have shorter contracts and higher churn, while others have longer contracts and lower churn [1].

These typologies help keep the thesis conceptually precise. They also create a bridge to the empirical idea of "identity." A platform that monetizes interconnection density, service layers, and high-touch customer needs can plausibly be mapped by investors into a different reference class than a platform that monetizes scale delivery and power procurement [10].

AI workload logic makes the same split intuitive. Training is commonly associated with locations prioritizing low-cost electricity and abundant land, while inference is more naturally associated with low latency and proximity to users [10]. This does not imply a clean binary in practice. It provides a mechanism by which "connectivity-heavy" and "power-heavy" value drivers can coexist and become more salient.

2.3 Power, cooling, and infrastructure constraints

This section explains why power and cooling have become binding constraints that shape what can be built, where, and how quickly.

Infrastructure coupling becomes more visible when power density rises. Planning trade-offs emerge between usable IT area and the space required for power and cooling infrastructure. At typical operating levels around 1.5 kW per rack, a meaningful portion of floor space is already lost to supporting infrastructure; above roughly 7 kW per rack, more than half of the space can be consumed by power and cooling equipment [22].

Recent policy-oriented discussion frames density escalation in contemporary ranges and links it directly to AI-driven requirements. Historical averages are described around 4–5 kW per rack, rising to 8–10 kW in recent years, with some facilities above 20 kW. AI data centers are described as operating above 30 kW per rack, with emerging designs requiring 100 kW per rack [9].

Financial implications. Power procurement and cooling engineering become binding constraints that shape what can be built, where, and how quickly. Development requires significant investment and coordination with power and water suppliers as well as telecommunications carriers and ISPs. This represents a direct statement of multi-network coupling and a practical barrier to entry [1].

Large-scale development models can be interpreted as organizational responses to these constraints. Master-planned data center communities have emerged as a development model, emphasizing the role of land, power, and renewables in site strategy [15].

2.4 Energy measurement and its relevance to pricing

This section discusses why measurement uncertainty matters for how investors classify data centers.

Energy use has become a dominant external interface for the sector. The key point is not only that energy demand is rising, but that **measurement quality and projections vary widely**, creating space for narrative shifts to outrun robust quantification.

A critical review of data center energy estimates documents wide-ranging projections for 2030, with a spread that can be almost forty-fold across studies [16]. The same review shows how quickly publication volume has grown and how heterogeneous the author ecosystem is, which helps explain why investors and policymakers can face inconsistent benchmarks [16].

Sustainability framing also increasingly appears as a system-level issue tied to the scale of hyperscaler buildouts, local impacts, and the broader energy transition context [13].

Why this matters for pricing identity. When measurement is noisy and external constraints are salient, classification heuristics can exert outsized influence on how investors map an exposure into "real estate," "infrastructure," or "tech." Uncertainty is not neutral. It creates room for narrative shifts that affect how the same asset is priced.

2.5 Public market classification and investor attention

This section explains why the REIT lens makes pricing identity observable.

Why the REIT lens matters. A listed REIT lens is useful because it makes pricing identity observable. Public markets continuously reclassify risk, and the REIT wrapper forces data center exposure to trade inside a broader listed real estate ecosystem.

The listed REIT universe has expanded beyond traditional sectors and now includes "new economy" categories, including data centers, as a meaningful share of market capitalization [11]. Listed data center REITs provide exposure to cloud computing and e-commerce growth, and the segment has exhibited historical outperformance and a valuation premium relative to broad REIT benchmarks [20].

This sits alongside a broader institutional trend of expanding real estate allocations into alternate sectors, including data centers, as part of diversification strategies [19]. The relevance for this thesis is that investor sector frameworks are already evolving, which creates room for identity shifts to appear first in public pricing behavior.

Why hybrid assets reprice nonlinearly. A key reason pricing identity can change abruptly is that investors do not process all information symmetrically. Limited attention implies that investors focus on a subset of available signals and that attention allocation itself can influence how information is incorporated into prices [6].

This matters for hybrid assets because classification itself becomes a signal. When an asset is reframed as tech-adjacent infrastructure, it can inherit attention flows and risk perceptions that were previously directed elsewhere. Co-movement patterns can then shift even if property-level leases evolve gradually, because the reference class used by investors changes.

2.6 Conceptual framework and transition to empirical design

This section summarizes the conceptual frame and introduces the empirical approach.

The conceptual frame. Data centers can be understood simultaneously as leaseable real assets with property economics **and** operational platforms whose value is delivered through reliability, interconnection, and service levels [1, 14].

The balance between these characteristics can shift over time as density, power coupling, and service layers become more central [9, 22].

Within that frame, "identity change" is not a claim that the asset stops being real estate. It is a claim that **the marginal driver of perceived value and risk can migrate** away from conventional sector comparables and toward technology and infrastructure comparables, and that this migration can be observed in public market behavior.

OpCo/PropCo as a forward-looking structural implication. A forward-looking structural implication is the possibility that platform and service components become more explicitly separated from the property wrapper. Strategic alternatives discussed in sector advisory work explicitly include PropCo and/or OpCo structures [15].

Precedents exist in other operating real estate sectors. In lodging, the vertical disintegration of integrated hotel companies into separate operating and property entities allows a listed REIT or PropCo to hold the real estate and receive contractual rents, while a distinct OpCo or manager bears most of the operating volatility and labor-intensive business risk [2].

Senior housing provides a parallel example: under RIDEA-style structures, a healthcare or senior housing REIT owns the facilities and leases them to a taxable REIT subsidiary, which in turn engages a third-party manager under a fee-based management agreement, so that property-level cash flows are divided between relatively stable lease income at the PropCo level and more variable operating income at the OpCo/manager level [8, 23].

Transition to Chapter 3. The next chapter translates this conceptual frame into testable empirical expectations using public pricing relationships. The empirical design operationalizes "pricing identity" through 52-week rolling correlations, regime-based two-factor identity regressions that separate REIT-benchmark exposure from technology-proxy exposure, and rolling betas as a time-varying confirmation of identity evolution.

Chapter 3

Data and Methodology

3.1 Empirical objective and interpretation boundary

This chapter describes how the thesis measures "data center identity" empirically. Before presenting the data and methods, it is important to clarify what the analysis can and cannot establish.

The thesis treats data center identity as a **pricing relationship observed in public markets**, not as evidence of any underlying structural change. The analysis does not claim to show that data centers adopted new technology, changed their tenant mix, or transformed their operations. Instead, it asks a narrower question: when investors price listed data center REITs, do they treat them more like traditional real estate or more like technology stocks? And does this treatment change over time?

To answer this question, the thesis examines how a basket of listed data center REITs co-moves with two reference points: (i) a broad REIT benchmark and (ii) technology equity proxies. If data center returns move closely with other REITs, the market is pricing them as real estate. If they move more closely with tech stocks, the market is pricing them differently.

Research questions and hypotheses. Two research questions guide the empirical design:

- **RQ1 (Regime-level identity shift):** Does the pricing identity of the listed data center REIT basket vary across macro regimes (2015–2019, 2020–2021, 2022–2023, 2024–2025)?
- **RQ2 (Time-varying identity confirmation):** Do rolling estimates confirm that identity changes gradually over time, rather than being constant?

Two hypotheses follow directly:

- **H1:** Pricing identity is not stable across regimes. Both co-movement patterns and factor exposures vary materially by period.
- **H2:** Conditional technology exposure is economically meaningful and shows consistent patterns across alternative technology proxies.

Key definitions. Throughout Chapters 3 and 4, two terms have precise meanings:

- **Tech alignment** refers to how sensitive data center returns are to technology stocks, after controlling for the REIT benchmark.
- **REIT alignment** refers to how sensitive data center returns are to the REIT benchmark, after controlling for technology stocks.

These are pricing sensitivities, not statements about what the companies actually do. The analysis stays strictly within this boundary.

3.2 Data sources and sample scope

3.2.1 Listed sample frame

The analysis focuses on listed REITs classified under the "Data Center" sector in the Green Street REIT universe. Green Street provides a standardized sector taxonomy widely used for institutional real estate benchmarking. Within this framework, the Data Center sector is represented by two names: Equinix (EQIX) and Digital Realty (DLR).

This does not mean only two companies operate data centers. Many firms have data center exposure. However, the thesis specifically examines the listed REIT category, the investable public market packaging of data center real estate, and the Green Street classification provides a consistent way to define that category.

3.2.2 Why the data center basket has two constituents

The data center basket contains EQIX and DLR because these are the two names in the Green Street Data Center REIT category. The thesis uses an equal-weight combination of these two stocks as the "DC basket."

The small number of constituents reflects recent industry consolidation. Between 2021 and early 2022, three significant data center REITs exited the public market through take-private transactions: QTS Realty Trust was acquired by Blackstone, CyrusOne was acquired by KKR and Global Infrastructure Partners, and CoreSite was acquired by American Tower. Details on these transactions are provided in Appendix A.

This consolidation is itself relevant to the thesis narrative: it illustrates the public-versus-private boundary in data center investment and explains why the listed pure-play data center REIT universe is narrow.

3.2.3 Price series source

The analysis uses weekly adjusted closing prices from a standard financial data vendor. "Adjusted close" prices account for corporate actions such as stock splits and dividends, so that price changes reflect total returns rather than just nominal price movements. The thesis computes returns directly from these adjusted prices without additional reconstruction.

3.2.4 Frequency and sample horizon

The sample covers approximately ten years of weekly data:

- **Start date:** January 1, 2015
- **End date:** December 11, 2025
- **Total observations:** 572 weekly data points

The data is organized on a Thursday grid. This is simply a calendar choice: the grid starts on January 1, 2015 (a Thursday) and advances by seven days. The Thursday anchor has no special market microstructure meaning. It follows mechanically from the start date.

One technical note: January 1, 2015 is the calendar anchor, not necessarily the first trading day of 2015. Because computing a return requires two price observations, return series begin only after the second weekly observation.

3.3 Weekly alignment and return construction

3.3.1 Weekly alignment rule

To create the weekly dataset, each asset's price history is mapped to the Thursday grid using a conservative rule: each Thursday is assigned the most recent available adjusted closing price observed **on or before** that date.

This rule prevents "look-ahead" contamination. If we accidentally used prices from after the Thursday date, correlations and regressions would reflect information that was not yet available. This would be a serious measurement error. The backward-looking rule ensures the weekly panel uses only information available as of each grid date.

3.3.2 Log returns

Weekly log returns are computed as:

$$r_{i,t} = \ln\left(\frac{P_{i,t}}{P_{i,t-1}}\right)$$

Log returns are used because they have a useful property: they add up over time. If a stock has log returns of 2% in week 1 and 3% in week 2, the two-week log return is simply 5%. This makes it easy to aggregate returns across periods and to average returns across multiple stocks.

Log returns are the primary input for all correlations and regressions in the thesis.

3.3.3 Simple returns

Simple returns are computed as:

$$R_{i,t} = \frac{P_{i,t}}{P_{i,t-1}} - 1$$

Simple returns are used only for creating intuitive wealth index charts (starting from a base of 100). They are **not** used for statistical analysis, because compounding log returns as if they were simple returns would introduce errors.

