Dynamic Modeling of Transaction Prices in Private Equity Markets

Srinivasan Selvam and Timothy Whittaker

Srinivasan Selvam
is a senior researcher at EDHEC Infra & Private Assets Research Institute in Singapore.
srinivasan.selvam@edhec.edu

Timothy Whittaker
is the research director at EDHEC Infra & Private Assets Research Institute in Singapore.
tim.whittaker@edhec.edu

KEY FINDINGS

- **Factor model–based private valuations:** A transaction-based factor model provides real-time, data-driven valuations, reducing reliance on appraisals and smoothing biases in traditional benchmarks.
- **Reliable insights with limited data:** Even with limited transaction data, the model can reconstruct private market valuation dynamics effectively, helping investors navigate illiquid markets and price assets with greater confidence.
- **Practical applications for investors:** By enabling dynamic portfolio valuation, enhancing benchmark accuracy, and aligning with fair-value accounting standards, this approach supports improved risk management, asset allocation, and performance measurement.

ABSTRACT

Problem: Illiquidity in private markets complicates the valuation of private companies. As a result, appraisals use subjective inputs, often resulting in smoothed and stale valuations, which in turn lead to unrepresentative benchmarks that misrepresent the risk and performance of private markets. **Solution:** A factor model calibrated with the latest transaction data can help refine sparse, noisy, and biased inputs into evolving factor prices that, on average, more accurately explain transaction prices. **Findings:** Key factors affecting valuation include size and country risk, which reduce valuation on average, and profitability, leverage, and public market valuation, which increase it. **Application:** The factor prices from the model can be applied to unlisted companies to produce a robust estimate of current market prices, consistent with fair-value accounting standards. The dynamic calibration of the model can also facilitate more frequent valuation and realistic risk estimates, by incorporating incremental information from each new observed transaction. Such an approach can improve asset allocation, performance attribution, and the monitoring of private investments.

Private markets are projected to reach \$13.1 trillion in assets under management (AUM) in 2024, compared to the \$123 trillion market capitalization of public equities (Goldstein 2024; McKinsey 2024). Private equity (PE) continues to serve as the primary investment vehicle for accessing private companies, accounting for 62.9% of the total AUM. By virtue of their sheer scale, private companies play an increasingly pivotal role in the global economy, contributing significantly to GDP, employment, innovation, and environmental impact (e.g., Morgenson 2021; Amess, Stiebale, and Wright 2015; Reclaim Finance 2022).

Despite their growing prominence, investors encounter significant challenges in valuing unlisted private companies, the core holdings within the vast PE market. These holdings lack robust and consistent valuation histories, impeding the development of meaningful, accurate, and representative benchmarks. In the absence of such benchmarks, fund-level indices reliant on appraised valuations have become the default solution. However, these appraisals are fraught with limitations, including smoothing, staleness, and poor representativeness, failing to capture the full breadth of private companies or the wealth of information observable in private markets.

In this article, we propose a factor model approach that robustly explains the valuations observed in private market transactions, effectively addressing the biases inherent in appraised or estimated valuations, such as the 'smoothness' problem. By dynamically estimating the model—recalibrating it each time a new transaction is observed—the framework incorporates all available information in real-time, mirroring the price formation process driven by individual market participants' decisions.

The factor model is calibrated using a global, extensive, and representative dataset of private company transactions. The selection of factors is grounded in established academic research, private market-specific characteristics, and insights from a survey of private equity fund managers conducted by the EDHEC Infra & Private Assets Research Institute. The key factors and their average effects on the price-to-sales ratio (P/S) are presented in Exhibit 1.

The statistically significant predictors of valuation in the model reveal that smaller, high-growth, profitable, highly leveraged, labor-intensive, and innovative firms command higher valuations. These findings are consistent with the view that size proxies for relative liquidity in private markets, thus making larger companies less attractive targets. Moreover, smaller and labor-intensive companies are more amenable to

EXHIBIT 1

Final Factors in the Model

Factor	Source	Avg. Effect on P/S (1999–2022)
Private Company Characteristics		
Size	Prior Research & Survey	Negative
Growth	Prior Research & Survey	Indeterminate
Profitability	Prior Research & Survey	Positive
Book Leverage	Prior Research & Survey	Positive
Labor Intensity	Prior Research	Positive
Patent	Prior Research & Survey	Positive
Hitech	Prior Research & Survey	Positive
Transaction Characteristics		
Deal Leverage	Private Market Features	Positive
Market Characteristics		
Market Valuation	Prior Research & Survey	Positive
Industry Valuation	Prior Research & Survey	Positive
Term Spread	Prior Research & Survey	Negative
Market Volatility	Prior Research	Positive
Sector Price Impact	Prior Research	Positive
Market Price Impact	Prior Research	Positive
PECCS Pillars		
PECCS Classes	Prior Research, Private Market Features, & Survey	Depends on Each Class

NOTES: The effect of size and growth is not statistically significant in ordinary least squares (OLS) models with all factors. Price impact is increasing in liquidity and is measured as the negative Amihud (2002) illiquidity measure.

typical PE value creation strategies, thus making such companies more valuable. Likewise, a high level of leverage along with innovation and growth, signal better future prospects, thereby supporting a higher current valuation. Furthermore, we find transactions are more valuable when market or industry valuations in public markets are elevated, term spreads are narrower, and public market liquidity and volatility are higher.

In addition, private companies in the education, financials, health, natural resources, and real estate sectors exhibit a valuation premium, while those in the remaining sectors are discounted at least in some periods in the sample. Similarly, companies employing a subscription revenue model, targeting end-consumers (individuals), or offering a hybrid mix of products and services achieve higher valuations. Notably, these effects are time-varying, reflecting evolving market conditions and investor preferences.

The remainder of the article is organized as follows: First, we examine the limitations of traditional private equity valuation methods that rely on low-quality multiples (e.g., public market multiples or reported data) or insufficient data points (e.g., a stale and limited set of recent transactions), emphasizing the gaps and potential distortions these approaches create. Second, we outline the motivation for our factor selection and describe the composition of our sample. Third, we detail our empirical approach. Fourth, we present the results of our asset pricing model applied to transaction data and discuss the economic interpretation of these findings. Fifth, we evaluate the robustness of the model. Finally, we conclude by discussing the practical applications of the model, including its use in shadow pricing private companies regularly, constructing benchmarks, and providing dynamic inputs to anchor valuation estimates.

ISSUES WITH TRADITIONAL APPROACHES

Why Are Private Market Valuations Challenging?

Under International Financial Reporting Standards (IFRS) 13 standards, the fair value of financial assets is unequivocally defined as the exit price on the valuation date—that is, the price at which the asset would sell in the market on that date, irrespective of its actual liquidity. However, this poses a significant challenge in private markets, where neither recently traded prices for the private company nor reliable benchmarks from comparable peers are readily available. Additionally, private market valuations are plagued by substantial information asymmetry and subjectivity, making each valuation exercise inherently complex and resource-intensive.

Given the inherent uncertainty surrounding future cash flows and the indefinite time horizons of private companies, market-based valuation methods are typically favored over income-based approaches. Nonetheless, challenges emerge when attempting to implement market-based valuations using a limited set of transaction-based multiples. Recent transactions are often sparse, noisy, and biased, as only a small subset of private companies transact regularly, and each valuation is influenced by unique, idiosyncratic factors. Consequently, such valuations often fail to provide investors with reliable or actionable insights.

To address the scarcity of data, practitioners often adjust recent transaction-based multiples for an extensive list of subjective differences, rendering the valuation process less scientific and more akin to an art. Similarly, listed company multiples serve as poor proxies due to fundamental differences between public and private markets and require further ad hoc and static illiquidity "adjustments." As a result, rather than enabling objective valuation estimation, current methods often provide a framework to rationalize predetermined valuations.

Lack of Recent Transaction Data

The severity of data scarcity is evident in Exhibit 2, which illustrates the limited number of transaction data points typically available in the market at any given time. For instance, MSCI's quarterly private capital holdings data (formerly Burgiss) show that only around 150 recent transactions are observed in a given quarter, with certain sectors scarcely represented.

Flaws in Reported Data

Reported data, such as those in the first two columns in Exhibit 2, present aggregate multiples that blend general partner (GP)-appraised valuations with actual transaction data. This aggregation results in a composite metric influenced by varying methodologies, assumptions, and decision-making processes of individual stakeholders. Even if the two components are separated, relying solely on reported valuation multiples remains problematic due to several well-documented shortcomings, including but not limited to the following:

- **Timeliness:** Due to delays in data collection, reporting, and aggregation, these multiples are often stale, lagged, and asynchronous. Furthermore, when GPs rely on static assumptions—such as a fixed internal rate of return (IRR) as the discount rate or failing to update valuation inputs to reflect market changes, the reported multiples can remain unchanged across several quarters.
- **Smoothness:** Behavioral biases among GPs result in reported valuations that are excessively smooth and deviate from reality, exhibiting low volatility across quarters. This artificially smooth behavior effectively obscures the true risk profile of the asset class.
- **Lack of granularity and robustness:** Aggregated multiples at the sector or regional level are insufficiently granular and less applicable to specific assets without significant ad hoc adjustments. For instance, substantial valuation differences often exist between subsectors within a single sector. Relying on aggregated sector multiples can result in significant valuation errors.
- **Combining disparate modeling approaches**: Reported multiples often rely on models, but these models are applied inconsistently by various GPs or analysts, introducing substantial noise. While the noise could theoretically cancel out if all GPs valued the same company, the uniqueness of each asset being valued exacerbates inconsistencies in methods and assumptions. These idiosyncratic biases amplify inaccuracies in the reported data.

EXHIBIT 2

Underlying Data That Captures Recent Transactions

	All Multiples		Entry Multiples		Exit Multiples	
Sector	Count	Median	Count	Median	Count	Median
All Sectors	1,209	12.4x	114	10.1x	43	12.2x
Comm. Services	62	13.8x	7	11.0x	2	N/A
Consumer Disc.	195	11.0x	14	7.6x	10	12.4x
Health	203	14.1x	22	12.1x	9	14.2x
Industrials	270	10.5x	29	9.1x	11	12.0x
Information Tech.	259	15.7x	23	11.1x	6	12.3x
Materials	62	10.6x	9	9.5x	1	N/A

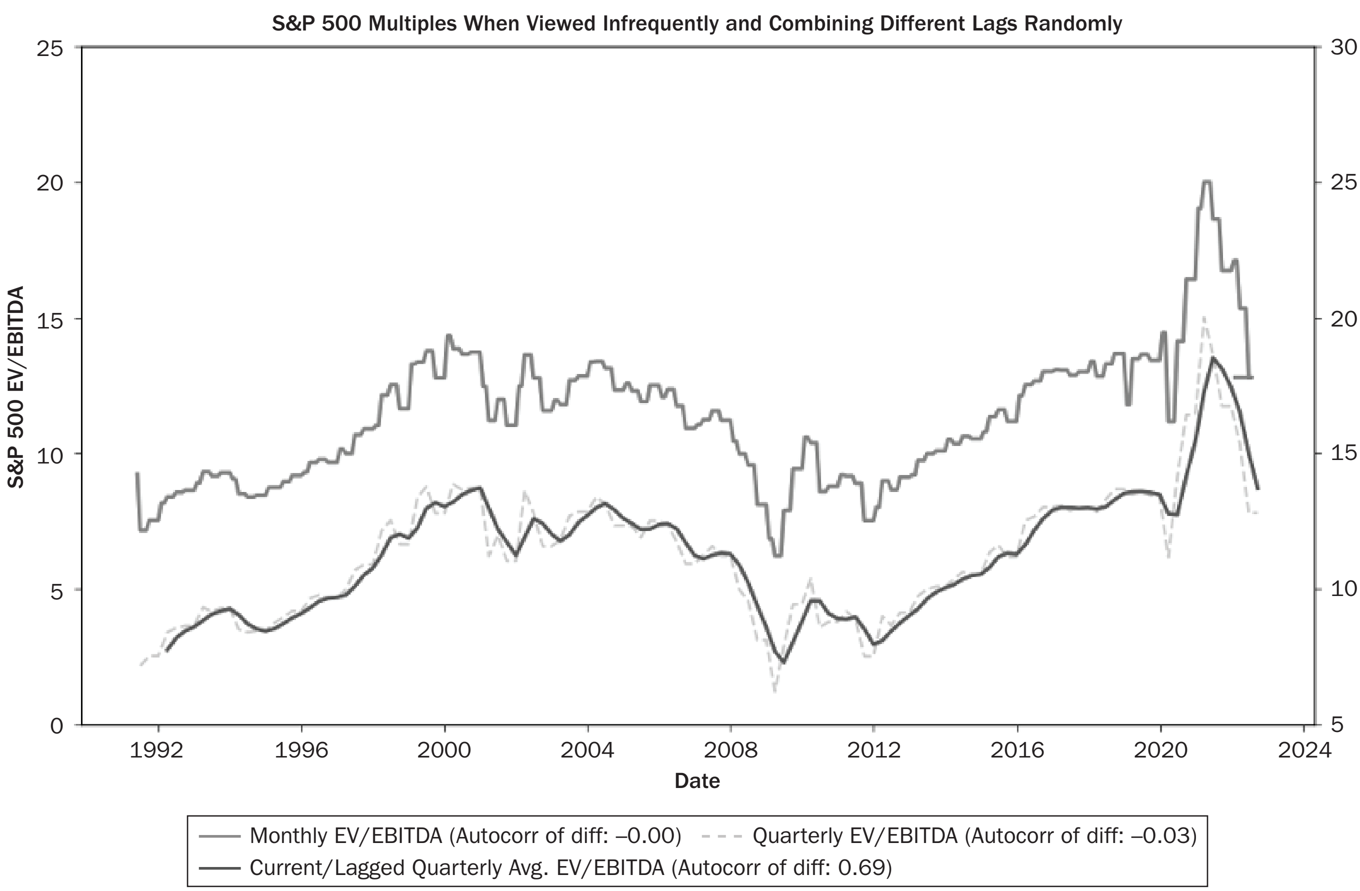
SOURCES: MSCI/Burgiss Q2, 2022 Report. Median EBITDA multiples from Burgiss Q2 2022.

However, despite the consensus on smoothness in private market performance, some studies such as Rudin and Farley (2022) extol the information content in private market reported data. Such data can provide better signals about future expected fundamentals than public markets, which suffer from Shiller's excess volatility problem (Shiller 1981)—that is, stock prices fluctuate way too much, which is unexplained by changes in earnings or dividend expectations. Although our study differs from Rudin and Farley (2022), we complement their work by proposing a regression-based approach using transaction data that avoid desmoothing and leverage adjustment assumptions (e.g., Rudin et al. 2019). Transactions in private markets arguably capture future earnings expectations with more clarity than reported valuations, thus complementing their work.

