APPLIED FINANCE LETTERS

VOLUME 12, 2023

MEASURING VALUATION UNCERTAINTY: A PCA APPROACH

JOHN GARCIA*[1]

1. California Lutheran University, USA

* Corresponding Author: John Garcia, Assistant Professor of Finance and Analytics, California Lutheran University, School of Management, 60 W. Olsen Rd, Thousand Oaks, CA, 91360. (+00 (805) 493 3814
* jgarcia@callutheran.edu

Abstract

Determining which companies are more difficult to value is a topic of significant interest in finance. While prior studies have employed various univariate proxies to classify firms into high- and low-valuation uncertainty groups, this study proposes a new approach to measuring valuation uncertainty. Specifically, I employ principal component analysis (PCA) to extract the first principal component from 11 valuation uncertainty proxies. The first principal component is proposed as a comprehensive measure of a firm's valuation uncertainty. The findings demonstrate that the PCA-derived valuation uncertainty index provides two key benefits over univariate valuation uncertainty proxies. First, integrating multiple valuation uncertainty proxies into a single metric improves our ability to quantify valuation uncertainty. Second, it assists in identifying the proxies that are most informative in measuring a firm's valuation uncertainty. Ultimately, the PCA-derived valuation uncertainty index can better enable market participants to measure a firm's valuation uncertainty.

Keywords: Valuation uncertainty, hard-to-value, principal component analysis

JEL Codes: G10, G12

1. Introduction

A firm's intrinsic value is calculated by discounting its expected future cash flow at the cost of capital. However, what makes a firm difficult to value? Over the years, finance academics and practitioners have explored this question and proposed many measures as proxies for the unmeasurable latent factor of valuation uncertainty. These proxies attempt to capture the uncertainty in estimating cash flows and discount rates, both of which affect the uncertainty in estimating a firm's value. Understanding a firm's valuation uncertainty has garnered increased interest in recent years, as we have observed dramatic price dislocations between a firm's intrinsic value and its market value for many hard-to-value firms.

Valuation uncertainty is a crucial factor that influences market participants' decision-making process in several ways. For instance, the uncertainty surrounding a firm's value can lead to divergent viewpoints among investors, which, in turn, negatively impacts stock prices and makes them more volatile in response to shifts in investor sentiment (Baker & Wurgler, 2006,2007). This effect is more pronounced during market dislocations, such as during the early stages of the COVID-19 pandemic, which exposed hard-to-value stocks to significant overreactions (Xiong et al., 2020).

Valuation uncertainty also plays a vital role in risk management because firms with higher valuation uncertainty have stock values that are more sensitive to changes in market conditions or economic shocks. This increased valuation complexity increases investors' propensity to use valuation heuristics, which are subject to more potent adverse effects stemming from behavioural biases, leading to a higher likelihood of valuation mistakes (Baker & Wurgler, 2007; Kumar, 2009). Further, valuation uncertainty negatively affects market makers' bid–ask spread decisions, negatively affecting a firm's liquidity (Glosten & Milgrom, 1985). Therefore, valuation uncertainty is a fundamental concept that affects market participants' decision-making, and it is important for investors to understand the factors that contribute to valuation uncertainty and consider them in their decision-making.

Valuation uncertainty has been extensively studied in the finance literature. However, previous research has predominantly relied on many indirect and arbitrary proxies to measure a firm's valuation uncertainty. This has revealed that understanding the drivers of a firm's valuation uncertainty encompasses multiple dimensions, and a single univariate measure cannot comprehensively capture it. Moreover, relying on indirect proxies in empirical models may cause significant issues as they may introduce measures that are highly correlated with irrelevant attributes, which can create noise or bias, compromising the accuracy of the results. This lack of a direct and reliable measure of valuation uncertainty can limit our ability to gain insights from models that may be affected by impure valuation uncertainty measures. Therefore, developing a purified holistic measure of valuation uncertainty is crucial to address these issues.

Principal Component Analysis (PCA) is a statistical technique that provides an effective way of constructing a comprehensive measure of valuation uncertainty. PCA achieves this by identifying and simplifying the common information content among variables, reducing the complexity of high-dimensional correlated data, and projecting them onto a smaller set of new variables. By extracting the shared information from multiple variables, we can identify uncorrelated latent structures from the data, providing a purified and holistic measure of valuation uncertainty.