3.3.4 Missingness and continuity

Some tickers enter the sample after the start date (IPOs). These tickers have missing values before their listing date and valid observations afterward. This is standard "left-censoring," meaning that early observations are missing, and does not pose problems for the analysis.

Two continuity points matter:

- **Individual tickers:** A ticker's return series begins only after it has two weekly price observations.
- **Constructed benchmarks:** The REIT benchmark (defined below) remains defined each week as long as at least one constituent has valid data.

In this dataset, no tickers disappear mid-sample. Membership changes occur through additions (post-IPO entries), not deletions. This helps the rolling-window analysis, which would be complicated if constituents vanished partway through.

3.4 Basket and benchmark construction

The analysis reduces the large universe of tickers into a small number of interpretable return series that directly address the research questions.

3.4.1 Data center basket (DC)

The data center basket is defined as the equal-weight average of EQIX and DLR log returns:

$$r_t^{DC} = 0.5 \cdot r_t^{EQIX} + 0.5 \cdot r_t^{DLR}$$

This matches the Green Street Data Center category. Equal weighting is transparent and avoids letting one firm's market cap dominate the category proxy.

3.4.2 REIT benchmark excluding data centers (REITxDC)

The REIT benchmark is the equal-weight average of all non-data-center REITs in the Green Street universe each week:

$$r_t^{REITxDC} = \frac{1}{N_t} \sum_{i \in U_t} r_{i,t}$$

where U_t is the set of non-DC REITs with valid returns at week t, and N_t is the count.

An important disclosure: this benchmark is **unbalanced**, meaning its membership changes over time as new REITs enter the sample after IPO. This is a common feature of equal-weight benchmarks and is explicitly acknowledged throughout the analysis.

3.4.3 Technology equity ETFs (XLK and SMH)

Technology exposure is measured using two alternative proxies:

- **XLK:** The Technology Select Sector SPDR ETF, which tracks the S&P 500 technology sector including software, hardware, semiconductors, and IT services [24].
- **SMH:** The VanEck Semiconductor ETF, which tracks the 25 largest U.S.-listed semiconductor companies [17, 25].

Using two proxies serves as a robustness check. If the results depend heavily on which proxy is used, the findings may be specific to that proxy rather than reflecting a genuine technology alignment. If results are similar across both proxies, the finding is more robust.

3.5 Empirical methods

This section describes the three statistical tools used to measure pricing identity. Each tool serves a distinct purpose.

3.5.1 Rolling correlations

The first tool is simple: 52-week rolling correlations between the DC basket and each reference series.

$$\rho_t^{(52)}(X, Y) = \mathrm{Corr}(\{X_{t-51}, \ldots, X_t\}, \{Y_{t-51}, \ldots, Y_t\})$$

The analysis computes correlations between DC and REITxDC, between DC and XLK, and between DC and SMH.

What rolling correlations show: They provide a transparent, visual picture of how co-movement evolves over time. If DC correlates more highly with REITs than with tech, the market is treating data centers as real estate. If the ranking reverses, something has changed.

What rolling correlations cannot show: They are pairwise measures. If REITs and tech stocks are themselves correlated, we cannot tell from correlations alone whether DC is responding to REITs, to tech, or to both. That decomposition requires the two-factor regression.

3.5.2 Regime partitions

The sample is divided into four regimes based on macroeconomic conditions:

Regime	Dates	Economic context
2015–2019	Jan 2015 – Dec 2019	Pre-COVID baseline
2020–2021	Jan 2020 – Dec 2021	COVID shock and aftermath
2022–2023	Jan 2022 – Dec 2023	Rate-hike stress period
2024–2025	Jan 2024 – Dec 2025	Post-hike period

These boundaries align with major policy shifts. The Federal Reserve began raising rates in March 2022, reached a peak of 5.50% in July 2023, and subsequently held or cut rates [3, 4].

The regime approach allows us to ask: did data center pricing identity differ across these distinct market environments?

3.5.3 Two-factor identity regression

The core statistical test is a two-factor regression estimated within each regime:

$$r_t^{DC} = \alpha + \beta_{REIT} \cdot r_t^{REITxDC} + \beta_{Tech} \cdot r_t^{Tech} + \varepsilon_t$$

where r_t^{Tech} is either XLK or SMH.

How to interpret the coefficients:

- β_{REIT} measures how much data center returns respond to the REIT benchmark, **holding the tech proxy constant**. This is "REIT alignment."
- β_{Tech} measures how much data center returns respond to the tech proxy, **holding the REIT benchmark constant**. This is "tech alignment."

A practical example: If $\beta_{REIT} = 0.80$ and $\beta_{Tech} = 0.20$, then when REITs go up 1% (tech unchanged), data centers tend to go up 0.80%. When tech goes up 1% (REITs unchanged), data centers tend to go up 0.20%. The market is pricing data centers primarily as real estate with modest tech sensitivity.

Statistical inference. Standard errors are computed using the Newey–West heteroskedasticity and autocorrelation consistent (HAC) estimator. This adjustment is standard for time-series regressions where residuals may be serially correlated or have changing variance. The implementation follows Wooldridge's treatment of HAC inference [26].

The Newey–West lag parameter is set to 8, meaning the estimator accounts for potential serial correlation in residuals up to 8 weeks (roughly two months). This is a conservative choice for weekly data. The lag parameter affects standard errors and statistical significance, not the coefficient estimates themselves.

3.5.4 Rolling two-factor betas

The regime regressions treat each period as homogeneous. To see how identity evolves more continuously, the same two-factor model is estimated on rolling 52-week windows:

$$r_\tau^{DC} = \alpha_t + \beta_{REIT,t} \cdot r_\tau^{REITxDC} + \beta_{Tech,t} \cdot r_\tau^{Tech} + \varepsilon_\tau, \quad \tau \in \{t-51, \ldots, t\}$$

This produces a time series of β_{REIT} and β_{Tech} estimates, showing how the pricing relationship drifts through time rather than jumping discretely at regime boundaries.

3.6 Construction choices and equal-weight rationale

3.6.1 Look-ahead prevention

The as-of mapping rule ensures the weekly panel never uses prices from after the Thursday grid date. Any mapping that sometimes pulled in future prices would contaminate the analysis with information not available at the time.

3.6.2 Log-return and simple-return separation

Log returns are used for all statistical analysis. Simple returns are used only for wealth index charts. Mixing these would introduce compounding errors.

3.6.3 Coefficient ordering in regression output

When using array-based regression functions, coefficients may be returned in a different order than the input columns. The analysis explicitly documents the extraction rules to ensure β_{REIT} and β_{Tech} labels are never accidentally swapped.

3.6.4 Equal-weight rationale

Both the DC basket and the REITxDC benchmark use equal weights rather than market-cap weights. For the DC basket (only two stocks), equal weighting is transparent and avoids letting one firm's size dominate the category proxy. The object is a category-level pricing identity, not a size-weighted portfolio return.

A market-cap-weighted design would answer a different question: the pricing identity of the larger firm's contribution to the category. The thesis instead fixes the object as the equal-weight proxy and tests whether its identity changes over time.

3.7 Limitations

Several limitations follow directly from the design:

- **Two-constituent DC basket.** Results describe how markets price the listed REIT category (EQIX and DLR), not the entire data center industry.
- **Unbalanced benchmark.** The REITxDC benchmark's membership expands over time as post-IPO REITs enter the sample. This is disclosed explicitly.

- **Constructed indices are not tradable.** The DC basket and REITxDC benchmark are averages of constituent log returns. These are statistical constructs for measuring pricing relationships, not implementable investment products.
- **Proxy interpretation.** XLK and SMH are used as alternative technology proxies for robustness. Neither uniquely represents "technology." Their role is to test whether findings depend on proxy choice.
- **Survivorship.** The sample contains no mid-period ticker deletions, suggesting delisted names may be excluded upstream. The thesis confines claims to observed pricing relationships within the available listed set.

3.8 Summary and transition to Chapter 4

This chapter defined the data and methods for measuring data center pricing identity:

- **Data:** A weekly, Thursday-aligned dataset from January 2015 to December 2025 (572 observations), covering a two-constituent data center basket (EQIX and DLR), an equal-weight REIT benchmark excluding data centers, and two technology proxies (XLK and SMH).
- **Methods:** Three linked tools: rolling correlations for visual co-movement patterns, regime-specific two-factor regressions for conditional decomposition with HAC inference, and rolling two-factor betas for continuous time-variation.

Chapter 4 presents the results in the same order: visual evidence first, then regime summaries, then conditional decomposition, and finally rolling confirmation. Throughout, interpretation remains strictly within the pricing-identity frame.

Chapter 4

Results and Interpretation

4.1 Interpretation framework

This chapter presents the empirical findings. Before examining the results, two points deserve emphasis.

First, all findings describe pricing relationships only. Nothing in this chapter is interpreted as evidence of technology adoption, tenant changes, or operational transformation. The analysis shows how public markets price listed data center REITs relative to other REITs and tech stocks. It shows nothing more.

Second, the key terms have precise meanings:

- **Tech alignment** = sensitivity to tech stocks, after controlling for REITs
- **REIT alignment** = sensitivity to REITs, after controlling for tech stocks

These are the β_{Tech} and β_{REIT} coefficients from the two-factor regression. When the chapter says "alignment shifted," it means these conditional sensitivities changed.

How to read the coefficients. A quick guide:

- $\beta_{REIT} = 0.80$ means: when REITs rise 1%, data centers tend to rise 0.80%, holding tech constant
- $\beta_{Tech} = 0.50$ means: when tech rises 1%, data centers tend to rise 0.50%, holding REITs constant
- If both are large and positive, data centers have a "hybrid" pricing identity
- If only β_{REIT} is large, data centers price like traditional real estate

- If only β_{Tech} is large, data centers price like tech stocks

The chapter follows the order previewed in Chapter 3: (i) wealth index context, (ii) rolling correlations, (iii) regime regressions, and (iv) rolling betas.

4.2 Wealth index context

This section provides visual context for the market environment. It is **not** used to draw conclusions about factor loadings. That comes later from the regressions.

4.2.1 Construction and level paths

Figure 4.1 shows base-100 wealth indices for the DC basket, REITxDC benchmark, XLK, and SMH, compounded from weekly simple returns starting January 2015.

The level paths reveal large differences in cumulative performance:

- **End-of-sample values (base 100):** DC $\approx$ 414, REITxDC $\approx$ 232, XLK $\approx$ 784, SMH $\approx$ 1433
- **Implied CAGRs:** DC $\approx$ 13.9%, REITxDC $\approx$ 8.0%, XLK $\approx$ 20.7%, SMH $\approx$ 27.5%

These differences explain why the thesis focuses on co-movement and conditional exposures rather than relative performance. Even when two assets have very different cumulative returns, their week-to-week co-movement can reveal how markets classify them.