In Exhibit 3, we illustrate the interplay between timeliness and smoothness in valuations. Specifically, we demonstrate how combining valuations with varying lags introduces a smoothing effect. First, we plot the monthly enterprise value/earnings before interest, taxes, depreciation, and amortization (EV/EBITDA) multiples for all S&P 500 stocks over time. The monthly differences in this series exhibit an autocorrelation of 0, indicating no persistence. Second, on the right-hand axis, we plot the same EV/EBITDA series using quarterly measurements, which appear similar to the monthly series. However, when we combine EV/EBITDA values from any of the

EXHIBIT 3

Biases from Combining Valuation of Different Lags of Multiples



SOURCE: Calculated from Bloomberg data on the S&P 500 Index.

last four quarters—assigning random weights to each lagged or current value—the resulting series becomes notably smooth, with a high autocorrelation of 0.69.

This analysis demonstrates how smoothness can inadvertently emerge in private valuations when data from different lags is combined. This issue arises when a data provider aggregates net asset values (NAVs) from multiple funds, each voluntarily reporting with varying lags, even in the absence of behavioral biases from GPs to smoothen valuations. The tendency of GPs to report less-volatile NAVs further exacerbates the smoothness of reported figures, amplifying the issue.

To further highlight the granularity and robustness challenges in reported data, we apply these metrics to private market transactions. Using the median EV/EBITDA values from MSCI/Burgiss multiples for Q2 2022, we compute implied prices for over 120 deals sourced from PitchBook. The absolute errors are calculated as the differences between the implied EV and actual EV based on the deal prices. Exhibit 4 presents these errors across private equity company classification standards (PECCS) activities and at the aggregate level. For sectors without specific estimates, aggregate medians from MSCI/Burgiss are applied. The results reveal significant inaccuracies, with absolute errors exceeding 50% for the majority of activities, highlighting the lack of robustness in reported values.

We extend the analysis to a subset of deals in the healthcare sector, where the granularity problem becomes particularly evident. The sample comprises two distinct types of healthcare companies with significant differences in their valuations. For instance, healthcare equipment and supplies consistently trades at higher EV/EBITDA multiples compared to healthcare providers and services. However, the lack of granularity in the aggregate healthcare multiples reported in the data leads to systematically biased errors in the implied EV multiples for these two subsectors, irrespective of which quartile multiple is applied.

The application of reported multiples to these two subsectors within the healthcare sector is shown in Exhibit 5. This analysis underscores the representativeness challenges that arise when broad sector-level multiples are applied to diverse subsectors.

EXHIBIT 4

Valuation Errors by Sectors in Deals Valued Using Reported Multiples

Activity	Absolute Error in Deal EV minus Reported Data Implied EV
Information and Communication	56%
Transportation	95%
Manufacturing	73%
Real Estate and Construction	81%
Hospitality and Entertainment	113%
Retail	60%
Professional and Other Services	68%
Utilities	96%
Health	46%
Financials	55%
Natural Resources	33%
Education and Public	30%
All Deals	**74%**

SOURCES: 120 PitchBook deals in 2022 and MSCI/Burgiss Q2 2022 report. Activity classification based on PECCS activities.

EXHIBIT 5

2022 Deals (PitchBook) and Reported Multiples (Burgiss 2022) in Healthcare Sector

2022 Healthcare Deals					Reported Healthcare Multiples			Estimation Error			
Co. Name	Sector	Deal Date	Deal Multiple	EBITDA M$	Top Q	Med.	Bottom Q	Upper Bound	Med.	Lower Bound	Avg. Abs. Error
Natus Medical	Health Care Equipment & Supplies	7/5/22	22.8x	48.6	18.3x	14.1x	10.9x	20%	38%	52%	37%
Artel		6/13/22	28.0x	5				35%	50%	61%	48%
IntriCon		5/24/22	51.2x	4.3				64%	72%	79%	72%
Hanger	Health Care Providers & Services	10/3/22	12.2x	100.74				–50%	–16%	10%	26%
Probo Medical		3/8/22	15.0x	30				–22%	6%	27%	18%
Tivity Health		6/28/22	19.5x	157.7				6%	28%	44%	26%
							Avg. Abs. Error	**33%**	**35%**	**46%**	**38%**

SOURCES: PitchBook deals in the healthcare sector in 2022 and MSCI/Burgiss Q2 2022 Report.

Inadequacy of Public Market Proxies

Private companies are deeply interconnected with the broader economy, as they must navigate market demand, secure financing, withstand business cycle fluctuations, and access capital markets for investor exits. These linkages expose them to factors similar to those affecting publicly listed companies. However, significant differences between the two asset classes render public peers poor proxies for valuation.

First, a large number of public companies have delisted, leaving public markets more concentrated and skewed toward larger firms (e.g., Doidge et al. 2018). Second, the use of leverage is more prevalent in private equity ownership compared to public companies (Chingono and Rasmussen 2015). Third, public markets allow investors to diversify portfolios with greater ease, making a liquid and easily diversifiable asset class a poor proxy for the inherently illiquid private market assets.

To illustrate these challenges, we compute the implied EV for over 120 PitchBook deals in 2022 using EV/EBITDA multiples from publicly listed companies. The dataset, sourced from (Damodaran 2024), provides EV/EBITDA multiples across 91 subsectors of US companies. By mapping these subsectors to the activities in the PitchBook dataset, we calculate the implied EVs. As shown in Exhibit 6, the valuation errors are even larger than those observed when using reported data. Additionally, over half of the sectors exhibits valuation errors exceeding 100%, indicating that private market companies transacted at significantly higher multiples than those of their publicly listed peers in 2022.

Consequences of the Valuation Challenges

The absence of reliable valuation histories in private markets hinders the development of meaningful, accurate, and representative benchmarks. Consequently, most available benchmarks rely heavily on appraised valuations, which are fraught with several well-documented biases. These include smoothing, staleness, and poor representativeness, and they fail to capture the full spectrum of private companies or all of the information observable in private markets. This lack of robust benchmarks perpetuates a vicious cycle of valuation inefficiencies, as many investors base their own valuations on these flawed benchmarks.

EXHIBIT 6

Valuation Errors by Sectors in Deals Valued Using Publicly Listed Multiples

Activity	Absolute Error in Deal EV minus Publicly Listed Peers' Implied EV
Information and Communication	70%
Transportation	107%
Manufacturing	78%
Real Estate and Construction	143%
Hospitality and Entertainment	179%
Retail	54%
Professional and Other Services	109%
Utilities	207%
Health	45%
Financials	355%
Natural Resources	11%
Education and Public	48%
All Deals	**103%**

SOURCES: 120 PitchBook deals in 2022 and Damodaran NYU dataset for 2022.

In summary, these valuation challenges contribute to, but are not limited to, the following market phenomena:

- Reliance on appraisal-based benchmarks that are unrealistically smoothed, lagged, and coarse.
- Misaligned risk expectations and observation of unrealistic Sharpe ratios.
- Distortions to asset allocation strategies, including the denominator effect.

THE FACTORS AND DATA

Our Approach

While several studies examine private equity fund performance using fund-level data (e.g., Korteweg and Nagel 2024; Harris et al. 2023; Kaplan and Schoar 2005), there is limited research on the valuation of private companies and the factors driving their performance. This gap is largely attributable to data scarcity. Unlike publicly traded equities, private companies transact infrequently, making it difficult to compute standard return metrics.

Transaction-based pricing models provide a potential solution to the limitations of conventional valuation methods including discounted cash flow (DCF) and multiples-based approaches. These models, including repeat sales indices (Bailey, Muth, and Nourse 1963) and hedonic pricing models (Malpezzi 2003), offer a more objective and statistically rigorous framework by leveraging actual transaction data.

For example, repeat sales indices, widely used in real estate, track price changes of assets sold multiple times. Similarly, hedonic pricing models, also prevalent in real estate, account for both internal characteristics and external factors affecting asset prices. Blanc-Brude and Tran (2019) develop a hedonic model to explain the prices of unlisted infrastructure assets based on their transaction prices and key characteristics. Their estimation employs a dynamic linear model that accommodates time-variation in factor prices. A similar approach can be applied to private companies, using observed transactions to estimate factor prices. Implementing such a model requires identifying a comprehensive list of potential factors.

Factor Motivation

Academic research and private market characteristics provide a foundation for identifying potential factors that influence the valuation of private companies. Below, we discuss key prospective factors and their expected impact on valuation.

- **Size:** Since Fama and French (1992), size has been recognized as a critical factor in asset pricing across various asset classes, including equities, bonds (Fama and French 1993), and infrastructure (Blanc-Brude and Tran 2019). Among private companies, a small-cap return premium may exist, as smaller firms are often perceived as riskier due to illiquidity, limited access to capital, and shorter operational histories, resulting in lower valuations. Conversely, size could positively influence valuation (or negatively affect returns) by reflecting the higher relative liquidity of smaller firms over lumpier larger ones. Also, the growth potential and entrepreneurial dynamism of smaller firms can provide a valuation boost. This dual perspective underscores the importance of empirical analysis to ascertain the precise relationship between size and valuation.
- **Leverage:** The relationship between leverage and valuation is multifaceted. Traditional finance theory suggests that higher leverage increases financial risk, thereby necessitating higher expected returns and resulting in lower valuations. However, empirical evidence is mixed, showing positive, neutral, and even negative associations between leverage and required returns (George and Hwang 2010; Gomes and Schmid 2010). In private equity, leverage is often employed strategically to enhance returns through financial engineering. Additionally, leverage decisions during acquisitions can serve as a signal of a company's perceived prospects, with higher leverage potentially reflecting both elevated risk and greater expected returns.
- **Growth and Profitability:** The seminal work of Fama and French (1992) emphasizes the relationship between growth and valuation, particularly the dichotomy between "value" and "growth" stocks. Growth stocks, typically exhibiting high price-to-book ratios, are expected to deliver superior future earnings growth, often commanding higher valuations. However, growth is inherently challenging to predict, and market inefficiencies or investor biases can contribute to mispricing. Similarly, profitability is a fundamental driver of valuation, as evidenced by studies such as Novy-Marx (2013) and Hou, Xue, and Zhang (2015), which highlight that more-profitable companies are generally perceived as more valuable while at the same time requiring a higher rate of returns, posing a puzzle.
- **Other Company Characteristics:** Factors such as human capital, technological capability, age, corporate governance, and industry-specific attributes can significantly influence a company's valuation. Value creation in private equity strategies often relies on operational, financial, and governance engineering, and these characteristics can present both opportunities and risks for executing such strategies effectively.
- **Sector:** An asset's industry plays a critical role in its valuation, as factors like investment opportunities, expected rates of return, competitive dynamics, and regulatory frameworks are often consistent within an industry. To account for industry effects, we employ the PECCS taxonomy developed by the EDHEC Infra & Private Assets Research Institute. PECCS is a multipillar, private market–specific taxonomy that categorizes private companies across multiple dimensions or risk factors, each with an independent influence on valuation. These dimensions include the company's industry (or activity), life cycle phase, revenue model, customer model, and value chain (PECCS 2023). Exhibit 12 presents the various pillars and classes in the PECCS taxonomy, along with their observed effects on valuation in the transaction data.

- **Other Controls:** Beyond these characteristics, market conditions, sector trends, and deal-specific structuring can also explain variations in observed valuations. For instance, strategic acquisitions (e.g., add-ons) may be valued differently from public-to-private transactions. Similarly, deals executed during economic expansionary periods may differ substantially from those completed during recessions. Additionally, specific sectors may exhibit periods of heightened or diminished activity, further influencing valuations.

The above factors were also presented to private equity fund managers in a 2023 survey conducted by the EDHEC Infra & Private Assets Research Institute. The fund managers were asked to identify the factors they deemed important, rank these factors by significance, and indicate whether they expected a positive or negative impact on valuation. The survey results, based on 95 responses, are presented in Exhibit 7.

The survey results align with expectations: growth, profitability, and revenue emerged as the top three factors positively influencing valuation, as indicated by the number of respondents. Furthermore, these three factors were consistently ranked as the most important, underscoring their critical role in valuation.

Sample Construction

We construct our sample of transactions by looking at PE investments in private companies using data from PitchBook™, a Morningstar company, spanning 1999–2022. For modeling purposes, the P/S ratio is preferred over other metrics such as price-to-EBITDA (P/EBITDA) or return-based measures. P/EBITDA is less suitable because negative EBITDA renders valuation ratios meaningless, and the nonstandard adjustments to EBITDA further complicate its use. Additionally, return metrics are excluded to avoid focusing solely on assets that are repeatedly traded, thereby neglecting the majority of transactions.

EXHIBIT 7
Survey Results on Factors Influencing Valuation

Factor	Response	Average Rank	Average Direction of Effect on Valuation
Profitability	67%	2.11	0.94
Growth	64%	1.80	0.95
Revenue	49%	2.30	0.94
Industry Valuation	40%	3.26	0.53
Size	32%	3.20	0.87
Leverage	27%	3.08	–0.46
Market Valuation	23%	3.32	0.55
Market Share	18%	3.41	1.00
Long Term Rate	16%	3.07	–0.87
Competition for Deals	14%	3.62	0.77
Synergy	11%	3.40	0.80
Technology	8%	3.50	1.00
Ownership Structure	7%	2.43	–0.43
Maturity	4%	2.50	0.50
Short-Term Rate	4%	3.50	–1.00
Human Capital	2%	1.00	1.00

SOURCE: Survey data on private equity fund managers' 95 responses on factors affecting valuation.

Our sample construction process is detailed in Exhibit 8. The percentages in each row represent the reduction in sample size compared to the previous step, culminating in a final global sample of 5,438 transactions.

Exhibits 9 and 10 illustrate the number and aggregate size of transactions by year and sector during the sample period of 1999–2022. Manufacturing and information and communication are the two largest activities in the sample in terms of deal count. However, in terms of transaction value, information and communication account for a larger proportion, reflecting the tendency for larger average transaction sizes within this activity.

Descriptive statistics for the explanatory variables derived from our factor motivation are presented in Exhibit 11. These variables are classified into two groups: key explanatory variables and optional explanatory variables. The key explanatory variables are always included in the model, while optional variables are incorporated only if they significantly enhance the model's explanatory power, as determined by a feature selection algorithm.