This study utilizes PCA to extract the first principal component from 11 widely used valuation uncertainty proxies, which reveals the underlying valuation uncertainty latent component. This approach results in a more precise and informative valuation uncertainty measure. This study contributes to the valuation uncertainty literature by presenting a more robust and informative measure of the valuation uncertainty latent construct. I argue that PCA offers two benefits. First, combining a wide range of valuation uncertainty proxies into one variable helps us better measure a firm's valuation uncertainty. Second, it can help determine the more meaningful proxies in measuring a firm's valuation uncertainty.

Previous research has identified several factors affecting a firm's valuation uncertainty, including Baker and Wurgler (2006, 2007), who demonstrate that younger companies with small market capitalizations, low profits, high growth, high financial distress, and low dividends are difficult to value. The rationale is that this type of firm makes estimating their future cash flows and discount rates more challenging. Other studies have presented similar proxies for valuation uncertainty as those used by Baker and Wurgler, such as small firms with high return volatility, low profitability, and high growth rates (Aboody et al., 2018; Hribar & McInnis, 2012). In addition, analysts' forecast dispersion is informative of valuation uncertainty (Güntay & Hackbarth, 2010), and dividend-paying firms tend to be easier to value (Pastor & Veronesi, 2016). Furthermore, the number of analysts covering a firm can be a proxy for its information environment, whereby higher analyst coverage reduces information asymmetry, thereby reducing valuation uncertainty (Ramnath et al., 2008).

Although prior studies have utilized univariate or bivariate measures to classify firms into valuation uncertainty groups, such approaches provide only a limited picture of the multifaceted concept of valuation uncertainty (e.g., Aboody et al., 2018; Baker & Wurgler, 2006,2007; Hribar & McInnis, 2012; Kumar, 2009). Additionally, the univariate valuation uncertainty measures employed in prior studies can be ad hoc and inadequate for fully capturing the elusive latent factor of valuation uncertainty. A comprehensive metric is necessary to capture the information content of various valuation uncertainty proxies found in the literature. The PCA approach discussed in this study provides such a metric by extracting shared information content across multiple proxies and identifying the uncorrelated latent structure, leading to a more comprehensive and dependable measure of valuation uncertainty. By moving beyond simplistic univariate or bivariate measures, we can obtain a deeper understanding of the complex nature of valuation uncertainty.

This paper is organized as follows. In section 2, I review the data and explain how PCA is applied to the valuation uncertainty proxies to extract the latent component. I present and discuss the PCA results in section 3 and offer some concluding remarks in section 4.

2. Data and methodology

2.1 Data

The initial step was to collect the proxies for a firm's valuation uncertainty. Specifically, I evaluate the firm-level valuation uncertainty proxies used in recent research and identify 11 proxies that are commonly used in the literature. The sample consists of U.S. firm-level data collected from Bloomberg and contains 1,062 publicly traded firms with observations for the fiscal period ending December 31, 2020. Table 1 presents summary statistics for the valuation uncertainty proxies used in this study[1].

Table 1: Descriptive Statistics

	Min.	Max.	Median	Mean	Std. Dev.
Analysts' EPS Dispersion	0.01	3.16	0.12	0.24	0.45
260-Day Share Price Volatility	24.72	190.54	57.89	60.99	18.94
EPS / Share Price	-0.55	0.24	0.04	0.03	0.09
EPS Volatility / Share Price	0.0005	0.2425	0.0078	0.0185	0.0334
EPS YOY Change (%)	-14.45	18.46	-0.01	0.05	3.38
Cash Flow Volatility / Share Price	0.0009	0.3087	0.0143	0.0269	0.0434
Annual Dividend Yield (%)	0	7.37	1.11	1.57	1.67
Total Analysts	1	57	10	12.52	8.59
Bloomberg 1-Year Default Prob.	0	0.164	0.002	0.0051	0.0114
Beta	0.14	2.61	1.1	1.11	0.28
Log (Market Capitalization)	3.99	14.62	8.46	8.6	1.65
Number of Firms	1,062				

Note: Table 1 reports descriptive statistics on the firm-level valuation uncertainty proxies. All data are as of December 31, 2020. A complete list and description of the variables used in the study are found in Appendix 1.