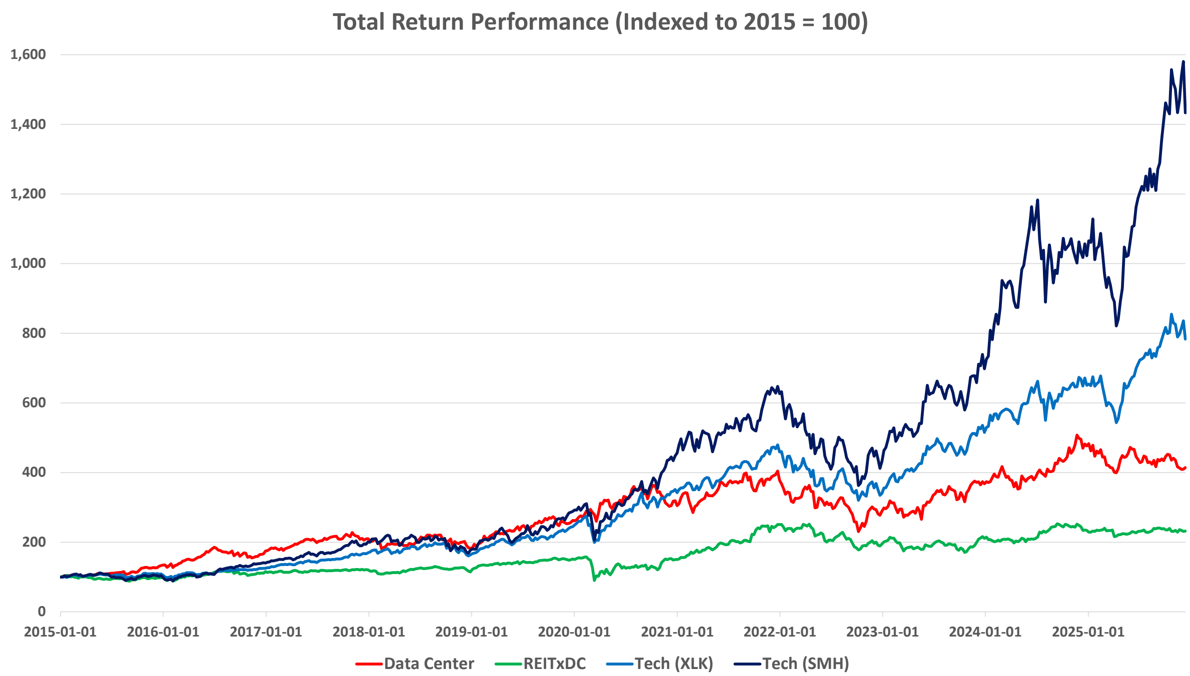

Figure 4.1: Base-100 wealth indices (simple-return compounding): DC basket vs. REITxDC vs. XLK vs. SMH (2015–2025 weekly). This figure is context only and is not used for factor interpretation.

4.2.2 Regime-level context

Table 4.1 shows simple-return total changes by regime, providing context for each period.

Table 4.1: Context only: simple-return total changes from the base-100 indices by regime (first-to-last grid date within each regime).

Regime	DC	REITxDC	XLK	SMH
2015–2019	+163%	+52%	+145%	+189%
2020–2021	+44%	+67%	+84%	+114%
2022–2023	+0%	-16%	+12%	+10%
2024–2025	+11%	+8%	+48%	+97%

A notable observation: 2022–2023 was a difficult period for the DC basket (essentially flat) and negative for the broader REIT benchmark, while tech remained slightly positive. This context matters for interpreting the correlations and betas that follow.

4.2.3 Drawdown context

For completeness, maximum drawdowns during the sample:

- DC: approximately -43% (peak late 2021 to trough October 2022)
- REITxDC: approximately -43% (peak February 2020 to trough March 2020), reflecting the severe COVID-19 market shock when broad REIT exposure experienced acute selling pressure amid concerns over rent collections and property valuations
- XLK: approximately -33% (peak late 2021 to trough October 2022)
- SMH: approximately -44% (peak late 2021 to trough October 2022)

These drawdowns describe the environment in which correlations and betas are estimated.

4.3 Rolling correlations

Rolling correlations provide the first evidence of how data center pricing identity evolved over time.

4.3.1 Rolling correlation paths

Figures 4.2 and 4.3 plot 52-week rolling correlations between the DC basket and each reference series. Two patterns stand out:

Pattern 1: Pre-2020, data centers co-moved primarily with REITs. The correlation between DC and REITxDC was consistently higher than the correlation between DC and either tech proxy.

Pattern 2: Post-2020, the ordering became mixed. In some periods, DC correlated more highly with XLK than with REITxDC. The simple "data centers are REITs" description no longer held consistently.

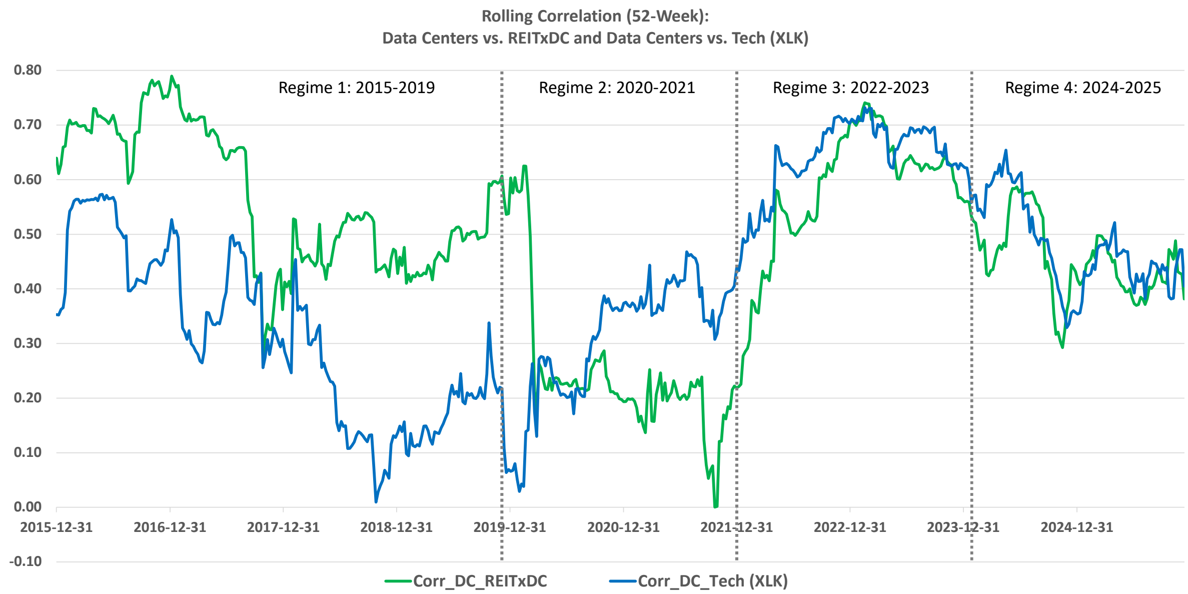

Figure 4.2: Rolling 52-week correlations: ρ(DC, REITxDC) and ρ(DC, XLK). The first valid observation appears after the 52-week window build-up (2015-12-31).

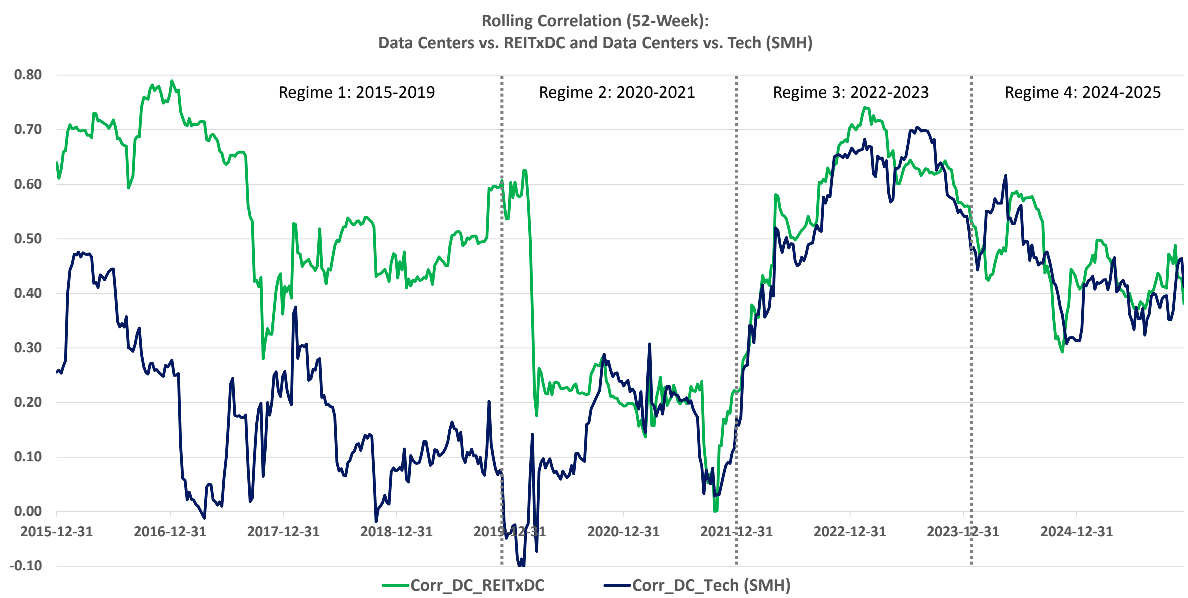

Figure 4.3: Rolling 52-week correlations: ρ(DC, REITxDC) and ρ(DC, SMH). The first valid observation appears after the 52-week window build-up (2015-12-31).

4.3.2 Regime summaries and interpretation

Table 4.2 summarizes the rolling correlations by regime.

Table 4.2: Rolling 52-week correlation regime summary (means on non-missing rolling windows). "Dominance share" is the fraction of weeks where ρ(DC, Tech) > ρ(DC, REITxDC).

Regime	N	Mean ρ(DC, REIT)	Mean ρ(DC, XLK)	Mean ρ(DC, SMH)	Share XLK > REIT
2015–2019	209	0.57	0.31	0.18	0%
2020–2021	105	0.24	0.31	0.13	74%
2022–2023	104	0.58	0.64	0.56	83%
2024–2025	102	0.46	0.48	0.44	66%

What this shows:

- **2015–2019:** Data centers clearly co-moved more with REITs. In 0% of weeks did tech correlation exceed REIT correlation.

- **2020–2021:** The ranking flipped. XLK correlation exceeded REIT correlation in 74% of weeks.
- **2022–2023:** All correlations were high. This happens in stressed markets where everything moves together.
- **2024–2025:** Mixed ordering persisted. Tech remained more frequently dominant, but the gap narrowed.

Important limitation: Correlations are pairwise. If REITs and tech are themselves correlated, we cannot tell from correlations alone whether data centers are responding to one, the other, or both. The two-factor regression addresses this by measuring **conditional** sensitivities.

4.4 Regime-level identity regressions

This section presents the core test: two-factor regressions that decompose data center returns into REIT alignment and tech alignment within each regime.

4.4.1 Results with XLK

Table 4.3 shows the regime regressions using XLK as the technology proxy.