EXHIBIT 8

Sample Construction

Step Description	% Reduction
Filtering within PitchBook	
1. Transaction Types: PE Excluding PIPE	–95.00%
2. Status is Completed	–1.32%
3. Deal Size of \$10 Million	–81.72%
4. Revenue of \$5 Million	–66.36%
Filtering during Processing	
5. Missing Revenue	–20.56%
6. Missing Founding Year	–6.97%
7. Missing Profitability	–30.84%
8. Missing Public Market Data	–0.08%
9. Missing Employees	4.94%
10. Exclude Outliers	–10.68%
Final Sample of 5,438 Transactions	

SOURCES: PitchBook and data-processing filters.

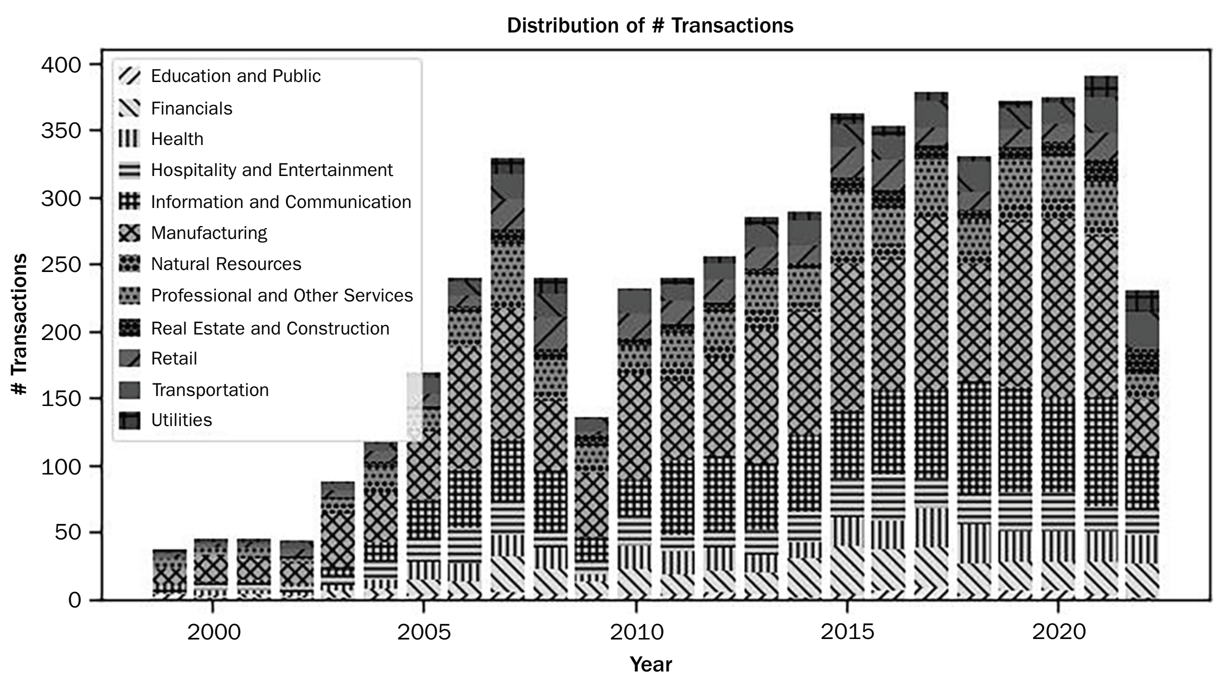
EMPIRICAL APPROACH

Feature Selection

The inclusion of numerous potential factors does not necessarily lead to an optimal model. While adding additional factors may reduce the residual

EXHIBIT 9

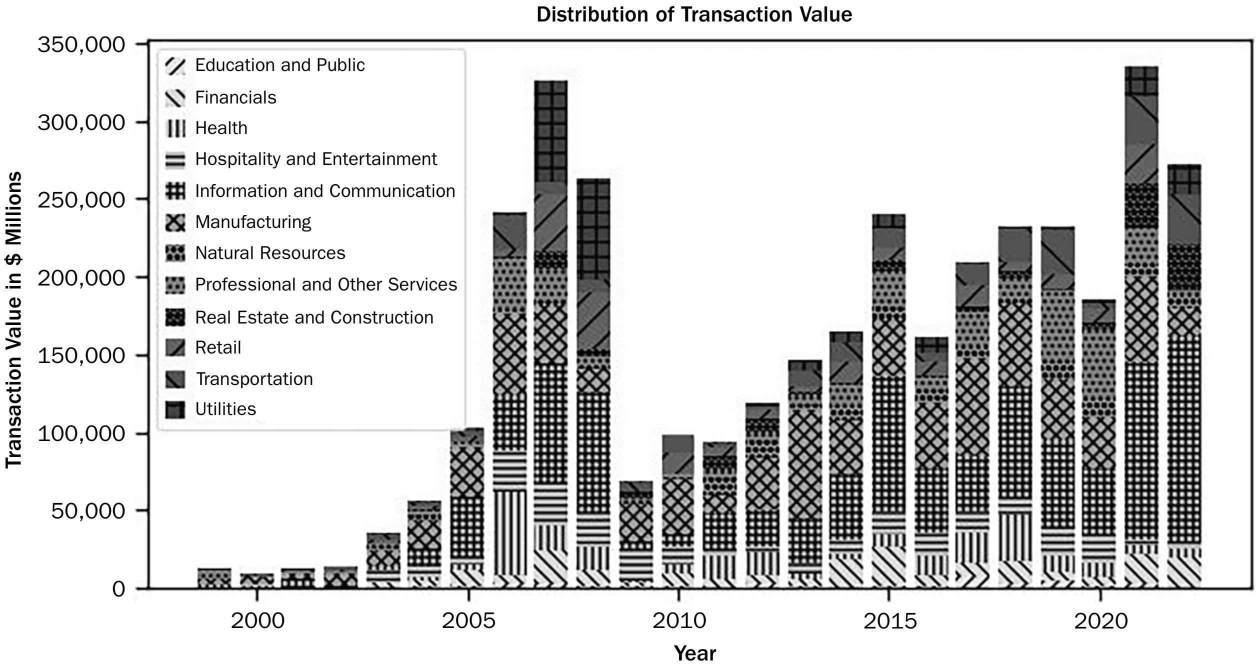
Transaction Counts by Year and Activity



SOURCE: PitchBook PE Transactions.

EXHIBIT 10

Transaction Value by Year and Activity



SOURCE: PitchBook PE transactions.

EXHIBIT 11

Summary Statistics for Variables Used in the Analysis

Variable	Obs	Mean	Std. Dev.	Min.	Median	Max.
P/S*	5,438	2.76	2.73	0.25	1.78	14.21
Key Explanatory Variables						
Size*	5,438	463.9	1,760.3	5.1	96.6	37,818.1
Growth*	5,438	0.47	7.04	–0.96	0.06	311.51
Profitability	5,438	14.48	26.89	–142.60	12.84	127.64
Book Leverage*	5,438	0.61	4.10	0.00	0.00	161.08
Market Valuation*	5,438	1.26	0.39	0.48	1.15	2.36
Industry Valuation*	5,438	1.82	1.67	0.10	1.28	12.29
Term Spread*	5,438	0.02	0.01	–0.05	0.02	0.08
Optional Explanatory Variables						
Age*	5,438	35.40	34.41	–1.00	23.00	214.00
Patent	5,438	0.26	0.44	0.00	0.00	1.00
Market Share*	5,438	0.28	1.38	0.00	0.04	51.72
Deal Leverage*	5,438	0.21	0.77	0.00	0.00	11.23
Herding*	5,438	0.15	0.10	0.00	0.12	0.42
Pe Back	5,438	0.62	0.49	0.00	1.00	1.00
Long-Term Interest Rate*	5,438	0.03	0.02	–0.01	0.03	0.16
Industry Concentration*	5,438	1,186.4	1,101.8	222.9	857.3	7,931.4
Forex Change	5,438	0.00	0.07	–0.43	0.00	0.33
VIX*	5,438	18.32	6.95	10.13	16.68	62.67
Gdp Growth	5,438	0.03	0.08	–0.34	0.04	0.43
Cpi Growth	5,438	0.02	0.02	–0.02	0.02	0.30

(continued)

EXHIBIT 11 *(continued)*

Summary Statistics for Variables Used in the Analysis

Variable	Obs	Mean	Std. Dev.	Min.	Median	Max.
Labor Intensity*	5,438	13.89	34.57	0.04	5.40	423.47
Control	5,438	0.71	0.31	0.01	0.80	1.00
Addon	5,438	0.17	0.37	0.00	0.00	1.00
Public co	5,438	0.06	0.23	0.00	0.00	1.00
Similarity	5,438	0.12	0.02	0.06	0.11	0.27
Size Factor	5,438	1.41	9.07	–25.37	0.32	54.00
Value Factor	5,438	–1.22	13.40	–35.01	–1.85	73.84
Momentum Factor	5,438	1.40	14.70	–56.40	2.96	54.93
Profitability Factor	5,438	2.81	7.73	–37.24	2.32	58.31
Investment Factor	5,438	0.63	7.41	–11.01	–0.86	44.82
Size Private Factor	5,438	1.35	0.64	–0.42	1.32	3.10
Dry Powder*	5,438	7.12	3.66	2.46	6.42	20.83
Market Volatility*	5,438	0.04	0.02	0.01	0.04	0.09
Market Price Impact*	5,438	0.00	0.00	0.00	0.00	0.00
Sector Price Impact*	5,438	0.00	0.00	0.00	0.00	0.00
Hitech	5,438	0.23	0.42	0.00	0.00	1.00
Emerging Country	5,438	0.03	0.17	0.00	0.00	1.00

NOTE: Variables marked with * are subject to log or other transformations when used in regressions.

sum of squares in a regression, it can also introduce econometric issues such as multicollinearity and overfitting. To address this, we employ statistical approaches designed to identify a subset of predictors most relevant to valuation. The selected model aims to maximize explanatory power while maintaining robustness, ensuring it performs reasonably well on out-of-sample data—that is, factor prices derived from one set of transactions should remain relevant when applied to another sample.

To identify the best subset of regressors, we perform a least absolute shrinkage and selection operator (LASSO) method. Prior work has found that LASSO feature selection regressions perform well in feature selection and avoid overfitting (e.g., Hastie, Tibshirani, and Tibshirani 2017). The LASSO method regularizes model parameters by shrinking the regression coefficients, reducing some of them to zero. The feature (or variable) selection phase occurs after the shrinkage, where every nonzero value is selected to be used in the regression model.

In implementing the LASSO, however, we do include the required variables such as key characteristics and sector dummies irrespective of whether they are selected or not by the LASSO algorithm. This combined approach ensures the features selected have economic and statistical motivation to be considered, making the model simple, intuitive, and practical.

Dynamic Linear Models

A simple linear model such as the ordinary least squares (OLS) model captures only the average effect of a factor on valuation, overlooking temporal variations in investor preferences. For instance, business cycles, often attributed to systematic pricing errors in financial markets, cause expansions and recessions over time. Assuming a constant average effect for a factor ignores these dynamic valuation perspectives.

Consider the case of certain sectors becoming overvalued during speculative bubbles. When these bubbles burst, valuations plummet dramatically. A static OLS model fails to account for these nuances, assuming instead a uniform average effect of a sector on valuation over time.

To address this limitation, we estimate a dynamic linear model (DLM). A DLM allows factor prices—representing how much investors are willing to pay for a particular characteristic of a firm—to evolve over time. This approach reflects traditional price formation patterns in markets, where investor preferences vary temporally and drive valuation dynamics. Details of the DLM and its mathematical formulation are presented in the appendix.

RESULTS

OLS Estimation

Although estimating the factor model using OLS regressions provides valuable insights, it has notable limitations. These include the presence of autocorrelation in valuation metrics (e.g., the Ljung–Box test returns an χ^2 statistic of 69.2, rejecting the null hypothesis of no autocorrelation in P/S, violating OLS assumptions) and time variation in the correlation between P/S and explanatory variables (e.g., as shown in Exhibit 14).

Despite these limitations, OLS regressions serve as an initial step in exploring the relationship between factors and P/S, offering intuitive interpretations of these associations. The results of OLS regressions using only the required variables are presented in Exhibit 12.

The findings indicate that smaller, more-profitable, and highly leveraged private companies tend to transact at higher valuations, as discussed in the Dynamic Estimation section. Additionally, transactions are valued higher when public markets—particularly the same sector—are valued more favorably, and when term spreads are narrower. The coefficients for these variables are statistically significant at the 5% level or better, indicating their meaningful impact on valuation.

Regarding the PECCS indicators, compared to manufacturing (the omitted indicator), companies in the health, financials, and natural resources sectors are valued higher, while those in the retail and hospitality and entertainment sectors are valued lower. For life cycle phases, companies in the startup and growth phases command higher valuations compared to mature companies. In terms of revenue models, subscription-based models are valued at a premium relative to production-based models. Similarly, companies operating consumer-focused customer models achieve higher valuations than those with business-focused models. Finally, firms with service-oriented value chains are valued higher than those with product-oriented value chains.

Other Predictors

To complement the required variables, we evaluate a broad set of observable variables presented in Exhibit 11, using the LASSO algorithm framework.

Leveraging both the set of mandatory regressors, determined a priori based on theoretical relevance, and a larger pool of optional regressors, we implement the LASSO method. We use fivefold cross-validation for estimation. This approach systematically splits the training data into five parts, iteratively training the model on four folds and validating on the remaining fold, thereby not losing valuable training data that may be assigned as a test sample.

The LASSO hyperparameter α, which controls the strength of the penalty on model complexity, is tuned to minimize mean squared error across the validation folds. A high value of α allows more features while compromising on model complexity, and vice versa. We determine the optimal alpha as the value that minimizes the cross-validated mean squared error across the five folds. This method allows us to select eight additional variables that improve the model while being parsimonious and interpretable. The results are presented in Exhibit 13.

