Environmental Management Information Disclosure Characteristics of Venture Capital-Backed SMEs: Evidence from China

Jia Liu
Shandong University(Weihai), Weihai, Shandong, China

Abstract: In the context of global sustainable development, the relationship between environmental information disclosure and venture capital (VC) has received increasing attention. Based on a sample of 622 VC-backed small and medium-sized enterprises (SMEs) in China from 2019 to 2023, this study manually codes publicly available information and employs descriptive statistics, regression analysis, and multiple imputation to examine the association between environmental disclosure and financing amount as well as financing round. The findings reveal that the overall disclosure rate is only 29.03%, while firms in green industries exhibit a significantly higher disclosure rate (66.67%). After controlling for other factors, environmental disclosure is not significantly associated with either financing amount or financing round. Firm age emerges as the most significant factor affecting financing outcomes. Heterogeneity analysis shows that the financing effects of disclosure are marginally significant in non-green industries and angel rounds. This study uncovers the co-occurrence patterns between environmental disclosure and financing characteristics within the already-funded firm population, providing empirical evidence for understanding the role of ESG information in early-stage VC markets.

Keywords: Environmental Management Information Disclosure; Venture Capital; Small And Medium-Sized Enterprises (SMEs); Financing Amount; Financing Round

1. Introduction

Against the backdrop of global sustainable development and China's "dual carbon" goals, ESG (Environmental, Social, and Governance) concepts are notably permeating from listed companies to the venture capital (VC) field. Severe information asymmetry exists between SMEs and external investors, making non-financial information an important vehicle for value signaling. However, a fundamental empirical question remains insufficiently answered: within the population of firms that have successfully obtained VC funding, what patterns characterize environmental disclosure behavior? Do systematic differences exist in key financing characteristics—such as amount and round—between disclosing and non-disclosing firms? [1,2]

This study focuses on associational analysis in a "state of outcomes" rather than causal inference. Based on a sample of 622 VC-backed SMEs, we manually code environmental disclosure behavior and use descriptive statistics, regression analysis, and multiple imputation to examine the co-occurrence relationships between disclosure and financing amount as well as financing round. Theoretically, this paper enriches empirical evidence on corporate non-financial disclosure in early-stage financing scenarios [3,4]; practically, it provides a market-reality-based reference for entrepreneurs, investment institutions, and policymakers.

2. Literature Review

Existing research focuses on three main areas: (1) the motives, determinants, and economic consequences of corporate environmental disclosure [5-10], almost exclusively using listed company samples [11,12]; (2) factors influencing VC investment decisions, where empirical research on ESG factors lags [13,14]; and (3) the relationship between ESG and SME financing, mostly centered on public markets or debt financing, with micro-level empirical studies on the specific "SME-VC" scenario still nascent[22,23].

In summary, three gaps exist in current research: (i) a mismatch of research subjects-environmental disclosure studies focus on listed firms, while VC research focuses on unlisted startups, and the intersection lacks

in-depth exploration [13,14]; (ii) a single research perspective-the dominant paradigm is one-directional "impact" studies, neglecting the distributional characteristics and association patterns within the already-funded firm population; and (iii) the informational value of "absence" is overlooked-whether a firm's choice not to disclose environmental information sends a specific signal remains unanswered[22]. This study addresses these gaps.

3. Theoretical Analysis and Research Hypotheses

3.1 Theoretical Foundations

This study integrates three major theories, focusing on systematic association patterns within the already-funded firm population. Signaling theory suggests that proactive environmental disclosure can signal management quality and forward-looking orientation, making disclosing firms differ from non-disclosing firms in financing characteristics [15,16]. Legitimacy theory indicates that environmental disclosure is a strategy for actively managing legitimacy resources, so disclosure behavior may correlate with financing round and industry attributes [17,18]. The resource-based view emphasizes that disclosure requires organizational capabilities and financial slack, thus disclosure may naturally correlate with resource endowment variables such as firm age and size [19,20]. Given that the significant financing effects observed in listed company settings may weaken or disappear in early-stage VC scenarios, this study proposes descriptive hypotheses.

3.2 Research Hypotheses

H1 (Descriptive distribution):The proportion of environmental disclosure among VC-backed SMEs is below 50% (H1a); disclosure proportion is significantly higher in green industries than in non-green industries (H1c-1), but does not show a significant upward trend over time (H1b not supported).

H2:After controlling for other factors, firms that disclose environmental information differ significantly from non-disclosing firms in financing amount (expected positive).

H3:Disclosing firms differ significantly from non-disclosing firms in financing round (expected higher round).