Table 4.3: Regime two-factor identity regressions (Tech = XLK). HAC standard errors (lag=8) in parentheses. Stars reflect conventional two-sided thresholds based on HAC t-statistics (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

Regime	N	α	β_{REIT}	β_{Tech}	R^2
2015–2019	260	0.0025	0.80*** (0.08)	0.01 (0.06)	0.33
2020–2021	105	0.0008	-0.06 (0.08)	0.45*** (0.12)	0.15
2022–2023	104	0.0001	0.48*** (0.13)	0.52*** (0.13)	0.51
2024–2025	102	-0.0005	0.56*** (0.11)	0.31*** (0.08)	0.28

Interpretation:

- **2015–2019: Strongly REIT-aligned.** The REIT coefficient is 0.80 and highly significant; the tech coefficient is essentially zero. In practical terms, data center returns moved with REITs, not with tech. An investor could reasonably view data center REITs as "just another REIT sector."

- **2020–2021: The relationship flipped.** The REIT coefficient collapsed to near zero; the tech coefficient became 0.45 and significant. Data center returns now moved with broad tech, not with REITs. This is a striking regime-level identity shift.

- **2022–2023: A hybrid emerged.** Both coefficients are positive and significant (REIT ≈ 0.48, tech ≈ 0.52). Data centers were no longer purely REIT-like or purely tech-like. They exhibited sensitivity to both reference classes.

- **2024–2025: Mixed, with persistent tech alignment.** REIT alignment recovered (0.56) but tech alignment remained positive (0.31). Unlike the pre-2020 baseline, tech sensitivity did not return to zero.

Summary of the XLK specification: REIT-aligned pre-2020, tech-aligned in 2020–2021, hybrid in 2022–2023, and mixed with persistent tech sensitivity in 2024–2025.

4.4.2 Results with SMH

Table 4.4 shows the same regressions using SMH (semiconductors) as the technology proxy.

Table 4.4: Regime two-factor identity regressions (Tech = SMH). HAC standard errors (lag=8) in parentheses. Stars reflect conventional two-sided thresholds based on HAC t-statistics (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

Regime	N	α	β_{REIT}	β_{Tech}	R^2
2015–2019	260	0.0026	0.81*** (0.08)	-0.01 (0.04)	0.33
2020–2021	105	0.0024	0.05 (0.06)	0.11 (0.09)	0.04
2022–2023	104	0.0006	0.59*** (0.13)	0.32*** (0.12)	0.49
2024–2025	102	-0.0007	0.61*** (0.12)	0.21*** (0.06)	0.29

Interpretation:

- **2015–2019:** Same as XLK. Strongly REIT-aligned, zero tech alignment.

- **2020–2021:** Different from XLK. The SMH coefficient is only 0.11 and not statistically significant. The "tech tilt" observed with XLK does not appear when technology is defined narrowly as semiconductors. This tells us the 2020–2021 identity shift was primarily a **broad-tech** pricing relationship, not specifically semiconductor-driven.

- **2022–2023 and 2024–2025:** Both REIT and SMH coefficients are positive and significant. Unlike the early period, tech alignment is now robust across proxy definitions.

What the proxy difference reveals: The early tech-alignment signal (2020–2021) was specific to broad tech (XLK), not semiconductors (SMH). But post-2022, tech alignment appeared under both definitions, suggesting a more durable shift.

4.4.3 Interpretation

The regime regressions establish several findings:

1. **Pre-2020, data centers were priced as real estate.** Both proxies agree: strong REIT alignment, zero tech alignment.
2. **2020–2021 saw a tech shift, but it was proxy-sensitive.** The shift appeared strongly under XLK but not under SMH.
3. **Post-2022, a hybrid identity emerged.** Both REIT and tech coefficients are positive and significant under both proxy specifications. Data centers are no longer purely REIT-like.

The key contribution of the two-factor approach is separating conditional sensitivities. Correlations alone could not determine whether data centers were responding to REITs, tech, or both. The regression shows that post-2022, the answer is "both."

4.5 Rolling two-factor betas

The regime regressions treat each period as homogeneous. This section examines whether identity shifts appear as gradual trends or discrete jumps using rolling 52-week estimates.

4.5.1 Rolling betas with XLK

Figure 4.4 plots rolling estimates of β_{REIT} and β_{Tech} using XLK as the technology proxy.

Key patterns:

- **Pre-2020:** Rolling β_{REIT} is persistently high (around 0.78 on average); rolling β_{Tech} is low (around 0.12). Data centers priced like REITs.

- **2020–2021:** A pronounced shift occurs. Rolling β_{REIT} compresses toward zero (regime mean 0.10); rolling β_{Tech} rises (regime mean 0.30). Near end-2021, the tech loading reaches about 0.67 while the REIT loading is near zero.
- **Post-2022:** Both loadings become economically meaningful. The regime means for 2022–2023 are $\beta_{REIT} \approx 0.38$ and $\beta_{Tech} \approx 0.55$. For 2024–2025, they are ≈ 0.51 and ≈ 0.44. The hybrid identity is not a discrete regime artifact. It appears as a persistent pattern in the rolling estimates.

End-of-sample state: By December 2025, β_{REIT} is about 0.51 and β_{Tech} is about 0.33. Data centers retain meaningful exposure to both reference classes.

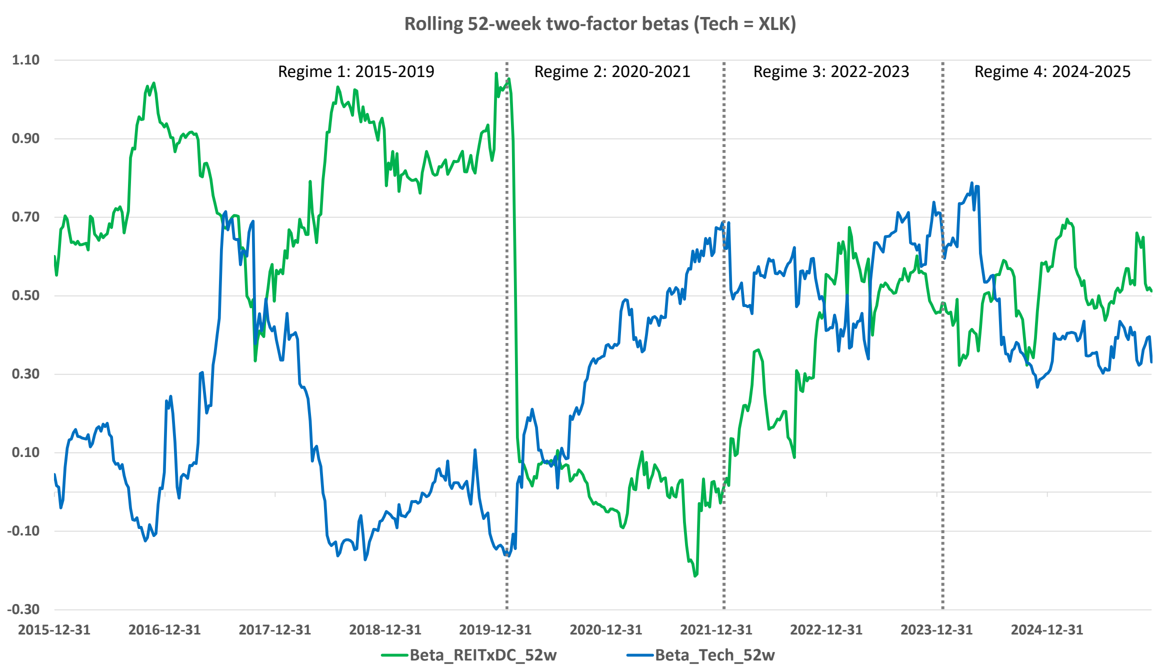

Figure 4.4: Rolling 52-week two-factor betas (Tech = XLK): β_{REIT} and β_{Tech} over time.

4.5.2 Rolling betas with SMH

Figure 4.5 repeats the exercise using SMH.

Key patterns:

- **Pre-2020:** Same as XLK. Strong REIT alignment (mean ≈ 0.81), near-zero SMH alignment (mean ≈ 0.03).
- **2020–2021:** Unlike XLK, the SMH loading remains near zero (mean ≈ 0.01). The early tech shift does not appear when technology is proxied by semiconductors.
- **Post-2022:** The SMH loading becomes positive and meaningful. Regime means are ≈ 0.30 for 2022–2023 and ≈ 0.28 for 2024–2025. The hybrid pattern is robust across proxies in the later period.

What this confirms: The 2020–2021 tech shift was broad-tech-specific. But post-2022, tech alignment appears under both XLK and SMH, suggesting a more fundamental change in how markets price data centers.

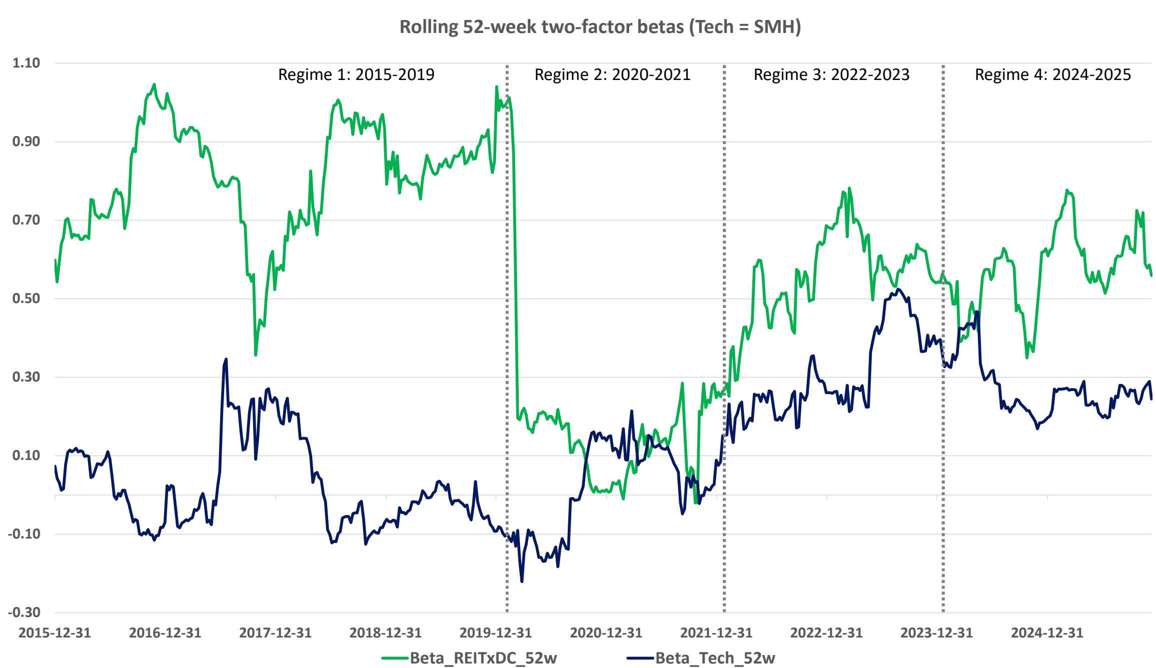

Figure 4.5: Rolling 52-week two-factor betas (Tech = SMH): β_{REIT} and β_{Tech} over time.

4.5.3 Model fit and regime means

Figures 4.6 and 4.7 show rolling R^2 for each specification. Model fit varies across regimes, being highest during 2022–2023 (the stressed macro period where all correlations were elevated).