EXHIBIT 12
OLS Regression Results

Explanatory Variables	Estimate	*t*-Statistic	Std. Error
Intercept	3.246**	(2.59)	1.255
Size	–0.102***	(–11.23)	0.009
Growth	0.005	(0.18)	0.028
Profitability	0.007***	(15.37)	0.000
Book Leverage	0.065***	(6.78)	0.010
Market Valuation	0.067***	(5.83)	0.011
Industry Valuation	0.374***	(10.55)	0.035
Term Spread	–3.909**	(–2.19)	1.783
PECCS Indicators			
Activity Education & Public	0.073	(0.64)	0.114
Activity Financials	0.151**	(2.05)	0.073
Activity Health	0.175***	(2.87)	0.061
Activity Hospitality & Entertainment	–0.163**	(–2.13)	0.077
Activity Information & Communication	0.079	(1.29)	0.061
Activity Natural Resources	0.282***	(4.02)	0.070
Activity Professional & Other Services	–0.112*	(–1.68)	0.066
Activity Real Estate & Construction	0.054	(0.63)	0.085
Activity Retail	–0.202***	(–2.95)	0.068
Activity Transportation	–0.118	(–1.55)	0.077
Activity Utilities	–0.107	(–1.12)	0.096
Lifecycle Growth	0.056**	(2.16)	0.026
Lifecycle Startup	0.220***	(5.02)	0.044
Revenue Model Advertising	0.036	(0.63)	0.058
Revenue Model Subscription	0.140***	(3.16)	0.044
Revenue Model Reselling	–0.005	(–0.11)	0.049
Customer Model Consumer Focused	0.105***	(3.70)	0.028
Value Chain Services	0.114**	(2.24)	0.051
Value Chain Hybrid	0.341	(4.69)	0.073
Observations		**5,438**	
Adjusted R^2		**0.205**	

NOTES: Variables are transformed as indicated in Exhibit 11. Significance levels: *p < 0.1, **p < 0.05, ***p < 0.01.

The interpretation of the coefficients in the optimal OLS model, based on the required variables, remains similar, although size loses statistical significance. Additionally, all variables added from the LASSO algorithm are statistically significant at the 5% level or better. The results indicate that private companies command higher valuations when they are purchased with higher deal leverage, are labor-intensive, possess patents, belong to high-tech sectors, and are acquired during periods of increased public equity market liquidity and volatility. The optimal model yields an adjusted R^2 of 0.27. Furthermore, the model's variance inflation factor (VIF) is only 1.88, indicating no significant multicollinearity issues.

Dynamic Estimation

To examine the extent of time-variation in factor prices, the sample period is divided into 10 equal-sized groups, each containing approximately 543 observations. Pearson correlation coefficients between P/S and the key explanatory variables are computed for each period.

EXHIBIT 13

Optimal OLS Regression

Explanatory Variables	Estimate	*t*-Statistic	Std. Error
Intercept	2.418*	(1.89)	1.279
Size	–0.002	(–0.18)	0.011
Growth	0.002	(0.06)	0.027
Profitability	0.007***	(16.97)	0.000
Book Leverage	0.053***	(5.72)	0.009
Market Valuation	0.082***	(7.36)	0.011
Industry Valuation	0.282***	(8.04)	0.035
Term Spread	–6.576***	(–3.69)	1.781
Deal Leverage	0.136***	(13.38)	0.010
Labor Intensity	0.091***	(10.55)	0.009
Patent	0.200***	(7.33)	0.027
Market Volatility	0.064**	(2.52)	0.026
Sector Price Impact	0.025**	(2.17)	0.012
Market Price Impact	0.245***	(4.45)	0.055
Hitech	0.167***	(5.70)	0.029
Includes PECCS Indicators		**Yes**	
Variance Inflation Factor		**1.881**	
Observations		**5,438**	
Adjusted R^2		**0.266**	

NOTES: Variables are transformed as indicated in Exhibit 11. Significance levels: *p < 0.1, **p < 0.05, ***p < 0.01.

The results, presented in Exhibit 14, reveal two notable patterns. First, the length of the periods (represented by rows) varies significantly, highlighting that private company transactions occur in clusters. Second, the correlations exhibit time-dependent variation that is inconsistent across variables. For instance, growth has the highest correlation with P/S during the 2016–2017 period, while no other explanatory variable reaches a similar peak during that time frame, except for labor intensity.

To address the static nature of OLS regressions and account for the time-varying behavior of factor prices, we estimate a DLM using the selected explanatory variables. Specifically, coefficients are estimated on each transaction date, incorporating new information as transactions occur. The filtered coefficient estimates from the dynamic model capture all available information up to the transaction date, mirroring the way private equity investors might integrate their information sets into pricing decisions.

For presentation purposes, we report the smoothed coefficient estimates, which leverage the entire sample to produce smoother trends in predictions by incorporating all data from the sample period. This smoothing process provides a more stable view of factor price dynamics over time.[1] The distributions of factor prices during the sample period, as derived from the dynamic estimation, are presented in Exhibit 15.

Note that the interpretation of the economic magnitudes of the effects depends on whether other variables are held constant and whether the predictor is log-transformed. For predictors that are not log-transformed, the interpretation is derived by exponentiating the coefficient and subtracting one, yielding the percentage change in P/S for a one-unit change in the predictor. Conversely, for log-transformed predictors, the coefficient is directly interpreted as the percentage change in P/S for a 1% change in the predictor variable.

[1] The filtered DLM results closely align with the smoothed results and are therefore not reported separately.

EXHIBIT 14

Time-Varying Correlation between P/S and Key Explanatory Variables

Period	1	2	3	4	5	6	7	8	9	10	11	12
1999–2005	–0.28	0.06	0.31	0.28	–0.07	0.13	–0.07	0.32	0.21	–0.09	0.09	0.08
2005–2007	–0.19	–0.08	0.29	0.22	–0.15	0.23	–0.06	0.26	0.15	–0.03	0.04	0.01
2007–2010	–0.29	0.00	0.20	0.25	0.04	0.17	–0.10	0.37	0.16	–0.08	0.04	0.04
2010–2013	–0.28	0.01	0.12	0.29	–0.11	0.14	–0.04	0.36	0.24	–0.07	0.10	0.02
2013–2014	–0.24	0.02	0.21	0.22	–0.05	0.19	0.00	0.33	0.12	–0.06	0.08	–0.01
2014–2016	–0.28	0.04	0.27	0.28	–0.16	0.19	0.02	0.29	0.13	0.03	0.01	–0.02
2016–2017	–0.28	0.10	0.16	0.20	–0.11	0.17	–0.04	0.35	0.27	0.00	0.06	–0.01
2017–2019	–0.24	0.06	0.15	0.20	–0.23	0.24	–0.04	0.34	0.25	0.08	0.04	–0.04
2019–2020	–0.23	0.07	0.23	0.23	–0.22	0.25	0.01	0.26	0.15	–0.01	0.17	0.06
2020–2022	–0.24	0.05	0.12	0.24	–0.16	0.29	0.02	0.31	0.18	–0.09	0.15	–0.01

NOTE: Variables are as follows: 1. size, 2. growth, 3. profitability, 4. book leverage, 5. market valuation, 6. industry valuation, 7. term spread, 8. deal leverage, 9. labor intensity, 10. market volatility, 11. sectoral price impact, and 12. market price impact.

EXHIBIT 15

Descriptive Statistics of Smoothed Factor Prices from the Dynamic Linear Model

Variable	Mean	Median	Min.	Max.	Std. Dev.
Intercept	0.0177	0.0181	0.0134	0.0220	0.0023
Size	–0.0083	–0.0090	–0.0582	0.0242	0.0151
Growth	0.0266	0.0148	–0.0533	0.1239	0.0598
Profitability	0.0096	0.0094	–0.0077	0.0284	0.0060
Book Leverage	0.0595	0.0591	0.0156	0.1012	0.0185
Market Valuation	–0.0368	–0.0411	–0.0496	0.0156	0.0126
Industry Valuation	0.1553	0.1661	0.0156	0.1904	0.0356
Term Spread	0.0114	0.0113	0.0081	0.0156	0.0012
Deal Leverage	0.1202	0.1255	0.0156	0.1607	0.0257
Labor Intensity	0.0937	0.0942	0.0156	0.1530	0.0299
Patent	0.1769	0.1882	0.0156	0.2402	0.0615
Market Volatility	–0.0113	–0.0149	–0.0302	0.0207	0.0131
Sector Price Impact	0.0195	0.0181	–0.0080	0.0817	0.0189
Market Price Impact	0.0245	0.0228	–0.0028	0.0620	0.0154
Hitech	0.1409	0.1640	0.0156	0.1969	0.0528

NOTE: PECCS indicators are included in the DLM and also have time-varying effects on valuation but are not reported here.

The effects of the key factors in the results and their economic interpretations are discussed below:

- **Size:** Size has a negative effect on valuations measured as P/S, indicating that larger companies are valued lower per dollar of revenue compared to smaller companies. This contrasts with public markets, where small caps are generally valued lower than large caps due to their greater exposure (or β) to systematic factors.[2]

[2] Investors in public markets undervalue small caps to account for their higher perceived risk, attributable to a higher cost of capital (Banz 1981), greater information asymmetry and illiquidity (Armstrong et al. 2011; Damodaran 2005), and increased variability in future cash flows (Hodgson and Stevenson-Clarke 2000).

In private markets, however, size is observed to have the opposite effect on valuation. Two primary economic reasons can explain this phenomenon. First, larger companies are more illiquid, making investments in them lumpier and concentrating the GP portfolio, which increases the investor's risk. Second, a significant source of private equity value creation—growth in revenue and profits—may be harder to achieve for larger companies. Traditional revenue and profit-increasing strategies are often already implemented in larger firms, making further enhancements more challenging. Both explanations align with the observed empirical evidence of lower valuations for larger private companies. Additionally, the illiquidity of private markets and the higher undiversifiable risk of large private companies impose binding limits to arbitrage (Shleifer and Vishny 1997), sustaining the size-related valuation wedge.

The mean coefficient estimate indicates that a 10% increase in size leads to a 0.1% decrease in P/S. Furthermore, the negative effect of size on valuations has increased over the years, reflecting a widening small-cap premium in private market valuations.

- **Growth:** Theoretical models suggest that value firms may be riskier during economic downturns, requiring higher rates of return and thus exhibiting lower valuations (Zhang 2005). With more assets in place than investment options, value firms are more vulnerable to business cycle fluctuations, whereas growth firms can delay or forgo investments. Furthermore, systematic pricing errors may occur when investors evaluate growth opportunities, leading to overvaluation of growth companies, particularly during favorable economic conditions.

  While growth is expected to have a positive association with private company valuations, the sample provides no unconditional support for this relationship on average. Some periods, particularly around 2010, are characterized by a negative effect of growth on valuation. However, in post-2015 years, a clear shift emerges, with investors paying higher premiums to acquire private companies experiencing robust growth.
- **Profitability:** Although productive assets are inherently valuable, they can also command higher required returns, as greater profitability is equivalent to a leveraged claim on revenues (Novy-Marx 2013), posing a valuation puzzle. Hou, Xue, and Zhang (2015) argue that if a firm's investment levels are fixed, higher profitability implies higher costs of capital. The reasoning is that, with lower costs of capital, firms would invest more and generate even greater profits.

  Consequently, studies based on publicly listed companies predict an ambiguous relationship between profitability and current valuation. In the context of private companies, however, our results reveal a positive association between profitability and valuation. This divergence may reflect the greater use of leverage by GPs and their ability to facilitate capital requirements, mitigating the demand for a higher cost of capital for productive assets.

  Our findings indicate that, holding other variables constant, a one-unit increase in profitability increases the P/S by 1.0%.
- **Leverage:** Theoretical and empirical studies of publicly listed companies present divergent conclusions regarding the effect of leverage on valuation. In private markets, however, leverage is often endogenously determined by GPs, with lower leverage employed for riskier portfolio holdings and higher leverage applied to less risky assets. For instance, value creation in private equity often involves a two-stage process: derisking a company by improving its cash flows, followed by rerisking it through new debt issuance such as for dividend recapitalizations. Moreover, financing conditions and investment opportunities are highly correlated, suggesting that leverage may serve as a

signal of lower required returns and more-favorable investment opportunities, thereby positively influencing valuation.

Our findings support this hypothesis. Leverage exhibits a positive effect on the valuation of private companies, consistent with the initial OLS regression results. Moreover, both book leverage and deal leverage have a positive effect on valuation. The average coefficient indicates that a 10% increase in book leverage and deal leverage, respectively, leads to a 0.6% and 1.2% increase in P/S.

- **Market conditions:** Market conditions significantly influence the valuation of private companies. The model incorporates several measures to capture these conditions, including the market valuation factor, term spread, industry valuation, and public market volatility and liquidity metrics. The results indicate that private company valuations are higher in environments characterized by lower country risk, higher sector valuation in public markets, lower aggregate market volatility and valuation, and higher liquidity in public markets in the aggregate and in the specific sector of the private company.

Among other factors, we find that labor intensity, patents, and high-tech classifications positively influence valuation. PECCS classifications also impact valuation, with their effects exhibiting greater dynamism compared to other factors. This observation aligns with the view that sector preferences are more time dependent.

ROBUSTNESS

Predictions versus Raw Valuations

To evaluate how closely the predicted valuations align with the raw transaction-based valuations, we compute the 12-month moving average P/S and plot the two series. For context, the monthly P/S of key public market benchmarks are also included. As shown in Exhibit 16, the moving average of the model-predicted valuations closely tracks the raw valuations, with the two series exhibiting a high correlation coefficient of 0.98.

Additionally, the average transaction valuation at any given time is strongly correlated with public market valuations, as evidenced by a correlation coefficient of at least 0.78. These results confirm that the model does not introduce mechanical artifacts in the predicted valuations and is well aligned with the raw data, providing further confidence in its robustness.