4. Research Design

4.1 Sample Selection and Data Sources

Using Zero2IPO Private Equity Database (PEDATA) and ITJuzi as data sources, we select Chinese mainland SMEs that received angel, Series A, or Series B VC funding between 2019 and 2023. Excluding listed companies, firms older than 10 years, and extreme financing events, we obtain an unbalanced cross-sectional dataset of 622 unique firms. Financing events and basic firm information come from the databases; firm age, registration location, etc., are supplemented via Qichacha; environmental disclosure data are collected through Python web scraping and manual coding[21].

4.2 Data Source System

The study adopts a multi-source data cross-validation strategy. Financing event and basic firm information mainly come from Zero2IPO PEDATA and ITJuzi, including financing date, round, amount, and investor names. Firm age, registration location, and industry classification (CSRC industry codes) are checked and supplemented via Qichacha or Tianyancha. PEDATA is the primary database due to its superior completeness in VC event coverage and standardization of disclosed amounts. For financing events missing in PEDATA but present in ITJuzi (about 12% of the initial event pool), ITJuzi data are used as supplements. When the two databases conflict on amount or round for the same event, PEDATA is prioritized; if PEDATA reports a range (e.g., "tens of millions") while ITJuzi reports an exact value, the exact value is used. All integrated data are manually cross-validated.

Environmental disclosure data are obtained via Python scraping, manual collection, and coding [21]. Public channels for information collection include: (i) corporate official websites (environmental disclosures around the financing month); (ii) official WeChat public accounts (historical articles one year prior to financing month); (iii) company profiles on major job platforms (e.g., BOSS Zhipin, Lagou); and (iv) official press releases or authoritative financial media reports related to the financing event (e.g., Pedaily, 36Kr).

Investor characteristics are determined by searching investor names on their official websites, fund prospectuses, and public interviews to identify whether they have explicit

ESG investment themes or impact investment funds.

4.3 Key Variable Measurement and Manual Coding

4.3.1 Dependent variables:

Financing amount (Ln_Amount): Natural logarithm, winsorized at 1% and 99%.

Financing round (Stage): 1 = Angel/Seed, 2 = Series A (including Pre-A, A+), 3 = Series B (including Pre-B, B+).

4.3.2 Independent variable:

Environmental disclosure (Disclose): Coded 1 if, within a 6-month window before the financing month, the firm's official website, WeChat public account, job platform, or financing news release contains substantive content regarding specific environmental management actions, green products/technologies, environmental certifications, or quantified environmental targets (not vague mentions). Coding reliability: percent agreement ≥90%, Cohen's Kappa ≥0.80.

Boundary case handling: Difficult-to-judge cases are separately recorded and archived. Final decisions follow the codebook principles, with typical judgment logic explained in the appendix.

Coding archive: For each "disclose" case, screenshots of the supporting text and URLs are archived to ensure auditability.

Table 1. Coding Manual Examples

Judgment	Example Text	Basis
Disclose(1)	"The company completed an energy-saving renovation in 2022, saving 1.735 million kWh annually, reducing CO2 emissions by 2,575.8 tons."	Specific action + quantitative data
Disclose(1)	"Our products use biodegradable materials and have passed ISO14001 environmental management system certification."	Green product + authoritative certification
Non-disclose(0)	"We are committed to green development, practicing environmental concepts, and building a beautiful China."	Vague statements only, no specifics
Non-disclose(0)	(No environment-related content on corporate website)	No information

Table 2. Variable Definitions and Summary Statistics

Variable	Definition	N	Mean	SD
Dependent: Financing characteristics				
Ln_Amount	Log of financing amount (CNY 10k); midpoints for ranges	428	0.135	1.546
Stage	Financing round: 1=Angel/Seed, 2=Series A, 3=Series B	622	1.542	0.599
Independent variable				
Disclose	1=discloses environmental information, 0=otherwise	620	0.29	0.454
Control variables				
Ln_Age	Log of firm age (years)	580	2.956	0.265
green	Dummy: 1=green industry, 0=non-green	622	0.079	0.269
Region	BTH, YRD, PRD, Other	622	–	–
Year	Financing year	622	–	–

4.3.3 Control variables (Table 2):

Ln_Age: Natural logarithm of firm age (years)

Green: Dummy variable for green industries (1 = green, 0 = non-green)

Region: Beijing-Tianjin-Hebei (BTH), Yangtze River Delta (YRD), Pearl River Delta (PRD), Other

Year: Financing year

4.4 Model Specification

All models use robust standard errors to mitigate heteroskedasticity.