The Pearson correlations between the study variables are presented in Table 2. Several measures have moderate (0.30 to 0.49) to strong (0.50+) correlations, suggesting that they may be measuring the same latent factor (the valuation uncertainty of a firm). Consequently, isolating the common latent component measured by the 11 valuation uncertainty proxies could provide a more accurate measure of valuation uncertainty than any single proxy.

[1] To mitigate the effect of extreme outliers, analysts' EPS dispersion, EPS/share price, and EPS YOY change were winsorized at the 1% and 99% level, and the annual dividend yield, EPS volatility/share price, and cash flow volatility/share price were winsorized at the 99% level.

Table 2: Correlation Matrix

		Analysts' EPS Dispersion	1	2	3	4	5	6	7	8	9
1	**260-Day Share Price Volatility**	0.22***									
2	**EPS / Share Price**	-0.12***	-0.25***								
3	**EPS Volatility / Share Price**	0.26***	0.53***	-0.51***							
4	**EPS YOY Change (%)**	-0.15***	-0.08***	0.24***	-0.17***						
5	**Cash Flow Volatility / Share Price**	0.12***	0.40***	-0.05	0.52***	-0.05*					
6	**Annual Dividend Yield (%)**	-0.07**	-0.07**	0.10***	0.02	-0.04	0.06*				
7	**Total Analysts**	-0.05	-0.24***	-0.08***	-0.05	-0.03	-0.15***	-0.01			
8	**Bloomberg 1-Yr. Default Prob.**	0.20***	0.64***	-0.36***	0.68***	-0.17***	0.46***	0.04	-0.07**		
9	**Beta**	0.03	0.69***	-0.16***	0.27***	-0.03	0.26***	0.01	-0.07**	0.40***	
10	**Log (Market Capitalization)**	-0.21***	-0.49***	0.04	-0.26***	0.02	-0.26***	0.02	0.80***	-0.29***	-0.19***

Note: Table 2 presents the correlation matrix for the variables used in the study. The numbers are the Pearson correlations. A complete list and description of the variables used in the study are found in Appendix 1. ***, **, and * denote significance at the .01, .05, and .10 levels, respectively.

2.2 Methodology

Using the 11 valuation uncertainty proxies, principal component analysis was applied to extract the latent structures in the valuation uncertainty proxies. PCA is a technique used to find patterns in data to decrease redundancy in univariate analysis when collinear data are employed (Adbi & Williams, 2010). PCA restructures the datasets with correlated variables into uncorrelated components of the original variables. In addition, new values are computed for each orthogonal component, which can be utilized to replace the original correlated variables and be used as a more holistic valuation uncertainty index.

Eigen decomposition is the standard method for conducting PCA. Decomposition is conducted on an $n \times n$ matrix representing the relationship between all pairs of n variables (in this case, the valuation uncertainty measures). Then, depending on the nature of the variable, either the covariance or correlation matrix is input into the eigen decomposition (Adbi & Williams, 2010). Because the valuation uncertainty measures used in the study are on various scales, a correlation matrix is utilized. Furthermore, the valuation uncertainty measures are normalized by scaling each measure before performing the eigen decomposition. Each derived eigenvector is treated as a principal component; that is, the numeric values in each eigenvector are the coefficients of each principal component, which can be thought of as the valuation uncertainty proxy weights.

A varimax PCA rotation is employed to better interpret the component loadings[2] (i.e., how closely the valuation uncertainty proxies relate to the principal component). Varimax rotation is an orthogonal rotation that maximizes the squared variance of the component loadings (Adbi & Williams, 2010). Therefore, it prioritizes either very high or extremely low loadings, making it simpler to understand the latent structure represented by each component.

[2] The PCA analysis was re-run using an oblique rotation and the results were unchanged.