Accuracy of Predictions across PECCS Pillars

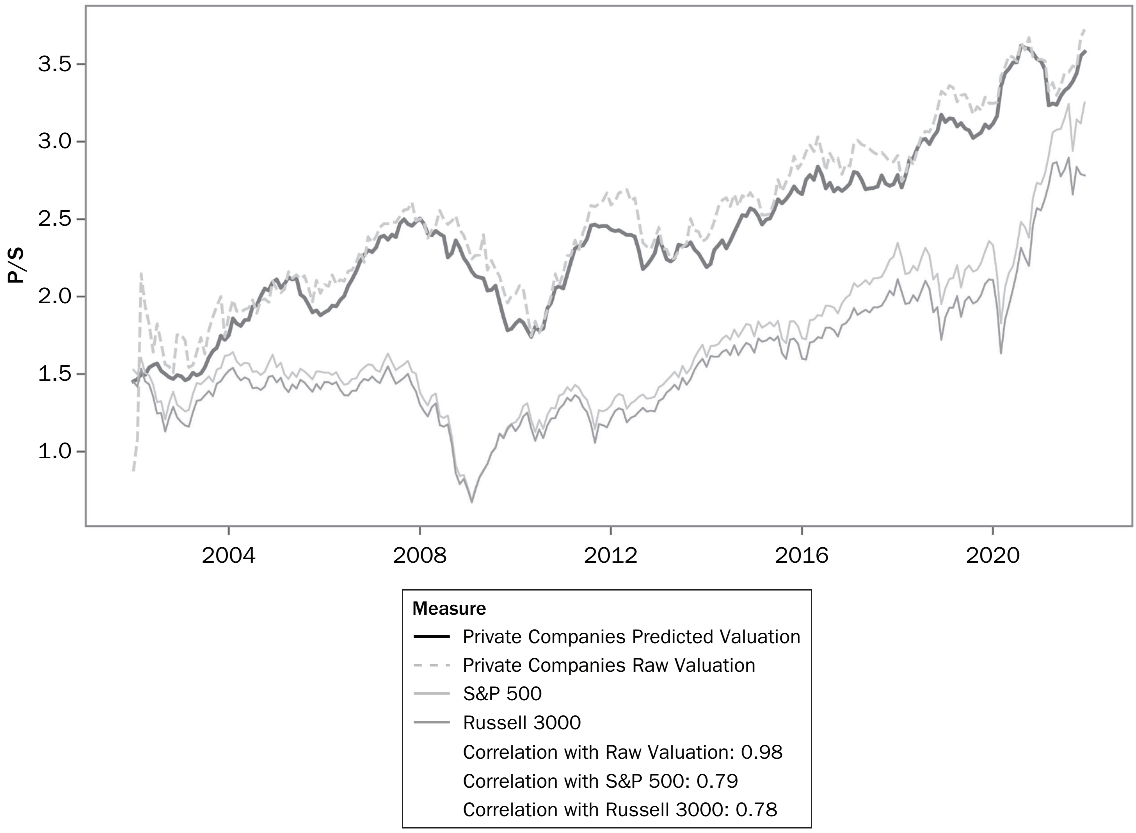
We also assess the accuracy of the predicted valuations within PECCS classes. Notably, while the model is designed to minimize estimation errors in the full sample, the estimation errors within individual PECCS classes are also remarkably low. Almost all errors fall within ±10%, providing reassurance that the model's predictions are reasonable and consistent across segments of PECCS. The summary of these errors is presented in Exhibit 17.

The most commonly used valuation metric in private markets is EV/EBITDA, as GPs can alter the capital structure and are therefore more focused on operational profitability without accounting for interest costs. However, our model uses P/S because it is statistically more robust, stable, and unaffected by loss-making companies. This raises a potential concern that our predictions for EV/EBITDA might be biased.

To address this concern, we compute the EV using the book value of debt and the predicted equity valuation, and then divide the total by EBITDA to derive a predicted EV/EBITDA. We compare this predicted ratio to transaction-implied ratios. As shown

EXHIBIT 16

Moving-Average Predicted versus Raw Valuation and Public Market Valuation



SOURCE: Based on predictions from the DLM smoothed model of transaction valuations.

EXHIBIT 17

PECCS Mean Estimation Errors

PECCS Pillar	PECCS Class	Mean Estimation Error	PECCS Class	Mean Estimation Error	PECCS Pillar
	Education	0.5%	Startup	0.7%	**PECCS**
	Financials	2.0%	Growth	–1.2%	**Lifecycle**
	Health	2.4%	Mature	2.3%	**Phase**
	Hospitality & Ent.	–1.3%	Advertising	0.3%	
	Info. Comm.	–5.0%	Reselling	4.7%	**PECCS**
PECCS Activity	Manufacturing	2.9%	Production	3.1%	**Revenue**
	Natural Resources	10.2%	Subscription	–7.6%	**Model**
	Prof. Services	3.0%	B2B	0.8%	**PECCS Cust.**
	Real Est. & Constr.	9.7%	B2C	1.3%	**Model**
	Retail	–0.3%	Hybrid	0.9%	**PECCS**
	Transportation	6.5%	Products	0.5%	**Value**
Full Sample		1.0%	Services	3.9%	**Chain**

SOURCE: Based on predictions from the DLM smoothed model of observed transactions.

in Exhibit 18, the average predicted and observed ratios are consistent, alleviating concerns about bias in our predictions.

Model Residuals

The factor model is designed to capture the systematic effects of observable factors on valuation while isolating the idiosyncratic "noise" inherent in transactions. Analyzing the model's residuals provides insight into whether this objective has been achieved. As shown in Exhibit 19, the residuals are centered around zero, with their distribution closely resembling a normal distribution, indicating that the model performs as intended.

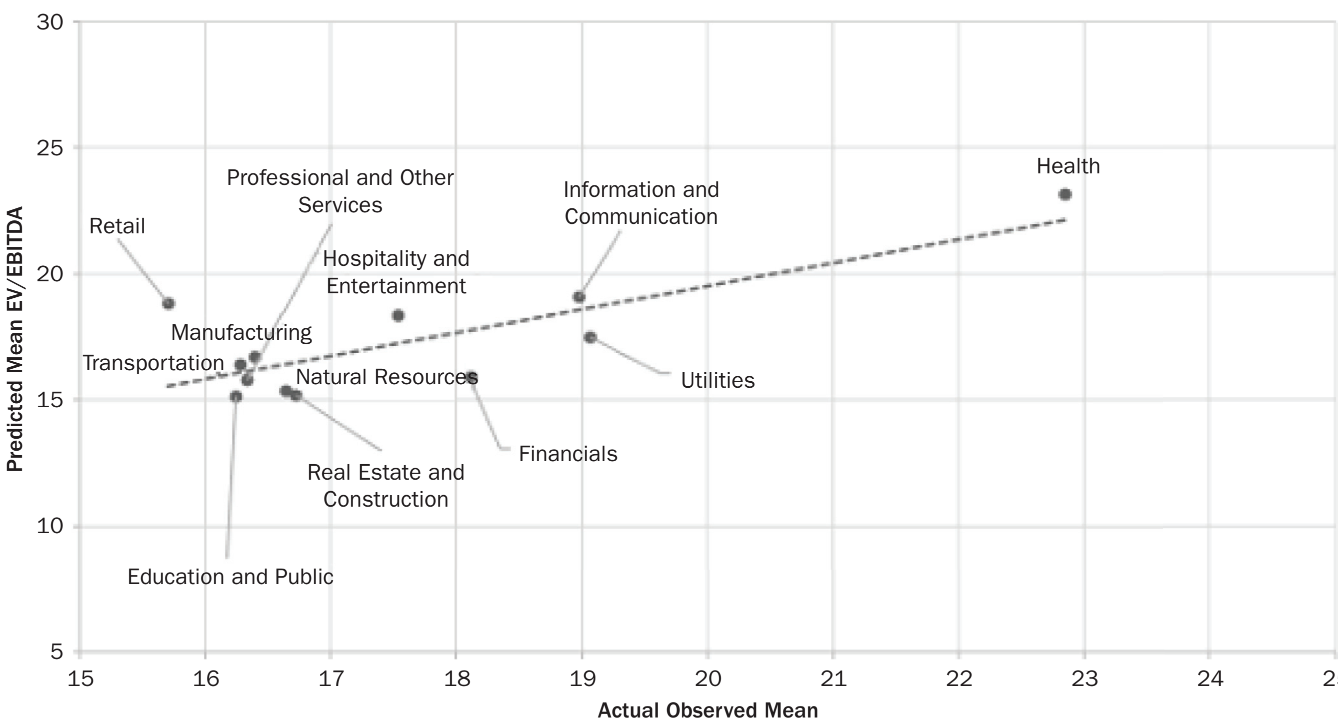
Transaction Reliance

The nonreliance on reported data in constructing the factor model avoids the obvious reporting biases in private markets. However, by requiring transaction data, the model may be subject to other biases, such as transaction clustering. That is, transactions may be more observable in hot than cold markets. Clustering does happen more when valuations are high and vice versa, as can be seen in Exhibit 9. For example, the year 2007 represented a peak in the number of transactions that was not surpassed till 2015. If that is the case, it raises the concern whether the approach is more biased to favorable valuation periods, as clustering tends to provide more data during favorable periods.

Despite being a concern, the dynamic linear method of estimating factor prices already accounts for the noisiness and sparseness of the training data and aims to

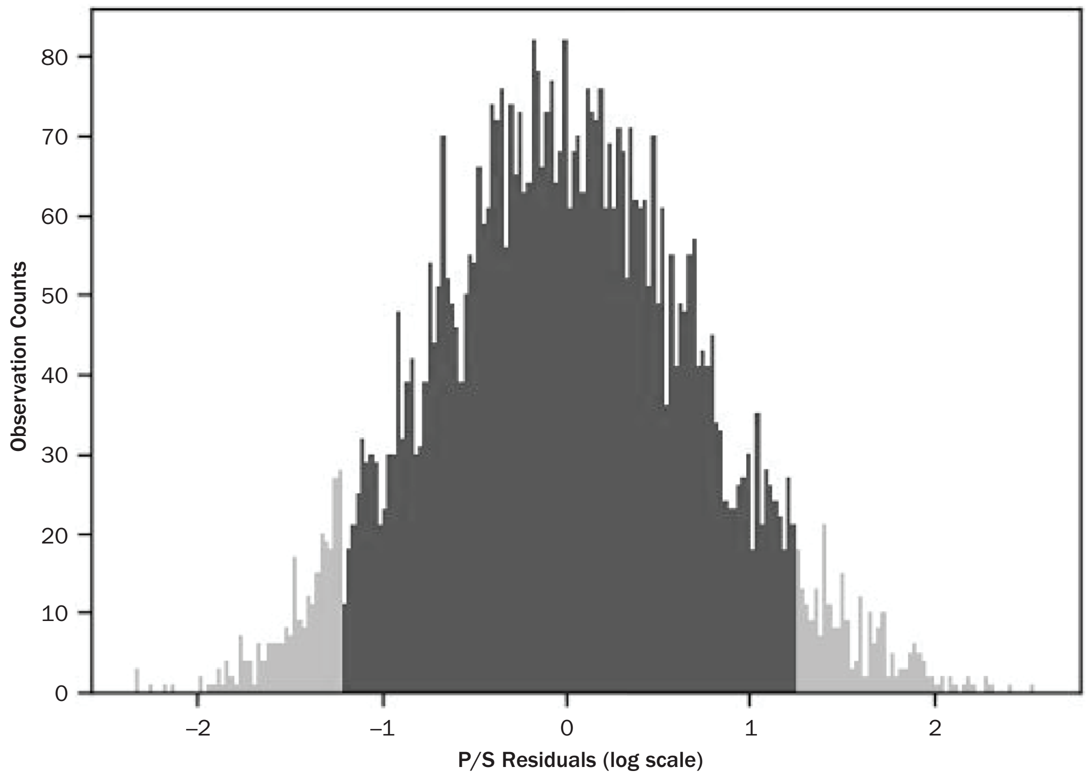
EXHIBIT 18

Predicted versus Actual EV/EBITDA by PECCS Classes



SOURCE: Based on predictions from the DLM smoothed model of transaction valuations.

EXHIBIT 19
Distribution of Model Residuals



SOURCE: Based on predictions from the DLM smoothed model of transaction valuations.

estimate the unbiased factor prices. However, it still is critical to understand what could be the bias due to clustering. To mitigate this concern, we perform the following. First, we examine what is the extent of clustering in the observed data. Second, we assume hypothetical scenarios where part of the transaction data is not observed and reconstruct the series of predicted valuations and see how that relates to all of the observed valuations, similar to Exhibit 16.

Transaction Clustering. To measure the clustering of transactions, we use the Lorenz curve and the Gini coefficient to assess the time-series patterns in transactions.

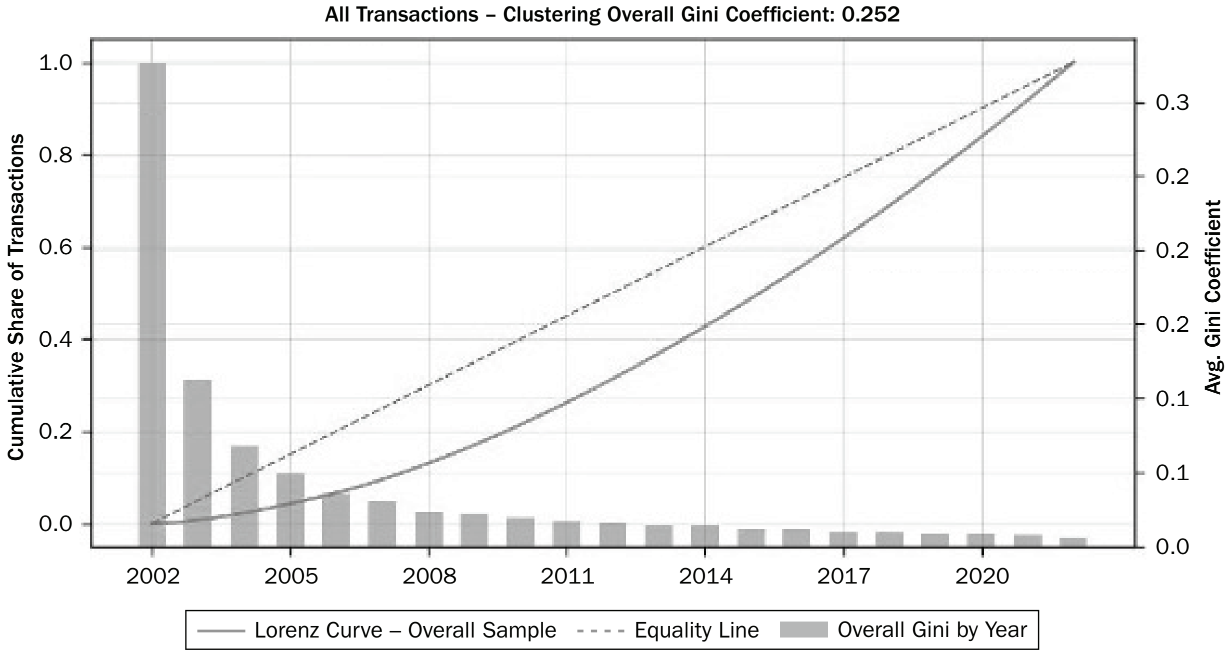
The Lorenz curve is a graphical representation of the distribution of a variable, typically used to measure inequality or concentration. It plots the cumulative share of a variable (y-axis) against the cumulative share of time or another independent variable in the x-axis. This curve can then be compared with an equality line—that is, a 45-degree diagonal line, which represents perfect uniformity. The distance between the Lorenz curve and the equality line indicates the degree of clustering—greater deviation from the diagonal suggests stronger concentration in specific periods.