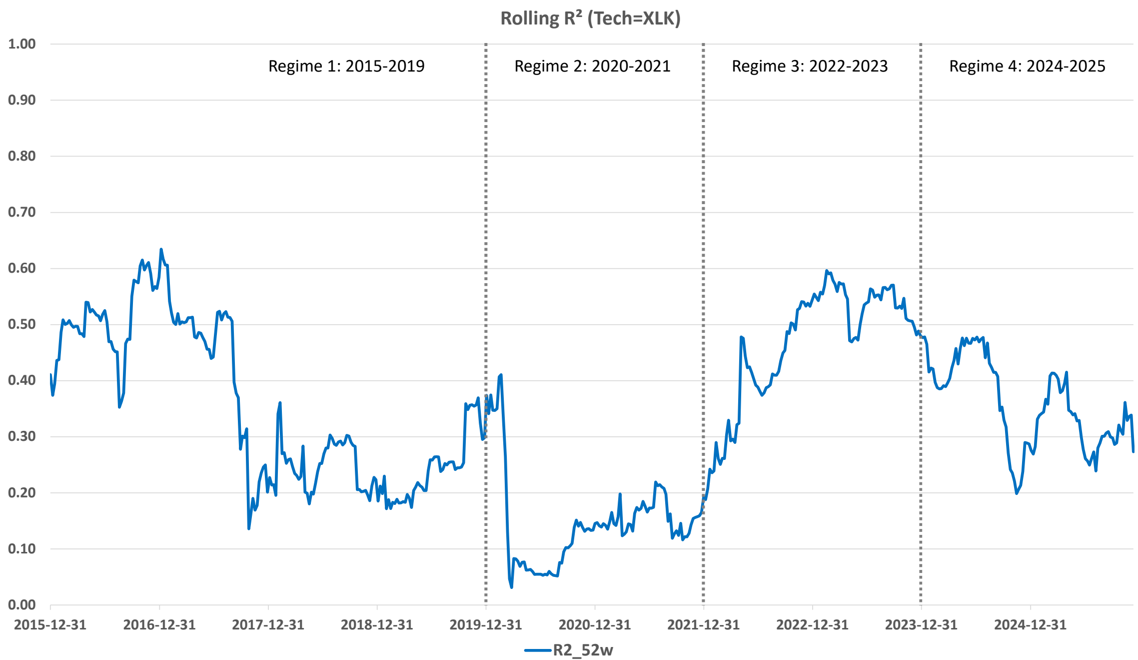

Figure 4.6: Rolling 52-week R^2 for the two-factor model (Tech = XLK).

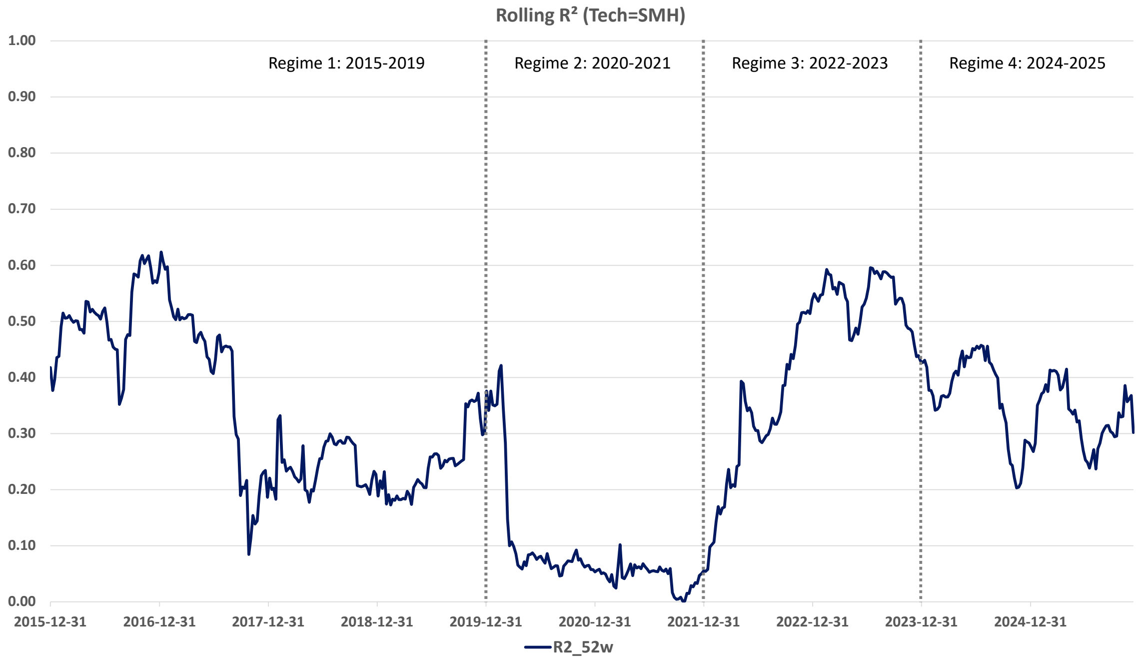

Figure 4.7: Rolling 52-week R^2 for the two-factor model (Tech = SMH).

Table 4.5 summarizes regime means of rolling betas.

Table 4.5: Regime means of rolling 52-week two-factor betas (OLS within rolling windows). Values are means over rolling windows that fall within each regime.

Tech proxy	Regime	N_{roll}	Mean β_{REIT}	Mean β_{Tech}	Mean R^2
XLK	2015–2019	209	0.78	0.12	0.36
XLK	2020–2021	105	0.10	0.30	0.15
XLK	2022–2023	104	0.38	0.55	0.47
XLK	2024–2025	102	0.51	0.44	0.36
SMH	2015–2019	209	0.81	0.03	0.35
SMH	2020–2021	105	0.21	0.01	0.09
SMH	2022–2023	104	0.55	0.30	0.43
SMH	2024–2025	102	0.57	0.28	0.35

4.5.4 Interpretation

The rolling evidence confirms two findings:

Finding 1: Identity is time-varying. Both β_{REIT} and β_{Tech} change materially over time. The pricing relationship is not constant.

Finding 2: Proxy choice matters early, less so later. The 2020–2021 tech shift appears under XLK but not SMH. However, post-2022, tech alignment appears under both proxies. The later-period hybrid identity is more robust.

4.6 Robustness and limitations

4.6.1 Alternative technology proxies

The use of two proxies (XLK and SMH) is a central robustness check. The findings are:

- **Baseline (pre-2020):** Both proxies agree. Strong REIT alignment, zero tech alignment.
- **Early shift (2020–2021):** XLK shows tech alignment; SMH does not. The signal is proxy-sensitive.
- **Later period (post-2022):** Both proxies show positive tech alignment. The hybrid identity is robust.

This pattern is itself informative. "Tech alignment" is not a single invariant concept. It depends on how technology is defined. The findings suggest the early shift was a broad-tech phenomenon, while the later hybrid structure is more general.

4.6.2 Constructed-series limitations

- **Two-constituent DC basket:** Results describe the listed REIT category (EQIX and DLR), not the entire data center industry.
- **Unbalanced REIT benchmark:** Membership expands over time as new REITs enter. It is a statistical reference, not a fixed index.
- **Pricing relationships only:** Findings do not identify structural drivers, tenant shifts, or causal mechanisms.

4.6.3 Why the chapter emphasizes conditional betas

The chapter places greater weight on β_{REIT} and β_{Tech} than on raw correlations because:

- Correlations are pairwise and cannot separate REIT vs. tech effects when the two anchors are themselves correlated.
- Conditional betas directly measure the thesis' definitions of REIT alignment and tech alignment.
- Rolling betas show whether identity drift is persistent or episodic.

4.7 Summary

This chapter answers the two research questions.

RQ1 (Regime-level identity shift): Answered.

Data center pricing identity varies materially across regimes:

- **2015–2019:** Strongly REIT-aligned, zero tech alignment.
- **2020–2021:** Tech-aligned under XLK, but not under SMH.
- **Post-2022:** Hybrid. Both REIT and tech coefficients are positive and significant under both proxies.

RQ2 (Time-varying identity confirmation): Answered.

Rolling betas confirm that identity changes gradually over time. The shift is not a discrete regime artifact but a persistent trend visible in the rolling estimates.

H1 and H2: Confirmed.

- **H1:** Pricing identity is regime-dependent. Co-movement and factor exposures vary materially across periods.
- **H2:** Conditional technology exposure is economically meaningful post-2022 and robust to alternative proxy definitions.

The bottom line: The listed data center REIT basket's pricing identity is not constant. It evolved from primarily REIT-aligned in the baseline period to a regime-dependent mixture of REIT and tech exposures in later periods. The strength and timing of tech alignment depend on how technology exposure is proxied, but the hybrid structure post-2022 is robust across definitions.

Chapter 5 interprets these findings in valuation terms and addresses alternative explanations.

Chapter 5

Discussion

5.1 Overview and Interpretive Objectives

This chapter interprets the empirical findings from Chapter 4 and connects them to broader implications for valuation, capital allocation, organizational form, and portfolio considerations.

Chapter 1 positioned a narrow empirical puzzle: the same listed vehicles are packaged as real estate, but appear to be increasingly interpreted through a technology lens in public markets. Chapter 2 explained why this shift is plausible for data centers specifically, where the delivered product combines a leaseable building with a reliability and interconnection platform. Chapter 3 fixed a strict interpretation boundary, targeting pricing relationships rather than causal operating transformation. Chapter 4 delivered the core evidence within that boundary.

This chapter extends the contribution without widening the claim set. Four objectives guide the discussion. First, the regime and time-varying patterns are translated into valuation terms that remain consistent with the thesis definition of identity as conditional alignment. Second, alternative explanations that a skeptical reader might raise are addressed explicitly. Third, the pricing identity results are connected to organizational form, including the OpCo/PropCo pathway as a structural implication. Fourth, implications for capital allocation and portfolio management are developed.

Throughout, the chapter maintains the interpretation boundary established in Chapter 3. The findings describe how markets price data center REITs. They do not prove that the underlying buildings became technology firms.

5.2 Interpreting Identity Shift in Valuation Terms

This section restates the Chapter 4 evidence in valuation language rather than statistical language.

In the baseline regime (2015–2019), the listed data center REIT basket was priced primarily through a REIT reference class. The conditional REIT beta was large and significant, while the conditional technology beta was essentially zero. An investor could reasonably treat data center REITs as "just another REIT sector" during this period.

In later regimes (post-2022), the same basket was increasingly priced through a hybrid reference class that incorporates both a REIT anchor and a technology anchor. Both conditional betas became positive and significant. This hybrid structure carries a concrete implication: the market is no longer using a single comparable set to price these assets. Instead, pricing reflects a blended narrative where discount rate expectations, growth optionality, and benchmark flows can be transmitted through both real estate and technology channels.

What this means for valuation. When an asset's reference class shifts, several valuation-relevant quantities can change. Comparable company selection becomes less obvious. Discount rate benchmarks may need blending across sectors. Sensitivity to sector-specific flows (such as REIT index rebalancing or technology fund flows) becomes harder to predict.

The contribution is therefore about classification and comparables, not about a definitive causal driver. A change in conditional loading is treated as evidence that the public market reference class used to price these vehicles is no longer stable. It is not treated as proof that the underlying buildings became technology firms.

5.3 Why Proxy Differences Matter

This section explains why the proxy sensitivity observed in Chapter 4 is informative rather than inconvenient.