Model 1 (H2-Financing amount): OLS regression

`Ln_Amount_i = $\beta_0$ + $\beta_1$·Disclose_i + $\gamma$·Controls_i + $\varepsilon$_i`

Model 2 (H3-Financing round): Ordered logit model

`P(Stage_i > j) = F($\alpha$_j - ($\beta_1$·Disclose_i + $\gamma$·Controls_i))`, for j = 1, 2

Propensity Score Matching (PSM):** Radius matching is used as an auxiliary analysis to enhance comparability. After matching, the overall bias B = 24.6% (below the 25% threshold), indicating effective mitigation of selection bias.

Robustness checks: Alternative model specifications (quantile regression, Tobit, probit, ordered probit), subsample analyses, changes in variable definitions (3-month and 12-month disclosure windows), placebo tests (randomly generated disclosure status).

Survivorship bias: The sample is limited to firms that have already received VC funding. We cannot observe firms that failed to obtain financing due to non-disclosure. This may bias estimates of the true effect of disclosure. Therefore, conclusions should be interpreted as "association patterns within the already-funded population" rather than population causal effects.

5. Empirical Results and Analysis

5.1 Descriptive Statistics

Among the 622 VC-backed SMEs, the overall environmental disclosure rate is 29.03% (180 firms), indicating that proactive environmental disclosure is not yet common among early-stage firms. By industry, among the 49 green industry firms, 66.67% disclose, significantly higher than the 28.10% in non-green industries (χ^2=10.567, p=0.001). These results support H1a and H1c.

Comparing key variable means by disclosure status: the disclosure group has a slightly higher mean Ln_Amount (0.327 vs. 0.058), slightly lower mean Ln_Age (2.919 vs. 2.971), and slightly higher mean Stage (1.583 vs. 1.525). Significance is tested further.

Temporal trend: Disclosure rates by year: 2019: 30.84%, 2020: 31.18%, 2021: 28.00%, 2022: 30.88%, 2023: 25.00%. Logistic regression (year coefficient = -0.058, p=0.359) and linear probability model (coefficient = -0.012, p=0.361) show no significant upward trend. H1b is not supported.

5.2 Baseline Regression Results

Table 3. Baseline Regression Results

Variable	(1) OLS (Financing Amount)	(2) Ordered Logit (Financing Round)
Core variable		
disclose	0.244 (0.168)	-0.008 (0.193)
Control variables		
ln_age	-0.543*(0.293)	-4.234*** (0.640)
green	-0.472*(0.285)	0.044 (0.315)
Region (ref: BTH)		
Other	-0.161 (0.293)	0.125 (0.324)
PRD	-0.457*(0.262)	0.195 (0.296)
YRD	-0.178 (0.209)	-0.155 (0.247)
Year (ref: 2019)		
2020	-0.108(0.241)	-0.700** (0.299)
2021	-0.436** (0.216)	-0.551** (0.279)
2022	-0.551** (0.235)	-0.338 (0.294)
2023	-1.212*** (0.255)	-0.470 (0.298)
Constant	2.404*** (0.902)	
cut1		-12.999*** (1.864)
cut2		-9.753*** (1.810)
Observations	406	582
R^2/Pseudo R^2	0.111	0.098

* p<0.1, ** p<0.05, *** p<0.01; robust standard errors in parentheses.

Financing amount (H2 not supported): Disclose coefficient = 0.244 (p=0.168), not significant.

Financing round (H3 not supported): Disclose coefficient = -0.008 (p=0.868), not significant.

Control variables: Ln_Age is significantly negative in both models. Green is marginally negative in the amount model (p<0.1). Year effects show decreasing amounts after 2019.

5.3 Robustness Checks

Propensity Score Matching (radius matching): After matching, covariate means bias decreased from 18.3% to 7.4%, median bias from 19.0% to 7.3%, overall bias B = 24.6% (acceptable). ATT = 0.1447 (p≈0.114), still not significant. Results consistent with baseline.

Placebo tests: Randomly generated disclosure dummy variables yield non-significant coefficients (p>0.1), ruling out random drivers.

Subsample analyses: Excluding green firms or using only 2019–2021 sample does not change conclusions.

6. Heterogeneity Analysis

6.1 By Green Industry

Green industry subsample (small N: 31 for amount, 46 for round): Disclose coefficient for amount = -0.884 (p=0.121); for round = 1.216 (p=0.127). Not significant.
Non-green industry subsample: Disclose coefficient for amount = 0.357 (p=0.043), significant at 5%; for round = -0.045 (p=0.825), not significant.
The positive amount effect exists mainly in non-green industries.