Similarly, the Gini coefficient is a statistical measure of dispersion used to quantify inequality in a distribution, commonly applied to measure income or wealth inequality. In our context, we adapt it to measuring the temporal distribution of transactions.

As shown in Exhibit 20, the Lorenz curve plotted on the primary y-axis (left) shows that the curve bows below the equality line, thus indicating the presence of clustering. However, we can also see that through the years, the wedge between the curve and the line diminishes, indicating that as private markets have evolved and

EXHIBIT 20

Clustering in All Transactions



SOURCE: PitchBook PE Transactions.

grown, clustering is becoming less evident and transactions are happening more uniformly over time.

Likewise, a more uniform distribution of transactions through time would result in a low Gini coefficient. In contrast, a high Gini coefficient signifies a strong concentration of transactions in specific periods, indicating clustering. We compute different levels of Gini coefficients including for the overall sample, by year, by activity, and by activity-year. The Exhibit 20 shows the Gini coefficients overall and by year, plotted on the secondary *y*-axis. We find that the overall coefficient is 0.252. The highest observed in any year is in 2002, the starting year, at 0.317, and it has been continuously decreasing over the years, consistent with what the Lorenz curve indicated.[3]

Repeating the exercise by PECCS activity, as shown in Exhibit 21, provides similar takeaways that clustering exists even in sectors and is more pronounced earlier than in the later parts of the sample period, with some variation across the sectors.

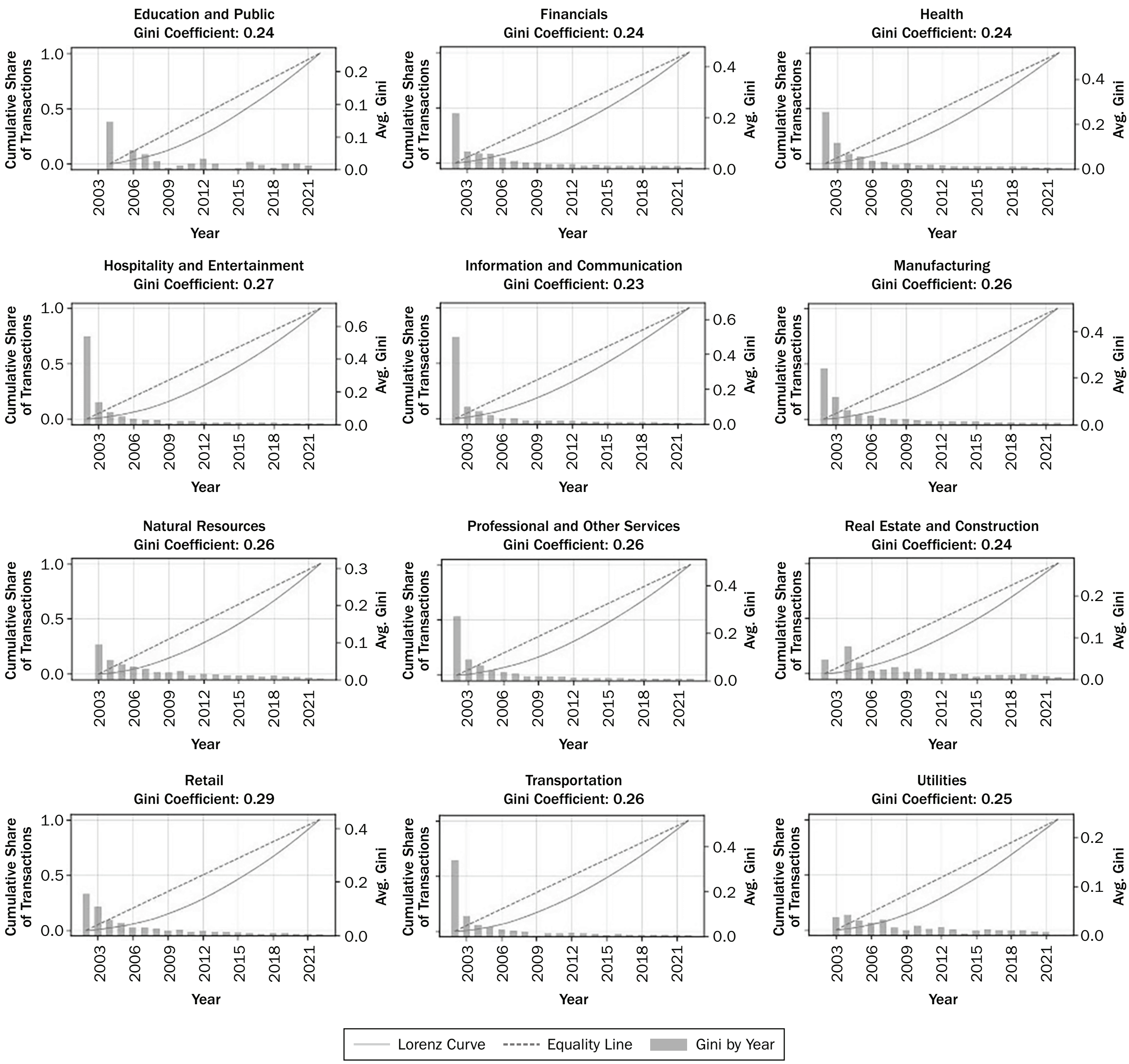
Observing Partial Transactions. To explore how well or poorly the model performs if we cannot observe all the transactions, we perform a counterfactual exercise by using the same set of factors determined as before and train the model on different subsets of the transaction data. Then we compare how the market dynamics look if we were to only use the subset of transactions and compare it to the market dynamics indicated by the observed valuation in all of the transactions.

First, we repeat the baseline model with all the transactions, shown in the top panel of Exhibit 22. We find that the correlation between the moving average of the predicted valuations and the moving average of the raw valuations is 0.98. To measure the fitness, we also compute a sum of the squared deviations between the two curves (i.e., raw and predicted) and standardize it by dividing by the square of the

[3] Because of the absence of transactions in some sectors before 2002, we use data only from 2002 in this analysis.

EXHIBIT 21

Clustering in Transactions by PECCS Activity



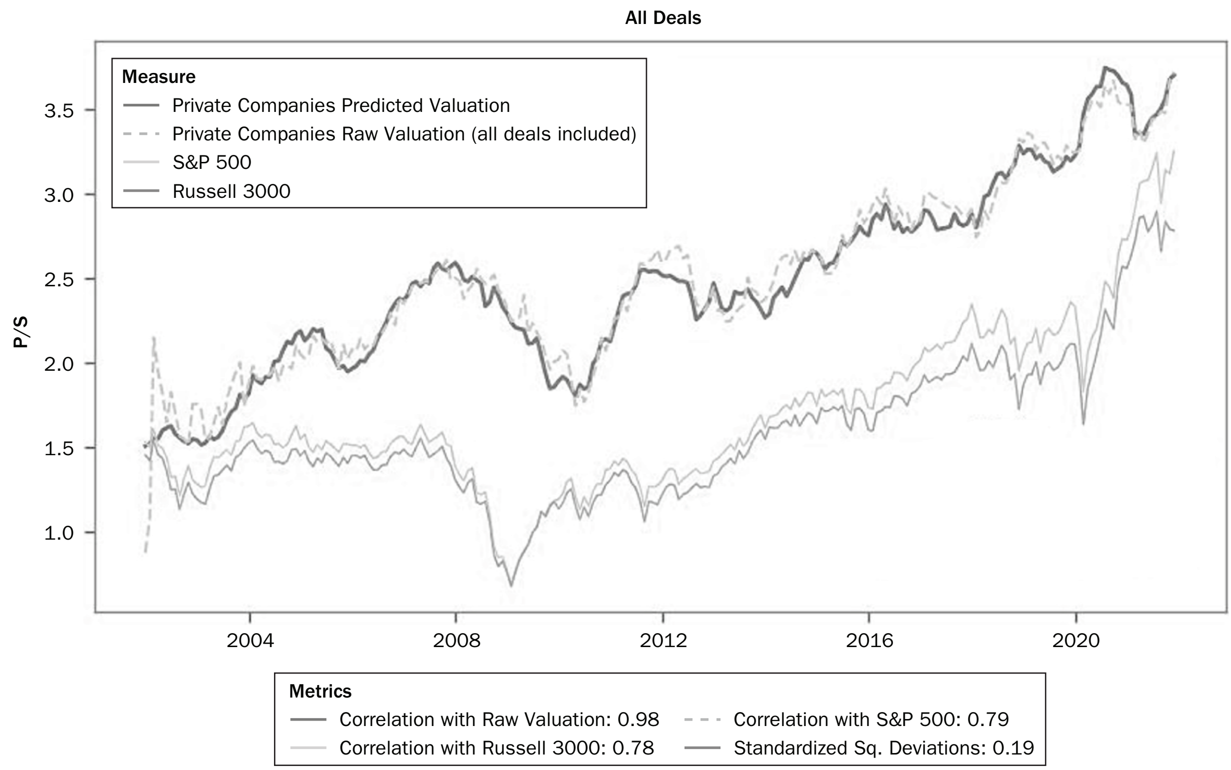
SOURCE: PitchBook PE Transactions.

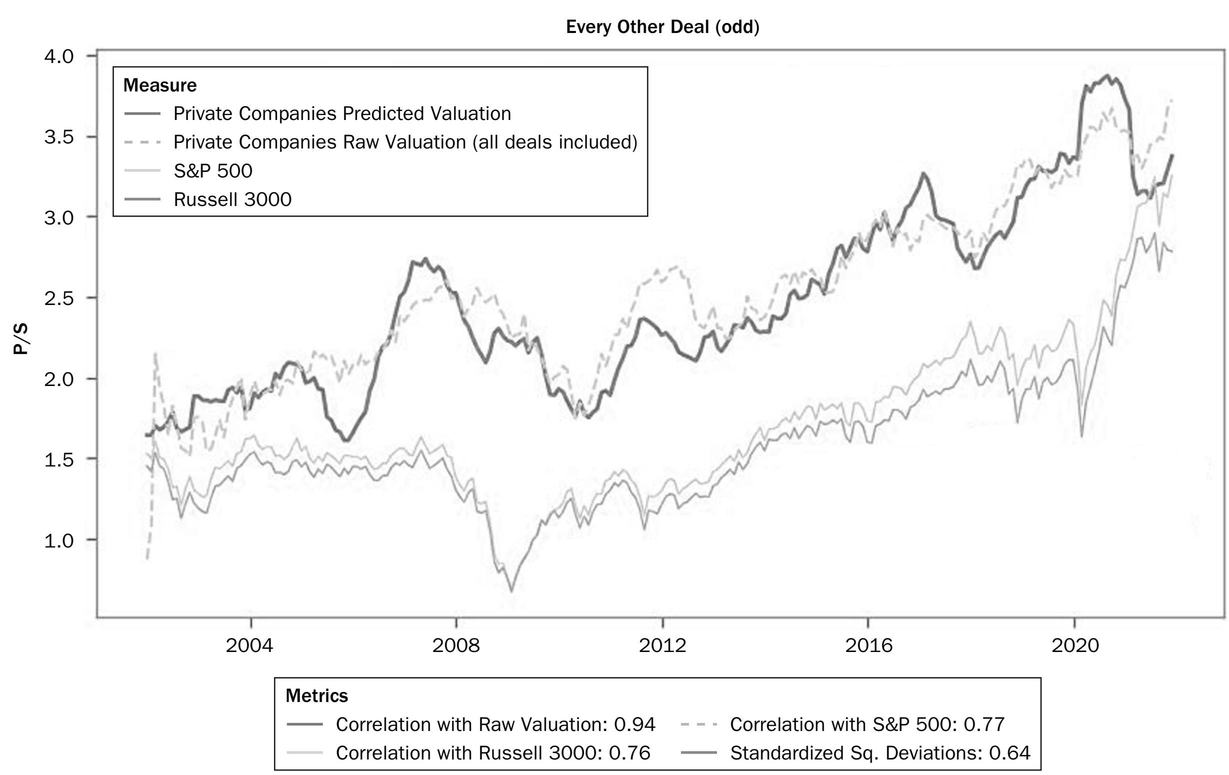
raw valuation. The baseline difference between the predictions on all data and the raw series is 0.19.

In the second panel, we repeat the exercise, but drop every other deal in the odd numbered row after sorting by date. In effect, we train the model on half of the data, imagining that half of the transactions is not observed. We see that the predicted curve has a correlation of 0.94, with the raw valuation series based on all the transactions. Also, the sum of the squared deviations has increased to 0.64.

EXHIBIT 22

Partial Observation and Predictions: 1

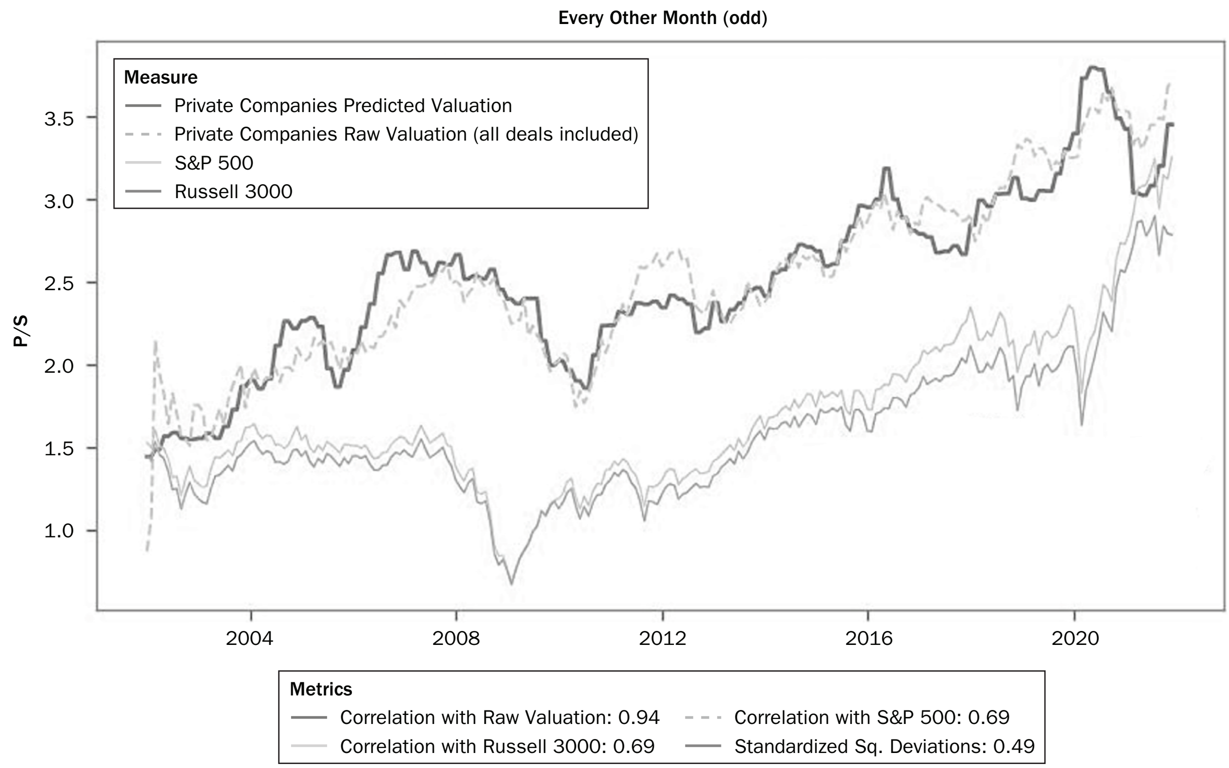




(continued)

EXHIBIT 22 *(continued)*

Partial Observation and Predictions: 1



SOURCE: Based on predictions from the DLM smoothed model of transaction valuations.