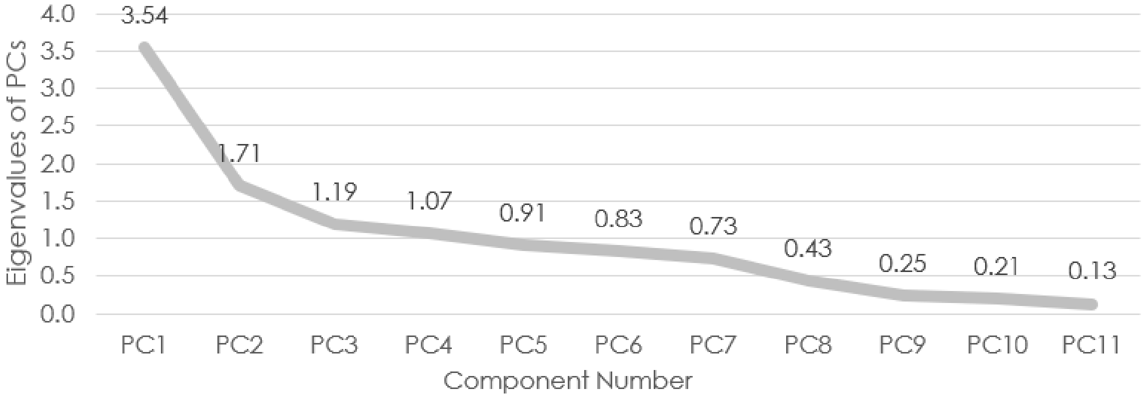
In addition, regression is used to integrate the measurements and derive the latent components[3]. The component scores explain the linear combination of measurements to generate each component. This regression method estimates the projection of each data point along each PCA component as a function of the original n measures for each PCA component. Finally, the calculated regression coefficients for a specific component are referred to as the component loadings and offer a way to linearly combine the *n* measurements to generate the latent component. By analyzing component loadings, we can determine which valuation uncertainty proxies have a more potent effect on the valuation uncertainty latent component.

Finally, for PCA to be reliable, three key assumptions must be met:1) sphericity or the existence of an identity matrix, 2) sampling adequacy or a sufficient number of observations related to the number of variables being evaluated, and 3) confirmation that there is a positive determinant of the correlation matrix. First, using the Bartlett test, I confirmed that the matrix is derived from a collinear population ($X^2 = 415.45$, $p<.01$). Second, Jung (2013) argues that 50 observations per measure are sufficient, and the minimum number required declines as the number of retained factors decrease. Given that there are over 98 observations per measure in this study (1,062 firms divided by 11 valuation uncertainty proxies) and that only one latent component is retained as a measure of a firm's valuation uncertainty, the number of observations is deemed to be sufficient to continue. Finally, a positive determinant of the correlation matrix is computed. Thus, all PCA assumptions were met. The PCA analysis is reported in the next section.

3. Empirical Results

Principal component analysis was performed[4]; Figure 1 presents the scree plot of the eigenvalues for the 11 latent components extracted using PCA. Principal component 1 (PC1) explains 32.2% (3.54/11) of the total variance in the valuation uncertainty proxies, whereas principal component 2 explains only 15.5% (1.71/11), suggesting that only the first principal component should be retained. Further, using Velicer's (1976) minimum absolute partial correlation (MAP) criterion, it is determined that one factor achieves the MAP of 0.05. Thus, only one principal component is retained, serving as the latent measure of a firm's valuation uncertainty.

Figure 1: Scree Plot



Note: The figure presents the eigenvalues of the principal components extracted using PCA.

---

[3] See DiStephano et al. (2008) for an overview of the different alternatives.

[4] The R package "psych" version 2.1.9 was used for the PCA analysis.

Table 3, column 1 presents the component loadings (i.e., correlation values in the range [-1,+1]) between the valuation uncertainty proxies and the first latent component. The measures have been sorted from the highest positive loadings on top to the lowest negative loadings at the bottom. Further, the root mean square of the off-diagonal residuals is 0.14 with a X^2= 2,287.47 and p<.01, confirming that 1 component is sufficient.