A broad technology proxy (XLK) and a narrower semiconductor proxy (SMH) represent different investor narratives. XLK captures a general technology classification channel that includes software, hardware, and IT services. SMH captures a more specific compute supply chain channel focused on semiconductor companies.

The timing difference is diagnostic. Technology alignment appeared earlier under XLK (2020–2021) but not under SMH until later (post-2022). This pattern is consistent with a staged narrative environment. The early stage aligns with a broad "cloud and digital infrastructure" framing where data centers were associated with general technology growth. The later stage aligns with an "AI-era compute constraint" framing where the marginal bottleneck shifted toward delivered megawatts, GPU supply chains, and execution capacity.

Preventing over-literal interpretation. This proxy point prevents an overly literal interpretation of "tech beta." Technology is not a single homogeneous factor in how public markets organize beliefs. Proxy choice operationalizes competing reference classes. The thesis result is therefore not merely "tech increased." The result is that the comparative anchor through which the sector is viewed became more elastic, and this elasticity became more visible after 2022.

The fact that post-2022 technology alignment appears under both XLK and SMH suggests the hybrid identity is not proxy-specific. It reflects a more fundamental shift in how markets classify these assets.

5.4 Alternative Explanations

This section addresses the most predictable critiques of the findings. The goal is to clarify what the current design already absorbs, what remains open, and what extensions would strengthen the defense.

5.4.1 The Generic Technology Rally Critique

A skeptical reader might argue that the results simply reflect a generic technology rally rather than a genuine identity shift. This critique is fundamentally a correlation story. The thesis object is not a correlation story. It is a conditional alignment story.

The two-factor specification conditions technology exposure on a REIT benchmark excluding data centers. In that construction, broad risk-on co-movement that affects both REITs and technology can raise raw correlations, but it does not automatically generate a stable incremental technology loading after conditioning on the REIT anchor.

The proxy differences strengthen this defense. A single narrative explanation claiming "technology drove everything" becomes less coherent when the timing and magnitude of alignment differs across plausible technology anchors, and when the later period exhibits a more clearly hybrid exposure structure. The empirical pattern is more consistent with a reference class reweighting interpretation than with a simple rally explanation.

5.4.2 The Interest Rate and Discount Rate Critique

Interest rates and discount rates are a credible alternative explanation because listed real estate is duration-sensitive, and the post-2022 period contained large discount rate movements. The current two-factor identity regression does not include an explicit rate factor. This limitation remains real and is better treated as an acknowledged boundary condition than as a hidden vulnerability.

The design still provides partial protection through the REIT benchmark anchor. REITxDC aggregates many channels through which discount rate changes and credit conditions affect listed real estate. Conditioning on REITxDC therefore asks a narrower question: holding the broad listed REIT complex constant, does incremental technology anchor alignment persist? That is a pricing identity question rather than a pure rates question.

Empirical work on listed real estate has documented that REIT returns exhibit measurable sensitivity to changes in interest rates, with interest rate betas varying across firms as financing structures and debt maturities differ [5].

A minimal extension could add a third factor such as weekly changes in the 10-year Treasury yield, and re-estimate the identity regression by regime. The defense target would be narrow: verifying that post-2022 conditional technology alignment does not vanish when a direct rates channel is included. This extension is noted as a direction for future work but is beyond the current scope.

5.4.3 Composition and Concentration

A concentrated public universe invites a compositional critique. The thesis already discloses that the public data center REIT basket contains only two constituents (EQIX and DLR). The appropriate response is not to deny the limitation but to interpret it correctly.

The empirical result is an estimate of listed wrapper identity, not a census of the entire asset class. That is consistent with the motivation in Chapter 1 that the public wrapper is a partial slice but is precisely where identity is observable in market prices.

Two points mitigate the concern. First, both constituents exhibit the same directional pattern. The identity shift is not driven by a single outlier. Second, the timing of the shift coincides with broader market and sector developments (the take-private wave in 2021–2022, the AI narrative intensification in 2023–2024) rather than with firm-specific events. This suggests the pattern reflects sector-level repricing rather than idiosyncratic noise.

5.5 Two Value Drivers: Connectivity Versus Power

This section connects the empirical results to the conceptual framework introduced in Chapter 2.

The data center product bundle contains two economically distinct attributes that are jointly priced in the public wrapper.

The first attribute is connectivity and interconnection. This attribute is tied to metropolitan proximity, low latency, network density, and ecosystem stickiness. Once embedded in an interconnection-rich environment, switching costs rise, and the value proposition becomes legible to investors as a service platform rather than as a single-tenant building. This attribute aligns with the "retail colocation" and "carrier-neutral" facility types described in Chapter 2.

The second attribute is power and deliverability. This attribute is tied to megawatt scale, grid interconnection rights, permitting lead times, and execution risk in capacity delivery. The dominant investor vocabulary for this attribute emphasizes constraints: power queues, utility negotiations, and capital intensity. This attribute aligns with the "hyperscale" and "wholesale" facility types, and has become more salient as AI workloads intensified power requirements [9].

Connecting to the results. The hybrid identity finding is consistent with a setting where both attributes became more salient to valuation narratives over time. In earlier periods, a REIT reference class could be sufficient for pricing when the dominant investor vocabulary was property-like (location, occupancy, rent rolls). In later periods, the market vocabulary expanded to include both platform-like growth optionality (connectivity) and infrastructure-like constraint risk (power). Hybrid alignment is then a natural pricing outcome even if the underlying real estate remains real estate.

The shift from XLK-only alignment (2020–2021) to XLK-and-SMH alignment (post-2022) may reflect this evolution. The early period emphasized general digital infrastructure growth. The later period added compute-specific constraints, making semiconductor exposure more relevant.

5.6 Capital Allocation: Public Versus Private Boundary

This section connects the pricing identity findings to broader capital allocation patterns.

Chapter 1 emphasized that listed data center REITs represent a partial slice of a much larger ecosystem. Most data center investment is funded by private capital rather than listed vehicles. A World Bank assessment reports that roughly 96 percent of cloud and data infrastructure financing comes from private sector investors [27]. The listed REIT sector is therefore a narrow window into the broader market.

Why identity drift matters for the public-private boundary. The identity drift evidence indicates that the public wrapper is priced through blended comparables. The public wrapper therefore inherits both the costs and benefits of hybrid classification.

In supportive environments, technology-aligned narratives can lower the equity cost of capital and reward visible growth pipelines. Public markets may assign higher multiples to companies perceived as technology platforms rather than traditional landlords.

In risk-off environments, the same narratives can increase drawdown sensitivity and penalize execution uncertainty. The 2022–2023 period illustrated this dynamic: data center REITs experienced drawdowns that resembled technology stocks more than traditional REITs.

Implications for private capital. When the marginal sector growth problem shifts toward power deliverability and large-scale development, the relevant risk is no longer only occupancy and rent roll stability. It becomes execution dispersion under constraint. That form of risk is more naturally warehoused in vehicles that can tolerate long lead times, staged capital deployment, and joint venture structures. Private platforms, project-level capital stacks, and long-duration closed-end funds are organizational forms that can absorb this dispersion with less mark-to-market volatility pressure.

This logic is compatible with the thesis boundary because it is framed as a consistent implication, not as a proven causal mechanism.

5.7 Organizational Implications: OpCo/PropCo as a Structural Pathway

This section connects the identity drift findings to organizational form. The argument is framed as a consistent implication, not as a prediction.

The logic. If data center value increasingly derives from operating platform attributes (reliability, interconnection, managed services) rather than pure property attributes (location, square footage, occupancy), then organizational separation between the property wrapper and the operating platform becomes economically plausible. The hybrid identity finding provides motivation: if the market already prices data centers through both real estate and technology reference classes, separating the two components could allow each to trade at its appropriate multiple.

Precedents in other operating real estate sectors. This pattern has precedents in sectors where operating intensity is high and property income can be distinguished from platform income.

In lodging, integrated hotel companies have vertically disintegrated into separate operating and property entities [2]. A listed REIT or PropCo holds the real estate and receives contractual rents, while a distinct OpCo or manager bears operating volatility and labor-intensive business risk. This structure allows the property component to be valued for yield stability while the operating component is valued for brand and management capability.

Senior housing provides a parallel example. Under RIDEA-style structures, a healthcare or senior housing REIT owns facilities and leases them to a taxable REIT subsidiary, which engages a third-party manager under a fee-based management agreement [8, 23]. Property-level cash flows are divided between relatively stable lease income at the PropCo level and more variable operating income at the OpCo level.

Why this is plausible for data centers. Strategic alternatives discussed in sector advisory work explicitly include PropCo and/or OpCo structures for data centers [15]. The economic rationale follows from the dual-attribute framework in Section 5.5.

A PropCo would hold the physical assets: land, buildings, power infrastructure, and grid interconnection rights. It would receive stable lease payments tied to power capacity and space. Valuation would emphasize yield, lease duration, and tenant credit quality. The natural reference class would be traditional real estate or infrastructure.

An OpCo would operate the platform: interconnection services, managed services, customer onboarding, and network ecosystem management. It would capture operating upside tied to customer growth, cross-connect revenue, and value-added services. Valuation would emphasize growth, platform economics, and customer lifetime value. The natural reference class would be technology or digital infrastructure.

Connecting to identity drift. The hybrid identity finding suggests the market is already implicitly blending these two valuation frameworks. Explicit organizational separation could make the blending transparent and allow each component to attract its natural investor base. PropCo investors seeking yield would not be forced to accept technology-like volatility. OpCo investors seeking growth would not be constrained by REIT distribution requirements.

Interpretation boundary. This section does not claim that OpCo/PropCo separation will occur or should occur. It claims that the identity drift evidence is consistent with organizational structures that separate property income from platform income. Whether such separation creates value depends on execution, tax efficiency, and market conditions that are beyond the scope of this thesis.

5.8 Portfolio and Practical Implications

This section translates the identity drift findings into portfolio context.

Benchmarking implications. When conditional betas vary across regimes, a single-sector benchmark becomes less informative. A data center REIT that is benchmarked against a broad REIT index may appear to generate "alpha" when it is actually tracking technology movements, or vice versa. A blended benchmark that allows weights to change by regime aligns more closely with observed pricing behavior.

Hedging considerations. A position that is assumed to be hedged by a REIT short can retain residual technology exposure when conditional technology alignment rises. Conversely, a position assumed to be hedged by a technology short can retain residual real estate exposure. The rolling beta evidence in Chapter 4 shows that these conditional exposures are not constant through time, which complicates hedge ratio estimation.

Diversification context. The regime-level evidence suggests that data center REITs offered different diversification properties in different periods. In 2015–2019, when REIT alignment dominated, data center exposure provided diversification relative to technology. In 2020–2021, this diversification benefit diminished as technology alignment increased. Post-2022, the hybrid structure suggests partial diversification: data centers co-move with both anchors, but not perfectly with either.

Relative value context. The time-varying betas create interpretable spread relationships. The difference between data center returns and REITxDC returns, optionally adjusted by rolling technology betas, represents a measure of technology-specific pricing. This spread widened during periods of high technology alignment and compressed during periods of REIT alignment. The thesis notes this pattern but does not claim profitability without a dedicated backtest.