The corresponding values when we train on odd calendar months instead are shown in the bottom panel, and we obtain the same 0.94 correlation and deviations of 0.49.

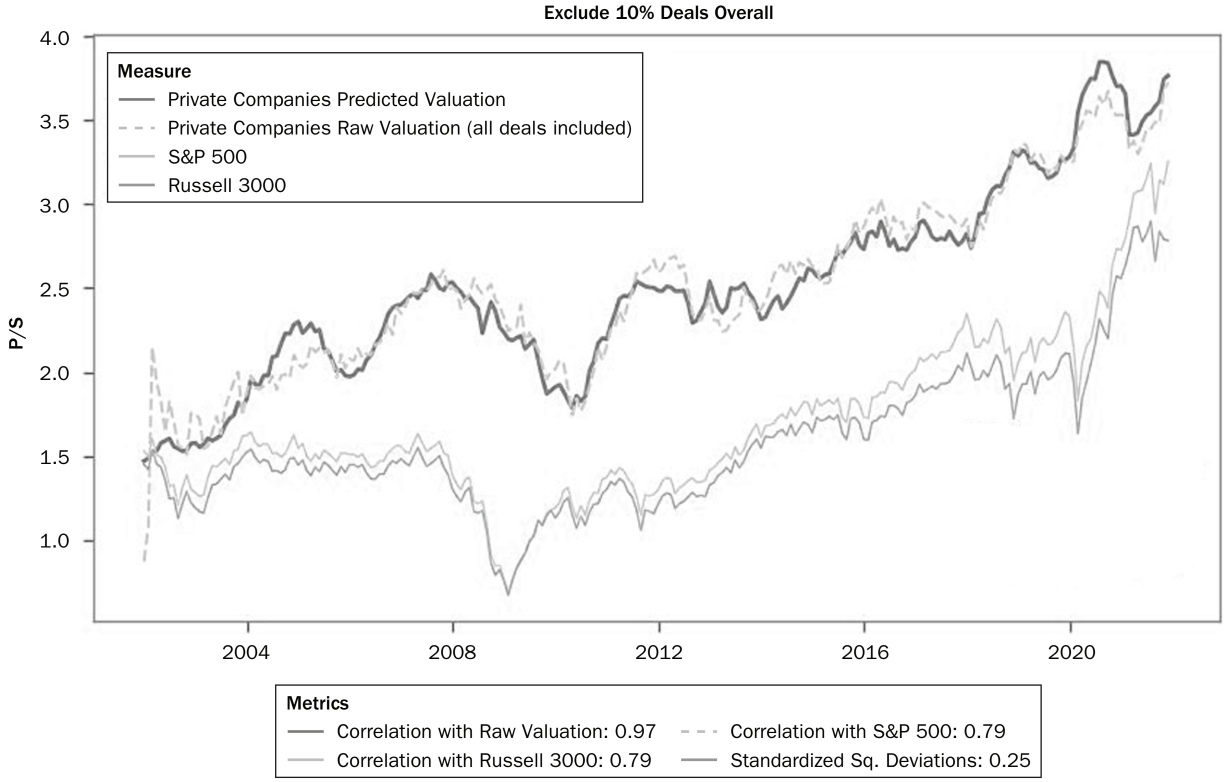
We repeat this exercise several times, assuming only different cuts of transactions are viewed. A sample of these results is shown in Exhibit 23, where we randomly exclude 10%, 25%, and 75% of the transactions, and we can see that the correlation remains high across the three panels, but deviations and patterns start to differ as more transactions are dropped. However, what provides solace is that even when 75% deals are excluded, the errors in the predicted series do not stray too far from the raw series and tend to catch up quickly. That is, observing even one in four transactions is enough to form a roughly precise estimate of valuations in private markets.

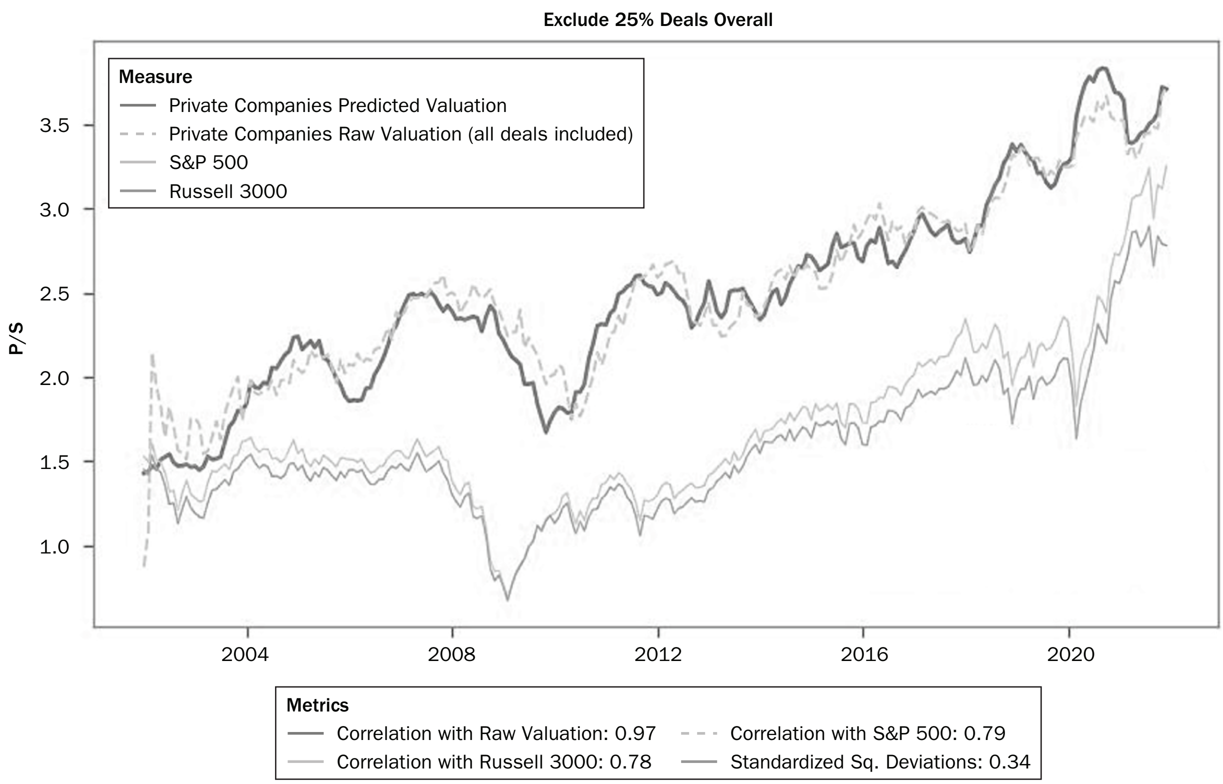
For the sake of brevity, we report the results of all the combinations of observing a partial set of transactions in Exhibit 24. The takeaways are that as fewer and fewer transactions are observed, the trends deviate from the ground truth (i.e., raw valuations) but still are extremely highly correlated, with a lowest correlation of 0.91, when 75% of transactions are excluded at random from the sample. Also, the deviations of the predicted series based on partial data and the full raw valuation series increase with greater loss in transactions used to model.[4]

[4] Note that the standardized deviations vary between 0 and infinite and are not bounded to 1 and hence can only be interpreted in a relative sense between the models.

EXHIBIT 23

Partial Observation and Predictions: 2

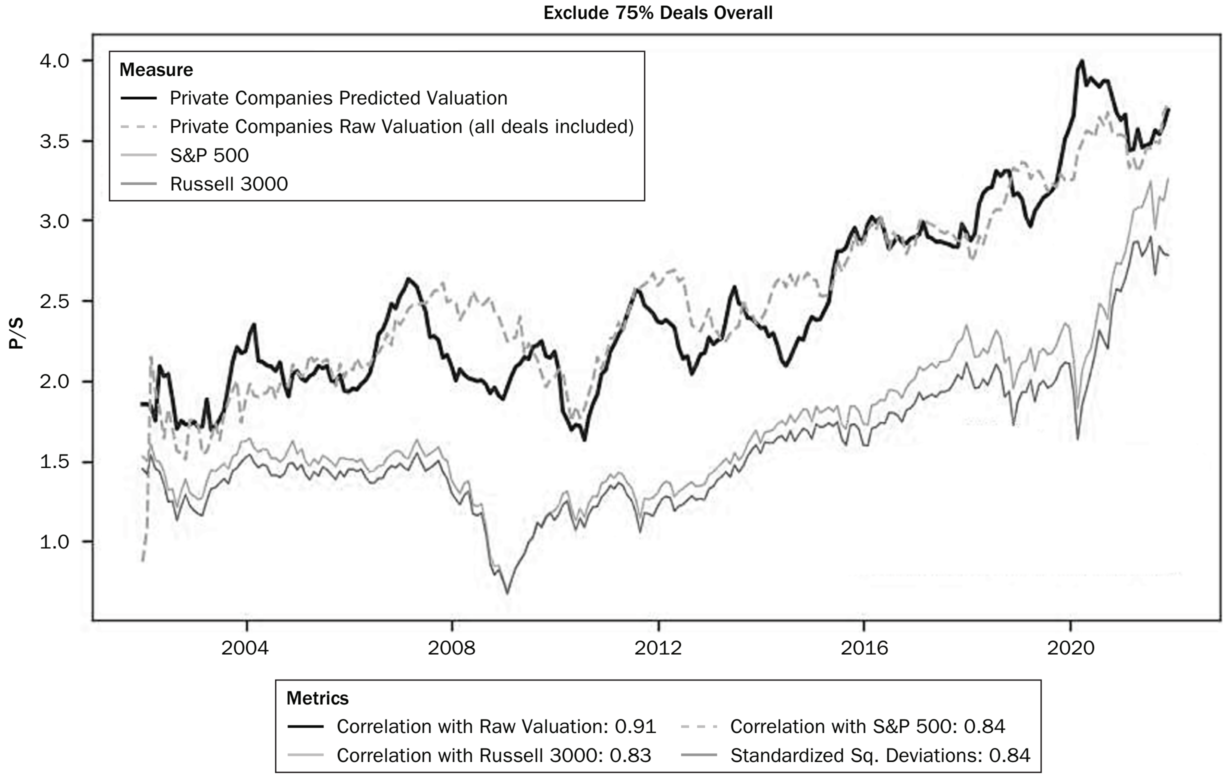




(continued)

EXHIBIT 23 *(continued)*

Partial Observation and Predictions: 2



SOURCE: Based on predictions from the DLM smoothed model of transaction valuations.

EXHIBIT 24

Effect of Not Observing Some Transactions on Predicted Market Dynamics

Training Data	Sample Size	Corr. with Raw Valuation	Corr. with S&P 500	Corr. with Russell 3000	Standardized Deviations
Full Sample	5,438	0.98	0.79	0.78	–
Annual 5% Dropped	5,179	0.98	0.79	0.78	0.20
Overall 5% Dropped	5,167	0.97	0.79	0.79	0.24
Annual 10% Dropped	4,905	0.98	0.80	0.80	0.21
Overall 10% Dropped	4,895	0.97	0.79	0.79	0.25
Annual 25% Dropped	4,090	0.97	0.81	0.80	0.33
Overall 25% Dropped	4,079	0.97	0.79	0.78	0.34
Every Odd Deal	2,719	0.94	0.77	0.76	0.64
Every Even Deal	2,719	0.93	0.75	0.74	0.74
Every Odd Calendar Month	2,691	0.94	0.69	0.69	0.49
Every Even Calendar Month	2,747	0.95	0.84	0.83	0.62
Annual 66% Dropped	1,859	0.95	0.78	0.77	0.56
Overall 66% Dropped	1,849	0.93	0.81	0.81	0.66
Annual 75% Dropped	1,369	0.92	0.76	0.75	0.87
Overall 75% Dropped	1,360	0.91	0.84	0.83	0.84

NOTE: Standardized deviations are computed as the sum of the squared deviations between the predicted valuation curve and raw valuation curve, divided by the square of the raw valuation.

SOURCE: Based on predictions from the DLM smoothed model of different sets of observed transactions.

CONCLUSION

The rapid expansion of private equity as an asset class has been accompanied by persistent opacity in the valuation of private companies, largely driven by the illiquidity of private assets. This lack of transparency has led to increasing skepticism among investors regarding the reliance on smoothed and potentially biased appraised valuations provided by GPs, which often fail to reflect market dynamics accurately.

Despite these challenges, evolving accounting standards and the growing demand for access to private investments necessitate more robust and transparent valuation methodologies. The solution proposed in this article represents a significant shift toward data-driven valuation models. By integrating recent transaction data, relevant company characteristics, and market factors, these models enable more accurate, objective, and timely valuations. Such advancements have the potential to enhance benchmarking, risk management, and portfolio allocation strategies for private market investors.

The factor model presented in this article is calibrated using a large and representative dataset of private company transactions. The proposed factors are grounded in prior academic research, institutional characteristics of private markets, and insights from a survey of GPs. Additionally, the dynamic estimation of factor prices—reflecting the premium or discount investors are willing to pay for specific factor exposures—captures shifts in investor preferences over time. This approach provides realistic risk estimates, even in the presence of biased or sparsely observed data.