Consistent with the literature, the valuation uncertainty proxies that positively relate to the valuation uncertainty latent measure (PC1) are as follows, from greatest to least effect: 260-day share price volatility, one-year default probability, EPS volatility, cash flow volatility, beta, and analysts' EPS dispersion. On the other hand, the valuation uncertainty proxies negatively related to the valuation uncertainty latent measure (PC1) include, from greatest to least effect, market capitalization, profitability (EPS/share price), the number of analysts covering the company, the EPS year-over-year change, and the annual dividend yield. Consequently, PC1 effectively isolates the common component among the 11 valuation uncertainty proxies, thus more holistically measuring the valuation uncertainty latent structure and thereby supporting its use as a comprehensive index of a firm's valuation uncertainty.

In addition, the component loadings for each valuation uncertainty proxy on the valuation uncertainty index (PC1) can be used to identify the most influential contributors to a company's valuation uncertainty. Table 3, column 2 displays the relative significance of each measure of valuation uncertainty (defined as the relative magnitude of the PC1 loading's absolute value). According to the PCA results, the three most influential proxies for a company's valuation uncertainty are the 260-Day share price volatility, the Bloomberg one-year default probability, and the EPS volatility; conversely, the three least influential proxies are the annual dividend yield, the EPS year-over-year change, and the analysts' EPS dispersion.

Table 3: PCA Results

Valuation Uncertainty Proxy	(1) Principal Component 1	(2) Relative Importance
260-Day Share Price Volatility	0.853	1
Bloomberg 1-Year Default Prob.	0.806	2
EPS Volatility / Share Price	0.785	3
Cash Flow Volatility / Share Price	0.612	4
Beta	0.592	5
Analysts' EPS Dispersion	0.338	9
Annual Dividend Yield (%)	-0.015	11
EPS YOY Change (%)	-0.219	10
Total Analysts	-0.347	8
EPS / Share Price	-0.437	7
Log (Market Capitalization)	-0.590	6

Note: Column 1 presents the loadings for each valuation uncertainty proxy to principal component 1 (PC1). PC1 is then computed for each firm as the sum weight of the PC1 component loading x the observed valuation uncertainty proxy measure. Column 2 presents the relative importance of each valuation uncertainty proxy in measuring the PC1 latent component. A complete list and description of the variables used in the study are found in Appendix 1.

Next, using the valuation uncertainty index (PC1), firms are sorted into terciles. Firms with a valuation uncertainty index of -0.407 or less are classified as “low valuation uncertainty,” firms with a valuation uncertainty index above -0.407 and up to 0.128 are classified as “average valuation uncertainty,” and firms with a valuation uncertainty index greater than 0.128 are classified as having “high valuation uncertainty.”

Table 4 presents the mean, median, and standard deviation of the 11 valuation uncertainty proxies by the valuation uncertainty index (PC1) tercile group. The results are consistent with prior literature, revealing that firms in the low valuation uncertainty group have the lowest analysts’ EPS dispersion, 260-day share price volatility, EPS volatility, cash flow volatility, one-year default probability, and beta. Additionally, these firms have the highest levels of profitability (EPS/share price), EPS year-over-year change, number of analysts covering the firm, and market capitalization, suggesting that they are relatively easier to value. Conversely, firms in the high valuation uncertainty group exhibit a reversed relationship, indicating that they are more difficult to value.

The final column of Table 4 reports the results of the independent t-test evaluating the difference in the means for the high versus low valuation uncertainty groups for each respective valuation uncertainty proxy. The p-values for all t-tests, except for annual dividend yield, are below 0.01, indicating that the difference between the high and low valuation groups is statistically significant. However, there is no consistent relationship between the annual dividend yield and the valuation uncertainty groups, and the corresponding t-test indicates that the difference is not statistically significant. Table 3 shows that the annual dividend yield has the lowest loading to the valuation uncertainty index of all the proxies (0.015), indicating that the measure is the least useful proxy for identifying hard-to-value firms.