5.9 Limitations and Extensions

This section restates limitations that affect interpretation and identifies directions for future work.

Factor parsimony. The two-factor design is intentionally interpretable, but it leaves macro channels only indirectly captured through market anchors. The most valuable single extension would be a direct interest rate control, implemented as a third factor. This would help separate discount rate effects from reference class migration.

Regime partition dependence. Regimes are chosen for interpretability based on macroeconomic conditions rather than endogenously estimated breakpoints. The contribution is the comparison across a transparent partition, not the identification of a single "true" break date. Alternative partitions could be examined for robustness.

Public wrapper scope. The results characterize listed pricing identity. They do not characterize the full private universe. Connecting public identity measures to private market evidence would require a separate empirical module with access to private transaction data.

Two-constituent basket. The narrow basket (EQIX and DLR) limits generalizability. Expanding the basket to include data-center-adjacent names (such as tower REITs or infrastructure REITs) could test whether the identity drift is specific to pure-play data centers or reflects a broader "digital infrastructure" repricing.

OpCo/PropCo evidence. The organizational implications in Section 5.7 are framed as consistent with the findings, not as independently tested hypotheses. Directly testing whether OpCo/PropCo structures create value would require firm-level event studies or cross-sectional comparisons that are beyond the current scope.

5.10 Summary

This chapter added four layers of interpretation without changing the empirical claim set.

First, identity drift was interpreted in valuation language. The shift from REIT-only alignment to hybrid alignment represents a migration in the market's reference class for pricing data center REITs. This affects comparable selection, discount rate benchmarking, and sensitivity to sector-specific flows.

Second, alternative explanations were addressed explicitly. The conditional specification already absorbs generic co-movement through the REIT anchor. The proxy differences strengthen the case against a single-factor "technology rally" explanation. Interest rate sensitivity remains a limitation that could be addressed through a three-factor extension.

Third, the findings were connected to organizational form. The dual-attribute framework (connectivity versus power) provides economic motivation for OpCo/PropCo structures. Precedents in lodging and senior housing illustrate how operating real estate sectors have separated property income from platform income. The identity drift evidence is consistent with, though does not prove, similar organizational evolution in data centers.

Fourth, implications for capital allocation and portfolio context were developed. The public-private boundary, benchmarking challenges, hedging complications, and diversification dynamics all follow from the time-varying nature of pricing identity.

Throughout, the discussion maintained the interpretation boundary from Chapter 3. The results describe pricing relationships. They do not prove structural transformation of the underlying assets.

Chapter 6

Conclusion

This thesis examined whether public markets have shifted the pricing identity of listed data center REITs from a traditional real estate reference class toward a hybrid real estate-technology reference class. The answer is yes, with important qualifications about timing, proxy choice, and interpretation boundaries.

6.1 Summary of Findings

The empirical analysis employed a two-factor conditional beta framework to measure how data center REIT returns align with a REIT benchmark (excluding data centers) and a technology benchmark across distinct market regimes. Three core findings emerged.

First, pricing identity shifted from REIT-only to hybrid. In the baseline period (2015–2019), data center REITs exhibited strong conditional alignment with the REIT benchmark and negligible alignment with technology proxies. By the post-2022 period, both conditional betas became positive and significant, indicating that markets now price these vehicles through a blended reference class.

Second, proxy choice reveals a staged narrative. A broad technology proxy (XLK) captured alignment earlier (2020–2021), while a narrower semiconductor proxy (SMH) showed significant alignment only in later periods (post-2022). This pattern is consistent with an evolving investor narrative: early framing emphasized general digital infrastructure growth, while later framing incorporated AI-era compute constraints and power deliverability.

Third, the identity shift is not explained by generic co-movement. The conditional specification controls for broad REIT market movements. The incremental technology loading therefore reflects sector-specific repricing rather than a simple correlation artifact.

6.2 Contributions

This thesis makes three contributions to the literature on real estate pricing and sector classification.

The primary contribution is empirical: documenting a measurable shift in how public markets classify data center REITs. Prior work has noted the operational distinctiveness of data centers, but has not quantified the pricing identity implications using a conditional factor framework with regime variation.

The second contribution is methodological: demonstrating that proxy sensitivity is informative rather than inconvenient. The timing difference between XLK and SMH alignment provides diagnostic value for understanding how investor narratives evolve.

The third contribution is interpretive: connecting pricing identity evidence to organizational form. The OpCo/PropCo framework, drawn from precedents in lodging and senior housing, offers a structural lens through which hybrid identity can be understood and potentially resolved.

6.3 Limitations

Several limitations bound the interpretation of these findings.

The two-factor design does not include a direct interest rate control. While the REIT benchmark partially absorbs discount rate effects, a three-factor extension would strengthen the defense against rate-driven alternative explanations.

The public data center REIT universe contains only two constituents (Equinix and Digital Realty). The findings characterize listed wrapper identity, not the full private universe. Generalizability to the broader asset class requires caution.

Regime boundaries were chosen for interpretability rather than estimated endogenously. Alternative partitions could yield different magnitudes, though the directional pattern is robust across reasonable alternatives.

6.4 Directions for Future Research

Several extensions would strengthen and extend this work.

A three-factor model incorporating Treasury yield changes would isolate discount rate effects from reference class migration. This extension is the most direct way to address the interest rate critique.

Connecting public pricing identity to private market evidence would test whether the identity shift reflects broader capital allocation patterns or is confined to the listed wrapper. Access to private transaction data would enable this comparison.

Event studies around OpCo/PropCo announcements in related sectors could test whether organizational separation creates value when hybrid identity is present. The lodging and senior housing precedents suggest testable hypotheses.

Expanding the asset basket to include tower REITs, fiber REITs, and other digital infrastructure names would test whether the identity drift is specific to data centers or reflects a broader "digital infrastructure" repricing phenomenon.

6.5 Concluding Remarks

Data centers occupy an unusual position in the real estate landscape. They are buildings, but they house infrastructure that enables the digital economy. They generate rental income, but their value increasingly derives from power delivery capacity and network connectivity. They trade as REITs, but markets appear to price them partly as technology companies.

This thesis does not resolve whether data centers "are" real estate or technology. That question may not have a stable answer. What the thesis does establish is that the market's revealed preference for classifying these assets has shifted over time, and that this shift is measurable in public market prices.

For investors, the implication is that benchmarking, hedging, and portfolio construction require attention to time-varying classification. Additionally, to the extent that data center REITs retain their high-dividend REIT characteristics while offering partial technology exposure, a pure-play data center allocation may provide incremental diversification benefits for income-oriented portfolios, including retirement-focused investment plans. For operators and sponsors, the implication is that organizational form choices may need to accommodate dual investor audiences. For researchers, the implication is that asset class boundaries are not fixed, and pricing identity can migrate even when the underlying physical assets remain unchanged.

The data center sector will continue to evolve as AI workloads intensify and power constraints bind more tightly. Whether the hybrid identity documented here stabilizes, deepens, or reverses remains to be seen. What is clear is that the question of identity—real estate, technology, or something in between—is no longer merely academic. It is priced into the market every day.

Appendix A

Data Center REIT Take-Private Transactions (2021–2022)

Between 2021 and early 2022, three significant data center REITs exited the public market through take-private transactions. These transactions explain why the listed pure-play data center REIT universe narrowed to two constituents (Equinix and Digital Realty) during the sample period.

Table A.1: Data center REIT take-private transactions (2021–2022)

Company	Ticker	Acquirer	Announced	Closed	Deal Value
QTS Realty Trust	QTS	Blackstone	Jun 2021	Aug 2021	$10.0B
CoreSite Realty	COR	American Tower	Nov 2021	Dec 2021	$10.1B
CyrusOne	CONE	KKR & GIP	Nov 2021	Mar 2022	$15.0B

Transaction details:

- **QTS Realty Trust:** Acquired by Blackstone Infrastructure Partners and Blackstone Real Estate Income Trust for $78.00 per share in an all-cash transaction. The purchase price represented a 21% premium to QTS's closing share price as of June 4, 2021.
- **CoreSite Realty:** Acquired by American Tower Corporation for $170.00 per share in an all-cash transaction. The acquisition was structured as a tender offer followed by a second-step merger.
- **CyrusOne:** Acquired by KKR and Global Infrastructure Partners (GIP) for $90.50 per share in an all-cash transaction. This represented the largest data center transaction in history at the time of announcement.

Source: Company press releases and SEC filings.

Appendix B

Green Street REIT Universe

This appendix lists all REITs included in the Green Street REIT universe used to construct the equal-weight benchmarks in this thesis. The Data Center sector (EQIX and DLR) forms the DC basket; all other sectors form the REITxDC benchmark.

Table B.1: Green Street REIT universe by sector

Sector	Tickers
Apartment	AVB, CPT, ELME, EQR, ESS, IRT, MAA, UDR, VRE
Cold Storage	COLD, LINE
Data Center	**EQIX, DLR**
Gaming	GLPI, VICI
Ground Lease	SAFE
Health Care	AHR, CTRE, HR, DOC, MPW, OHI, SBRA, VTR, WELL
Industrial	EGP, FR, LXP, PLD, REXR, STAG, TRNO
Life Science	ARE
Lodging	APLE, DRH, HST, PK, PEB, RLJ, RHP, SHO
Mall	MAC, SPG, SKT
Manufactured Home Park	ELS, SUI
Net Lease	ADC, BNL, EPR, EPRT, NNN, O, WPC
Office	AAT, BDN, BXP, CDP, CUZ, DEI, ESRT, HIW, HPP, JBGS, KRC, PDM, SLG, VNO
Self-Storage	CUBE, EXR, NSA, PSA, SMA
Single-Family Rental	AMH, INVH
Strip Center	AKR, BRX, CURB, FRT, IVT, KIM, KRG, PECO, REG, SITC, UE
Tower	AMT, CCI, SBAC

Notes:

- The Data Center sector (bolded) contains only EQIX and DLR throughout the sample period.
- Technology proxies (XLK and SMH) are ETFs, not REITs, and are not part of this universe.
- Some tickers entered the universe after 2015 due to IPOs, creating an unbalanced panel.

Appendix C

Data Construction and Methodology Details

This appendix provides detailed documentation of the data construction process and empirical implementation for replication purposes.

C.1 Data Source and Sample

Table C.1: Data specifications

Item	Specification
Data source	Yahoo Finance
Price type	Adjusted close (split and dividend adjusted)
Sample period	January 1, 2015 – December 11, 2025
Frequency	Weekly (Thursday grid)
Total observations	572 weekly data points

C.2 Weekly Alignment Rule

The time axis uses a fixed Thursday weekly grid. When a price observation is not available on the exact Thursday date, the nearest prior trading day price is mapped to the Thursday date. This "as-of" mapping prevents look-ahead bias.

C.3 Return Calculations

Log returns (used for correlations and regressions):

$$r_{i,t} = \ln\left(\frac{P_{i,t}}{P_{i,t-1}}\right)$$

Simple returns (used for wealth index charts only):

$$R_{i,t} = \frac{P_{i,t}}{P_{i,t-1}} - 1$$

Log returns are averaged across constituents to construct equal-weight basket returns. Simple returns are compounded to construct base-100 wealth indices.