Our findings highlight the significant role of key factors in private company valuation, including size, profitability, leverage, labor intensity, and technology. Other control variables such as transaction-specific deal leverage, contemporaneous market and industry valuations, stock market liquidity and volatility, term spreads, and PECCS classes also explain valuation differences. Several of these factors exhibit time-varying effects.

Applications

Our results have a wide range of practical applications:

- **Shadow pricing private markets:** The factor model can be applied to financial information for a large universe of private companies to estimate the market prices of unlisted private companies. Because the factor prices derived from the model are accurate on average and evolve over time, the generated valuations will be precise and robust.
- **Construction of indices:** Shadow pricing unlisted companies, combined with a systematic methodology for selecting private companies, enables the construction of indices and benchmarks at varying levels of granularity. These include global indices, country-specific indices, and PECCS class-level indices. The model's dynamic calibration allows these indices and benchmarks to be repriced frequently and without lag. Indices constructed in this manner provide highly relevant and representative benchmarks for LPs, GPs, and other investors, accurately reflecting the true volatility of private assets.
- **Anchoring valuations:** The computed shadow prices can serve as a foundation for constructing genuine comparables to support investment due diligence and asset valuation. The comparables method involves identifying assets with characteristics similar to the one being evaluated, including market segments and risk profiles. By generating numerous shadow prices, the model ensures accuracy on average, enabling the creation of comparables that closely match

the characteristics and risk profile of the target asset. Furthermore, when idiosyncratic characteristics are time-invariant, this approach provides a dynamic and frequently updated anchor for valuations, offering a precise starting point for assessment.

Among private companies, granularity and robustness of data are often viewed as trade-offs. However, the proposed factor model, when applied to a large universe of nontraded private companies, can deliver highly granular, accurate, and robust segment-level metrics. The production of such metrics, coupled with a novel taxonomy of private companies, enables more timely mark-to-market valuations, mitigates typical biases in private company valuation, and provides a clearer understanding of the true diversification benefits of a portfolio of private funds. Ultimately, this approach equips investors with the tools for better-informed portfolio allocation and monitoring decisions.

APPENDIX

SUBSET SELECTION

Factor selection is performed using subset selection algorithms. Specifically, we follow a forward-stepwise regression (e.g., Hastie, Tibshirani, and Tibshirani 2017), which reduces the number of models to be estimated by selecting the variable that is the most significant in reducing the residual sum of squares the most. This approach is advantageous over a brute-force approach, as it achieves the same results with far fewer estimations.

The forward-stepwise method produces the best models for each number of variables chosen, and from this set of optimal models, we use the Bayesian information criterion (BIC) to pick one best model. BIC is well known to favor models that are parsimonious.[5]

LASSO FEATURE SELECTION

Although forward-stepwise selection allows one to select the best linear model given the potential set of regressors, it is also possible that the regressors are related to valuation in a nonlinear manner. For examining that, one can estimate polynomial specifications of regressor combinations, and feature selection (i.e., narrowing down the list of potential variables) becomes a larger problem.

When second-order combinations of regressors are allowed, then the additional information from such variables to valuation is likely to be lower—that is, low signal-to-noise ratios, or in other words, marginal information from such variables with regard to valuation is lower.

In such scenarios, prior work has found that LASSO feature selection regressions perform well (e.g., Hastie, Tibshirani, and Tibshirani 2017). The LASSO method regularizes model parameters by shrinking the regression coefficients, reducing some of them to zero. The feature (or variable) selection phase occurs after the shrinkage, where every nonzero value is selected to be used in the regression model.

Thus, we use the LASSO method to select second-order terms that meaningfully improve the explanatory power of the model. To prioritize parsimony, we also construct principal components of the selected second-order terms to add the least number of additional variables while preserving the explanatory power of the model.

[5] Specifically, BIC penalizes the model for the number of parameters more than other model selection methods such as the Akaike information criterion or simply selecting based on higher R^2 (e.g., Schwarz 1978).

DYNAMIC LINEAR MODEL

The simple linear model proposed can be expressed as in Equation A-1:

$$Y_{i,t} = \beta_1 + \sum_{k=2}^{K} \beta_k X_{k,i,t} + \varepsilon_t \tag{A-1}$$

where $Y_{i,t}$ corresponds to a valuation ratio of asset i at time t with k characteristics/factors $X_{i,t}$ such as size, leverage, and so on. Here $\beta_{k,t}$ measures the price of each factor and is estimated in an unbiased manner when the error term ε_t is independently and normally distributed. However, in private asset markets neither could be assumed, with the error terms likely to capture systematic valuation errors, such as periods when larger or smaller firms may be preferred more, and hence overvalued.

In other words, investor preferences and market conditions may evolve, which may alter the price of factors over time. This will lead to the error term systematically capturing these trends, leading to nonnormally distributed errors, and hence biased estimates.

Thus, the coefficients or β are allowed to evolve over time to accommodate for time-varying factor prices, as shown in Equation A-2:

$$Y_{i,t} = \beta_{1,t} + \sum_{k=2}^{K} \beta_{k,t} X_{k,i,t} + \varepsilon_t \tag{A-2}$$

The errors above can be assumed to be independent and normally distributed, as the evolution of β can capture time-varying investor preferences for private company characteristics. The β can be modeled as a first-order autocorrelated process as below:

$$\beta_{k,t} = \beta_{k,t-1} + W_{k,t} \tag{A-3}$$

Defining θ_t as a vector of factor prices $\beta_{1,t}$, $\beta_{2,t}$, and so on till $\beta_{K,t}$, we can rewrite Equation A-2 as

$$Y_{i,t} = X_t'\theta_t + \nu_t \tag{A-4}$$

$$\theta_t = G_t\theta_{t-1} + W_t \tag{A-5}$$

where ν_t is the variance or noise of the pricing equation and is independent and identically distributed and W_t is the covariance matrix of the model's coefficients.

Equation A-5 is a state-space Markov chain model between observable valuation and autoregressive vector of factor prices. When G_t is the identity matrix, the regression coefficients can be thought of as being independent random walks—that is, each factor's effect on valuation is serially correlated but the different factors themselves are independent of each other. However, if that is not the case (like when factors jointly evolve and affect value, for example, or say the effect of size and leverage on valuation changes jointly over time), the covariance matrix W_t can capture such dynamics, which are not feasible in the linear model proposed in Equation A-1.

Equation A-5 is estimated using Bayesian techniques that can estimate K coefficients each time a new observation is available, which in this setting, is a new transaction in private companies, thus allowing one to learn and update unbiased estimates of factor prices each time a new biased and noisy transaction is observed.

The summary of the empirical approach is provided in Exhibit A1.

EXHIBIT A1
Econometric Framework

Step	Detail	Purpose
1	Split regressors into required and optional	Designate some regressors as required like size, leverage, market valuation, etc., as these variables inarguably determine private company valuation without requiring statistical justification
2	Forward-stepwise selection	Estimate the list of optional regressors that are key for valuation
3	LASSO feature selection	Examine whether there are nonlinear relationships between regressors and valuation, use LASSO to select key higher order terms, and use principal component analysis to further reduce features
4	DLM	For the model that includes required, selected optional, and principal component second-order variables as regressors, estimate time-varying coefficients

MODELING NONLINEAR RELATIONSHIPS

To model the nonlinear relationships, we experiment with polynomial specifications of our model. Rather than pick and choose variables that may exhibit nonlinear relationships, we experiment with all potential polynomial specifications to check if overall model performance is improved. To maintain the simplicity and intuitiveness of the model, we restrict the experimentation to second-order polynomial terms that include all second-order terms of predictor variables as described below.

To implement this approach, first all combinations of predictor variables in the second degree are created. That is for any two regressors a and b, a list of regressors that include a, b, a^2, b^2, and ab is created. Using this expanded set of regressors, we can implement a forward-stepwise subset selection. Using the best model chosen by the algorithm for each number of optional second-order variables, we realize a large gain in model performance. For example, the adjusted R^2 improves to 0.33 from 0.21 in the OLS setting.

However, a problem with this approach is that the large set of regressors generated by the combination of second-order terms, while improving model performance, can lead to overfitting, that is the model fitting the in-sample data extremely well but performing very poorly in out-of-sample predictions. To overcome this concern, we perform LASSO regressions to select the most predictive second-order terms and also use principal component analysis to reduce the number of predictors.

ACKNOWLEDGMENTS

This work was conducted under the auspices of the EDHEC Infra & Private Assets Research Institute, and the authors are grateful for the institutional support provided by the EDHEC Business School. The authors especially acknowledge the support of Noel Amenc and Frederic Blanc-Brude for their insightful guidance, constructive feedback, and encouragement throughout the research process.

REFERENCES

Amess, K., J. Stiebale, and M. Wright. 2015. "The Impact of Private Equity on Firms' Innovation Activity." Number 184. DICE Discussion paper.

Amihud, Y. 2002. "Illiquidity and Stock Returns: Cross-Section and Time-Series Effects." *Journal of Financial Markets* 5 (1): 31–56.

Armstrong, C. S., J. E. Core, D. J. Taylor, and R. E. Verrecchia. 2011. "When Does Information Asymmetry Affect the Cost of Capital?" *Journal of Accounting Research* 49 (1): 1–40.

Bailey, M. J., R. F. Muth, and H. O. Nourse. 1963. “A Regression Method for Real Estate Price Index Construction.” *Journal of the American Statistical Association* 58 (304): 933–942.

Banz, R. W. 1981. “The Relationship between Return and Market Value of Common Stocks.” *Journal of Financial Economics* 9 (1): 3–18.

Blanc-Brude, F., and C. Tran. 2019. “Which Factors Explain Unlisted Infrastructure Asset Prices?” EDHEC Infra & Private Assets Research Publication.

Chingono, B., and D. Rasmussen. 2015. “Leveraged Small Value Equities.” Available at SSRN 2639647.

Damodaran, A. 2005. “Marketability and Value: Measuring the Illiquidity Discount.” Available at SSRN 841484.

——. 2024. “Enterprise Value Multiples by Sector (US).” NYU Stern School of Business.

Doidge, C., K. M. Kahle, G. A. Karolyi, and R. M. Stulz. 2018. “Eclipse of the Public Corporation or Eclipse of the Public Markets?” *Journal of Applied Corporate Finance* 30 (1): 8–16.

Fama, E. F., and K. R. French. 1992. “The Cross-Section of Expected Stock Returns.” *The Journal of Finance* 47 (2): 427–465.

——. 1993. “Common Risk Factors in the Returns on Stocks and Bonds.” *Journal of Financial Economics* 33 (1): 3–56.

George, T. J., and C.-Y. Hwang. 2010. “A Resolution of the Distress Risk and Leverage Puzzles in the Cross Section of Stock Returns.” *Journal of Financial Economics* 96 (1): 56–79.

Goldstein, S. 2024. “World Market Cap Set to Hit New Record of More Than $123 Trillion.” MarketWatch. https://www.marketwatch.com/story/after-u-s-and-china-panic-world-market-cap-set-to-hit-new-record-says-bank-of-america-aa108ae4.

Gomes, J. F., and L. Schmid. 2010. “Levered Returns.” *The Journal of Finance* 65 (2): 467–494.

Harris, R. S., T. Jenkinson, S. N. Kaplan, and R. Stucke. 2023. “Has Persistence Persisted in Private Equity? Evidence from Buyout and Venture Capital Funds.” *Journal of Corporate Finance* 81: 102361.

Hastie, T., R. Tibshirani, and R. J. Tibshirani. 2017. “Extended Comparisons of Best Subset Selection, Forward Stepwise Selection, and the LASSO.” arXiv preprint, 1707, 08692.

Hodgson, A., and P. Stevenson-Clarke. 2000. “Earnings, Cashflows and Returns: Functional Relations and the Impact of Firm Size.” *Accounting & Finance* 40 (1): 51–74.

Hou, K., C. Xue, and L. Zhang. 2015. “Digesting Anomalies: An Investment Approach.” *Review of Financial Studies* 28 (3): 650–705.

Kaplan, S. N., and A. Schoar. 2005. “Private Equity Performance: Returns, Persistence, and Capital Flows.” *The Journal of Finance* 60 (4): 1791–1823.

Korteweg, A., and S. Nagel. 2024. “Risk-Adjusted Returns of Private Equity Funds: A New Approach.” *The Review of Financial Studies*, 67.

Malpezzi, S. 2003. “Hedonic Pricing Models: A Selective and Applied Review.” In *Housing Economics and Public Policy*, T. O’Sullivan, and K. Gibb, Eds., Oxford: Blackwell Publishing. https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470690680.ch5.

McKinsey. 2024. “McKinsey Global Private Markets Review 2024.” https://www.mckinsey.com/industries/private-capital/our-insights/global-private-markets-report-2024.

Morgenson, G. 2021. Working for Companies Owned by Well-Heeled Private-Equity Firms Can Mean Lower Wages for Employees. NBC News.

Novy-Marx, R. 2013. "The Other Side of Value: The Gross Profitability Premium." *Journal of Financial Economics* 108 (1): 1–28.

PECCS. 2023. "Private Equity Company Classification Standards (PECCS)." EDHEC Infra & Private Assets Research Publication.

Reclaim Finance. 2022. "Private Equity: Fuelling Climate Crisis from the Shadows." https://reclaim-finance.org/site/en/2022/09/16/private-equity-fuelling-climate-crisis-from-the-shadows/.

Rudin, A., and D. Farley. 2022. "Public and Private Equity Returns: Different or Same?" *The Journal of Portfolio Management* 48 (3): 117–127.

Rudin, A., J. Mao, N. R. Zhang, and A.-M. Fink. 2019. "Fitting Private Equity into the Total Portfolio Framework." *The Journal of Portfolio Management* 46 (1): 60–73.

Schwarz, G. 1978. "Estimating the Dimension of a Model." *Annals of Statistics* 6 (2): 461–464.

Shiller, R. J. 1981. "Do Stock Prices Move Too Much to Be Justified by Subsequent Changes in Dividends?" *American Economic Review* 71 (3): 421–436.

Shleifer, A., and R. W. Vishny. 1997. "The Limits of Arbitrage." *The Journal of Finance* 52 (1): 35–55.

Zhang, L. 2005. "The Value Premium." *The Journal of Finance* 60 (1): 67–103.

Disclosure

The authors declare no conflicts of interest with respect to the research, authorship, and/or publication of this article.