Table 4: Descriptive Statistics by PCA Terciles

	PC1 Tercile #1 Low Uncertainty (n=354)			*PC1 Tercile #2 Avg. Uncertainty (n=354)*			*PC1 Tercile #3 High Uncertainty (n=354)*			*High-Low Groups*	
	Median	Mean	Std. Dev.	Median	Mean	Std. Dev.	Median	Mean	Std. Dev.	Mean Difference	*t*-statistic
Analysts’ EPS Dispersion	0.08	0.11	0.12	0.12	0.18	0.21	0.19	0.44	0.7	0.32	8.64***
260-Day Share Price Volatility	45.3	45.38	7.57	58.07	58.54	7.61	75.58	79.04	19.64	33.66	30.10***
EPS / Share Price	0.03	0.04	0.03	0.05	0.04	0.04	0.03	0	0.15	-0.04	-4.26***
EPS Volatility / Share Price	0.0042	0.0063	0.0058	0.0071	0.0105	0.0108	0.0195	0.0387	0.0508	0.03	11.92***
EPS YOY Change (%)	0.06	0.43	2.76	0.03	0.41	3.01	-0.26	-0.67	4.1	-1.09	-4.15***
Cash Flow Volatility / Share Price	0.01	0.01	0.01	0.01	0.02	0.02	0.03	0.05	0.06	0.04	12.31***
Annual Dividend Yield (%)	1.07	1.4	1.39	1.38	1.74	1.72	0.8	1.56	1.86	0.15	1.24
Total Analysts	18	18.73	8.98	8	9.91	6.24	7	8.91	6.59	-9.83	-16.60***
Bloomberg 1-Year Default Prob.	0.0005	0.0009	0.0013	0.002	0.0026	0.0026	0.007	0.0119	0.0177	0.01	11.67***
Beta	0.89	0.91	0.2	1.12	1.12	0.19	1.3	1.3	0.29	0.39	20.92***
Log (Market Capitalization)	9.99	10.08	1.31	8.29	8.29	1.11	7.44	7.43	1.24	-2.65	-27.70***
Valuation Uncertainty Index (PC1)	-0.78	-0.82	0.28	-0.14	-0.14	0.15	0.56	0.96	1.13	1.78	28.69***

Note: Table 4 displays the median, mean, and standard deviation for the valuation uncertainty proxy measure (PC1) used in the study and the valuation uncertainty proxies used in the study by PC1 terciles, where the first tercile is designated the low valuation uncertainty group, the second tercile is the average valuation uncertainty group, and the third tercile is the high valuation uncertainty group. The last column presents the independent t-test statistic comparing the means for the high-uncertainty group to the low-uncertainty group. *** denotes significance at the 0.01 level.

As a robustness check, I compare the valuation uncertainty groups (derived using the PCA method described in this section) to the clusters formed using the partition around medoids (PAM) clustering algorithm, where the Manhattan distance is used to measure similarity. PAM provides a more robust version of the k-means algorithm: while *k*-means clustering aims to minimize intra-cluster distance, *k*-medoid minimizes the dissimilarities between points in a cluster and points considered to be the centres of that cluster, thereby producing a more robust result (Kaufman & Rosseeuw, 1990). Using the PAM algorithm and 11 valuation uncertainty proxies, firms are clustered into three groups, with the first cluster containing 383 firms, the second containing 371 firms, and the third containing 308 firms.

Table 5 provides a detailed analysis of the 11 valuation uncertainty measures for the PAM-derived cluster groups. The table presents the mean, median, and standard deviation for each measure, while the last column displays the results of the independent *t*-test. The *t*-test compares the difference in the means for the high versus low-valuation uncertainty groups for each valuation uncertainty proxy. The summary statistics in Table 5 reveal that the firms in Cluster 1 exhibit low valuation uncertainty, cluster 2 contains firms with average valuation uncertainty, and Cluster 3 contains firms with high valuation uncertainty. Notably, seven valuation uncertainty proxies exhibit a relationship consistent with their uncertainty classification (low, average, and high), including the five proxies with the highest loadings to PC1.

It is worth noting that the remaining valuation uncertainty proxies do not always exhibit a relationship consistent with their valuation uncertainty designation. This inconsistency can be attributed to the fact that these proxies represent the least informative measures for constructing the valuation uncertainty index (PC1), and are, therefore, less informative in driving the construction of the clusters. Consequently, the application of the PAM algorithm to these measures can produce mixed results. Nevertheless, the *t*-test evaluating the difference in the means for the high versus low valuation uncertainty groups indicates that the difference in means for all the valuation uncertainty proxies, except for the EPS year-over-year change measure, is statistically significant at the .01 level.