C.4 Index Construction

Data Center basket (DC):

$$r_t^{DC} = \frac{1}{2}(r_t^{EQIX} + r_t^{DLR})$$

REIT benchmark excluding data centers (REITxDC):

$$r_t^{REITxDC} = \frac{1}{N_t}\sum_{i \in U_t} r_{i,t}$$

where U_t is the set of non-DC REITs with valid returns at week t, and N_t is the count.

C.5 Evolution of REITxDC Benchmark Constituents

The REITxDC benchmark uses an unbalanced panel because some REITs entered the sample after 2015 due to IPOs. Table C.2 summarizes the evolution of constituent counts.

Table C.2: REITxDC benchmark constituent count over time

Year	REITxDC Count (Start)	REITxDC Count (End)
2015	71	73
2016	73	73
2017	74	76
2018	77	79
2019	79	79
2020	79	80
2021	80	81
2022	81	81
2023	81	81
2024	81	85
2025	85	86

C.6 Rolling Correlation

52-week rolling correlations are computed as:

$$\rho_t^{(52)}(X, Y) = \mathrm{Correl}(\{X_{t-51}, \ldots, X_t\}, \{Y_{t-51}, \ldots, Y_t\})$$

The first valid rolling correlation observation appears after the 52-week window build-up period.

C.7 Two-Factor Identity Regression

The regime-level identity regression is:

$$r_t^{DC} = \alpha + \beta_{REIT} \cdot r_t^{REITxDC} + \beta_{Tech} \cdot r_t^{Tech} + \varepsilon_t$$

where r_t^{Tech} is either XLK or SMH log return.

Standard errors are computed using the Newey–West heteroskedasticity and autocorrelation consistent (HAC) estimator with lag parameter set to 8 weeks.

C.8 Rolling Two-Factor Betas

Rolling 52-week betas are estimated using OLS on rolling windows. For each week t with a complete 52-week history, the two-factor model is estimated on observations $\{t-51, \ldots, t\}$, producing time series of $\beta_{REIT,t}$ and $\beta_{Tech,t}$.

C.9 Regime Definitions

Table C.3: Regime definitions

Regime	Start Date	End Date	Economic Context
2015–2019	2015-01-01	2019-12-31	Pre-COVID baseline
2020–2021	2020-01-01	2021-12-31	COVID shock and aftermath
2022–2023	2022-01-01	2023-12-31	Rate-hike stress period
2024–2025	2024-01-01	2025-12-11	Post-hike period

C.10 Software Implementation

All empirical analysis was executed in Microsoft Excel using:

- `LN()` for log returns
- `AVERAGE()` for equal-weight index construction
- `CORREL()` for rolling correlations
- `LINEST()` with `OFFSET()` for rolling regressions
- Custom formulas for Newey–West HAC standard errors

Appendix D

Correlation Matrices by Regime

This appendix reports full pairwise correlation matrices for the four key return series (DC, REITxDC, XLK, SMH) within each regime. These correlations are computed on weekly log returns over the entire regime period, not rolling windows.

D.1 2015–2019 (Pre-COVID Baseline)

Table D.1: Correlation matrix: 2015–2019

	DC	REITxDC	XLK	SMH
DC	1.0000	0.5779	0.2462	0.1417
REITxDC	0.5779	1.0000	0.4128	0.2678
XLK	0.2462	0.4128	1.0000	0.8522
SMH	0.1417	0.2678	0.8522	1.0000

Interpretation: DC correlates more strongly with REITxDC (0.58) than with either tech proxy (0.25 for XLK, 0.14 for SMH), consistent with a predominantly real estate pricing identity.

D.2 2020–2021 (COVID Shock and Aftermath)

Table D.2: Correlation matrix: 2020–2021

	DC	REITxDC	XLK	SMH
DC	1.0000	0.1824	0.3780	0.2042
REITxDC	0.1824	1.0000	0.6330	0.6990
XLK	0.3780	0.6330	1.0000	0.8692
SMH	0.2042	0.6990	0.8692	1.0000

Interpretation: DC–REITxDC correlation collapsed to 0.18, while DC–XLK correlation rose to 0.38. This is consistent with the regime regression finding that data centers tilted toward tech alignment during this period.

D.3 2022–2023 (Rate-Hike Stress Period)

Table D.3: Correlation matrix: 2022–2023

	DC	REITxDC	XLK	SMH
DC	1.0000	0.6436	0.6711	0.6175
REITxDC	0.6436	1.0000	0.6847	0.6095
XLK	0.6711	0.6847	1.0000	0.9226
SMH	0.6175	0.6095	0.9226	1.0000

Interpretation: All correlations are elevated (0.61–0.68 for DC with other series), reflecting stressed market conditions where asset classes move together. DC exhibits similar correlation with both REITxDC and tech proxies, consistent with the hybrid identity finding.

D.4 2024–2025 (Post-Hike Period)

Table D.4: Correlation matrix: 2024–2025

	DC	REITxDC	XLK	SMH
DC	1.0000	0.4115	0.3704	0.3511
REITxDC	0.4115	1.0000	0.1100	0.0215
XLK	0.3704	0.1100	1.0000	0.9332
SMH	0.3511	0.0215	0.9332	1.0000

Interpretation: REITxDC–tech correlations collapsed to near zero (0.11 for XLK, 0.02 for SMH), indicating that REITs and tech stocks moved largely independently. DC maintained moderate correlation with both REITxDC (0.41) and tech proxies (0.37, 0.35), consistent with continued hybrid identity.

D.5 Summary: Evolution of Key Correlations

Table D.5: Selected correlations across regimes

Pair	2015–2019	2020–2021	2022–2023	2024–2025
DC – REITxDC	0.58	0.18	0.64	0.41
DC – XLK	0.25	0.38	0.67	0.37
DC – SMH	0.14	0.20	0.62	0.35
REITxDC – XLK	0.41	0.63	0.68	0.11
REITxDC – SMH	0.27	0.70	0.61	0.02
XLK – SMH	0.85	0.87	0.92	0.93

Bibliography

[1] Alvarez & Marsal. Global data centre insights. Technical report, Alvarez & Marsal, 2024.

[2] Rosemary Batt, Eileen Appelbaum, and Tamar Katz. The role of public reits in promoting shareholder value and financialization: The case of propcos and opco reits in the u.s. lodging industry. Working Paper 189, Institute for New Economic Thinking, July 2022. URL `https://doi.org/10.2139/ssrn.4168375`.

[3] Board of Governors of the Federal Reserve System. Selected interest rates (daily) - H.15. H.15 statistical release (web page; PDF printout). Data Download Program: `https://www.federalreserve.gov/datadownload/Choose.aspx?rel=H15`, January 2026. Release date: January 9, 2026. Last update: January 9, 2026. Accessed: 2026-01-09.

[4] Board of Governors of the Federal Reserve System (US). Federal funds target range - upper limit [DFEDTARU]. FRED series page: `https://fred.stlouisfed.org/series/DFEDTARU`, January 2026. Retrieved from FRED, Federal Reserve Bank of St. Louis. Accessed: 2026-01-10.

[5] Dirk Brounen, David Ling, and Huib Vaessen. The interest rate sensitivity of public real estate. Research paper summary, EPRA Research & Indices, 2016.

[6] Honghui Chen, David M. Harrison, and Mahsa Khoshnoud. Investors' limited attention: Evidence from reits. *The Journal of Real Estate Finance and Economics*, 61:408–442, 2020. doi: 10.1007/s11146-018-9667-y.

[7] Michael Chui, Eric Hazan, Roger Roberts, Alex Singla, Kate Smaje, Alex Sukharevsky, Lareina Yee, and Rodney Zemmel. The economic potential of generative ai: The next productivity frontier. Technical report, McKinsey Global Institute, June 2023.

[8] Dene Dobensky, David Lee, Leesa Overmann, and Sharon Liu. A common question from healthcare reits: Should the trs lease be amended during tough times? KPMG, What's News in Tax (Washington National Tax), May 2020.

[9] Nora Wang Esram and Neal Elliott. Turning data centers into grid and regional assets: Considerations and recommendations for the federal government, state policymakers, and utility regulators. Technical report, American Council for an Energy-Efficient Economy (ACEEE), October 2024.

[10] Michael Fawcett. Contours of the cloud: Data centers, ai, and the future of real estate. Master's thesis, Massachusetts Institute of Technology, Cambridge, MA, September 2024. Submitted to the MIT Center for Real Estate.

[11] FTSE Russell. How reits deliver access to the new economy. Technical report, FTSE Russell, March 2020.

[12] Daniel Greene. Landlords of the internet: Big data and big real estate. *Social Studies of Science*, 52(6):904–927, 2022. doi: 10.1177/03063127221124943.

[13] Ian Hitchcock and Merritt Cahoon. Hyperscaler data center buildout: A sustainability bane, boon, or both? Technical Report NI R 25-04, Nicholas Institute for Energy, Environment & Sustainability, Duke University, 2025.

[14] Hogan Lovells. The world of data centers: Greenfield development and brownfield investments: Success factors, pitfalls and lessons learned. Report/Brochure, 2021.

[15] Houlihan Lokey. Real estate market update: Data centers (summer 2024). Technical report, Houlihan Lokey, 2024.

[16] George Kamiya and Vlad C. Coroamă. Data centre energy use: Critical review of models and results. Technical report, EDNA – IEA 4E TCP, March 2025.

[17] MarketVector Indexes. *MVIS US Listed Semiconductor 25 Index: Index Guide*, January 2026. Version 1.02 (01.2026).

[18] Nestor Maslej et al. Artificial intelligence index report 2025. Technical report, Stanford Institute for Human-Centered Artificial Intelligence, 2025.

[19] Graeme Newell and Muhammad Jufri Bin Marzuki. The increasing importance of the alternate real estate sectors for institutional investors. *Journal of General Management*, 49(1):46–64, 2023.

[20] PGIM Real Estate. Global data centers: Getting connected with a niche sector. Technical report, PGIM Real Estate, February 2021. Investment Research.

[21] Edward F. Pierzak. Data centers: High demand, slow deliveries, big development profits. Nareit (web page; PDF printout)., July 2024. Nareit Industry News. Published: 2024-07-05. Accessed: 2026-01-09.

[22] Neil Rasmussen. Guidelines for specification of data center power density. White Paper 120, APC by Schneider Electric, 2010. Revision 1.

[23] Peter Salvatore and Fisch Gogliormella. Senior housing facility management agreements. *New York Law Journal*, 251(92), May 2014. Transactional Real Estate column; RIDEA and management agreement context.

[24] State Street Global Advisors. Technology select sector SPDR ETF (XLK). State Street / SSGA product page (web page; PDF printout)., January 2026. Holdings as of Jan 8, 2026. Accessed: 2026-01-09.

[25] VanEck. SMH - vaneck semiconductor ETF (us): Holdings & performance. VanEck fund page (web page; PDF printout)., January 2026. As of January 9, 2026. Accessed: 2026-01-09.

[26] Jeffrey M. Wooldridge. *Econometric Analysis of Cross Section and Panel Data*. The MIT Press, Cambridge, MA, 2 edition, 2010.

[27] World Bank. Advancing cloud and data infrastructure markets. Technical report, World Bank, 2021.