Table 5: Descriptive Statistics by PAM Clusters

	Cluster #1 *Low Uncertainty (n=383)*			*Cluster #2* *Avg. Uncertainty (n=371)*			*Cluster #3* *High Uncertainty (n=308)*			*High-Low Groups*	
	Median	Mean	Std. Dev.	Median	Mean	Std. Dev.	Median	Mean	Std. Dev.	*Mean Difference*	*t*-statistic
Analysts' EPS Dispersion	0.09	0.15	0.21	0.11	0.24	0.48	0.18	0.35	0.57	0.2	6.00***
260-Day Share Price Volatility	45.97	46.89	8.68	57.34	58.09	9.06	76.72	80.76	19.56	33.87	28.24***
EPS / Share Price	0.03	0.03	0.03	0.05	0.05	0.07	0.03	0	0.15	-0.03	-3.59***
EPS Volatility / Share Price	0.0045	0.0088	0.0123	0.0077	0.0147	0.0227	0.0144	0.0343	0.0507	0.03	8.62***
EPS YOY Change (%)	0	0.14	2.7	0	-0.05	2.54	-0.09	0.08	4.7	-0.05	-0.17
Cash Flow Volatility / Share Price	0.01	0.01	0.02	0.02	0.03	0.03	0.02	0.04	0.06	0.03	8.10***
Annual Dividend Yield (%)	1	1.31	1.41	2.51	2.44	1.72	0	0.83	1.41	-0.48	-4.46***
Total Analysts	19	20.06	7.97	6	7.14	4.26	8	10.11	6.74	-9.95	-17.77***
Bloomberg 1-Year Default Prob.	0.0005	0.0013	0.0022	0.0022	0.0037	0.0077	0.0064	0.0113	0.0174	0.01	10.05***
Beta	0.92	0.95	0.22	1.07	1.06	0.18	1.35	1.35	0.28	0.4	20.55***
Log (Market Capitalization)	10.08	10.21	1.2	7.76	7.73	1.07	7.84	7.76	1.21	-2.44	-26.39***

Note: Table 5 displays the median, mean, and standard deviation for the valuation uncertainty proxies used in the study by the PAM clusters compiled from the 11 valuation uncertainty proxies. The last column presents the independent t-test statistic comparing the means for the high-uncertainty group to the low-uncertainty group. *** denotes significance at the 0.01 level.

To further assess the efficacy of the PCA-based valuation uncertainty index and the PAM clustering algorithm, I performed a comparison of the valuation group designation resulting from the two methods. I find that 73.26% of the firms are classified consistently as low, average, or high valuation uncertainty between the two methods. Additionally, the correlation between their respective classification groupings (1, 2, 3) is 0.78, indicating a high level of agreement between the two methods. These results suggest that the valuation uncertainty index constructed using PCA is a robust measure of a firm's valuation uncertainty, as it generates similar valuation group designations as the PAM clustering algorithm.

4. Conclusion

Understanding a firm's valuation uncertainty is essential for investors, analysts, and other market participants in making informed decisions. However, prior studies have relied on several univariate proxies to measure a firm's valuation uncertainty, providing only a partial picture of this complex, multifaceted, and elusive latent factor. To overcome this limitation, this study employs PCA to create a purified and comprehensive measure of valuation uncertainty that captures the information content of 11 proxies.

The findings demonstrate that the valuation uncertainty index derived from these 11 proxies accurately captures a firm's valuation uncertainty and aligns with the results of the PAM clustering algorithm. PCA provides two significant advantages. Firstly, it permits the merging of various proxies for valuation uncertainty into one inclusive metric, thus enhancing the ability to measure it more comprehensively. Secondly, it enables the identification of the most informative proxies for determining the factors that drive a firm's valuation uncertainty.

The findings have significant implications for market participants seeking to understand the impact of valuation uncertainty on financial markets. By offering a comprehensive measure of a firm's valuation uncertainty, this study contributes to an improved understanding of how to measure valuation uncertainty efficiently and holistically. Ultimately, the PCA-derived valuation uncertainty index can assist investors, analysts, and other market participants in better measuring a firm's valuation uncertainty.

References

Abdi, H., & Williams, L. J. (2010). Principal component analysis. Wiley interdisciplinary reviews: computational statistics, 2(4), 433-459.

Aboody, D., Even-Tov, O., Lehavy, R., & Trueman, B. (2018). Overnight Returns and Firm-Specific Investor Sentiment. *Journal of Financial and Quantitative Analysis*, *53*(2), 485–505. https://doi.org/10.1017/S0022109017000989

Baker, M., & Wurgler, J. (2006). Investor Sentiment and the Cross-Section of Stock Returns. *The Journal of Finance*, *61*(4), 1645–1680. https://doi.org/10.1111/j.1540-6261.2006.00885.x

Baker, M., & Wurgler, J. (2007). Investor Sentiment in the Stock Market. *Journal of Economic Perspectives*, *21*(2), 129–151.

DiStefano, C., Zhu, M., & Mindrila, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research, and Evaluation*, *14*(1), 20.

Glosten, L., & Milgrom, P. R. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, *14*, 71–100.

Güntay, L., & Hackbarth, D. (2010). Corporate bond credit spreads and forecast dispersion. *Journal of Banking and Finance*, *34*(10), 2328–2345. https://doi.org/10.1016/j.jbankfin.2010.02.019

Hribar, P., & McInnis, J. (2012). Investor Sentiment and Analysts' Earnings Forecast Errors. *Management Science*, Vol. 58, pp. 293–307.

Jung, S. (2013). Exploratory factor analysis with small sample sizes: A comparison of three approaches. *Behavioural processes*, *97*, 90-95.

Kaufman, L., & Rousseeuw, P. J. (1990). Partitioning around medoids (program pam). *Finding groups in data: an introduction to cluster analysis*, *344*, 68-125.

Kumar, A. (2009). Hard-to-Value Stocks, Behavioral Biases, and Informed Trading. *Journal of Financial and Quantitative Analysis*, *44*(6), 1375–1401. https://doi.org/10.1017/S0022109009990342

Pastor, L., & Veronesi, P. (2016). Uncertainty and Valuations: A Comment. *Critical Finance Review*, *5*(1), 129–134. https://doi.org/10.1561/104.00000022

Ramnath, S., Rock, S., & Shane, P. B. (2008). Financial analysts' forecasts and stock recommendations: A review of the research. Foundations and Trends in Finance, 2(4), 311–421.

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, *41*(3), 321-327.

Xiong, H., Wu, Z., Hou, F., & Zhang, J. (2020). Which firm-specific characteristics affect the market reaction of Chinese listed companies to the COVID-19 pandemic? *Emerging Markets Finance and Trade*, *56*(10), 2231-2242.

Appendix 1: Definition of Variables

Variable	Definition
EPS Dispersion	Measures the dispersion of the analyst EPS estimates around their mean value as of December 31, 2020. The standard deviation of analysts' quarterly EPS estimates as of December 31, 2020, is divided by the average EPS estimate (before extraordinary items).
260-Day Share Price Volatility	A measure of the risk of price moves for a stock calculated from the standard deviation of day-to-day logarithmic historical price changes. The 260-day price volatility equals the annualized standard deviation of the relative price change for the 260 most recent trading days' closing price, expressed as a percentage.
EPS / Share Price	The trailing 12 months' earnings per share are normalized by the share closing price.
EPS Volatility / Share Price	The quarterly EPS volatility for 2015 – 2020 normalized by the 12/31/2020 share closing price
EPS YOY Change (%)	The year-over-year change in the firm's quarterly EPS as of 12/31/2020.
Cash Flow Volatility / Share Price	The quarterly cash flow per share volatility for 2015 – 2020 normalized by the 12/31/2020 share closing price.
Annual Dividend Yield (%)	The sum of dividend per share amounts that have gone ex-dividend over the prior 12 months, divided by the current stock price, expressed as a percentage.
Total Analysts	The total number of analysts making recommendations for the firm as of 12/31/2020.
Bloomberg 1-Year Default Prob.	The probability of default of the firm over the next 1-year is calculated by the Bloomberg Issuer Default Risk model.
Log (Market Capitalization)	The log of the total current market value of a company's outstanding shares is stated.
Beta	Measures the volatility of the stock price relative to the volatility in the market index. Beta is the percent change in the price of the stock given a 1% change in the market index.