6.2 By Region
BTH region: Amount coefficient = 0.022 (p=0.960); round coefficient = 1.019 (p=0.057)-marginally significant for round.
YRD, PRD, Other: No significant effects.

6.3 By Firm Age (split at median Ln_Age = 3.0157)
Young and mature groups both show non-significant disclosure coefficients in amount models.

6.4 By Financing Round
Angel round subsample (Stage=1): Disclose coefficient = 0.386 (p=0.083)-marginally significant positive effect on amount.
A/B round subsample (Stage≥2): Disclose coefficient = -0.033 (p=0.901)-not significant.
The positive amount effect is concentrated in angel rounds.

7. Conclusion, Implications, and Limitations

7.1 Main Conclusions
(1) Environmental disclosure is not yet common among early-stage SMEs (overall 29.03%), but significantly higher in green industries (66.67%). (2) No statistically significant association is found between environmental disclosure and financing amount or round, though marginally significant positive effects appear in non-green industries and angel rounds. (3) Firm age is the most significant factor affecting financing outcomes-younger firms obtain larger amounts and later rounds. (4) Significant regional and temporal variations exist in the financing environment.

7.2 Theoretical Contributions and Practical Implications
The study delineates the functional boundaries of ESG information in early-stage VC financing, showing non-significant effects. Firms should weigh disclosure costs and benefits, prioritizing substantive environmental actions. Investment institutions should gradually establish internal ESG assessment criteria, focusing on substantive environmental performance. Policymakers could provide tangible incentives for firms that disclose and perform well.

7.3 Limitations and Future Research
Limitations include survivorship bias, limited causal inference, unaddressed disclosure quality, and small green industry sample size. Future research could employ Heckman models, panel data with difference-in-differences, and other more rigorous methods.

References
[1] Drobetz, W., El Ghoul, S., Guedhami, O., Hackmann, J. P., & Momtaz, P. P. (2024). Entrepreneurial finance and sustainability: Do institutional investors impact the ESG performance of SMEs? Journal of Business Venturing Insights, 22, e00498.
[2] Kleinert, S., & Hildebrand, M. (2025). Venture capitalists' decision-making in hot and cold markets: The effect of signals and cheap talk. Entrepreneurship Theory and Practice, 49(2), 571-598.
[3] Deng, M., Shao, Q., & Chen, S. (2025). The ESG–financial performance nexus in startups: A multi-level contingency framework in the emerging economy. Sustainability, 17(24), 11197.
[4] Zhou, L., & Rahat, B. (2025). How do environmental, social, and governance factors impact entrepreneurial finance? Evidence from European small and medium enterprises. International Journal of Entrepreneurial Behavior & Research, ahead-of-print.
[5] Li, Z., Li, F., & Zhao, Y. (2024). The financing effect of environmental information disclosure: Evidence from heavy polluters. Audit & Economics Research, 39(1), 117-127. (in Chinese)
[6] Chen, R., et al. (2022). The impact of ESG indices on cost of capital. Scientific Decision Making, (11), 24-40. (in Chinese)
[7] Yi, X., & Zhou, Y. (2023). Environmental information disclosure, media attention, and corporate inefficient investment: Based on heavy polluter data. Finance & Accounting Communications, (5), 56-64. (in Chinese)
[8] Fu, Y., Chen, J., & Liu, S. (2024). Green

innovation effects of carbon information disclosure. East China Economic Management, 38(7), 27-38. (in Chinese)
[9] Wang, Y., & Albitar, K. (2025). A novel metric for corporate environmental responsibility and its impact on investment inefficiency. International Journal of Finance and Economics, 30(3), 3003-3030.
[10] Cai, R. (2022). Environmental information disclosure quality of listed companies based on annual report text characteristics and its economic consequences. Beijing: Economic Science Press. (in Chinese)
[11] Wang, Y., & Tan, X. (2021). Financing constraints, environmental information disclosure quality, and R&D investment. Accounting Friends, (2), 56-64. (in Chinese)
[12] Mei, X., Ge, Y., & Zhu, X. (2020). Mechanism of environmental legitimacy pressure on corporate carbon information disclosure. Soft Science, 34(8), 14-20. (in Chinese)
[13] Xu, J., Li, F., & Liu, W. (2025). The impact of green finance on corporate greenwashing risk. Frontiers of Engineering Management Technology. (in Chinese)
[14] Uong, L. T. N., & Nguyen, N. D. (2025). The disclosure of environmental, social, and governance reports of small and medium-sized enterprises in the emerging market. Ho Chi Minh City Open University Journal of Science - Economics and Business Administration, 15(3), 107-123.
[15] Spence, M. (1973). Job market signaling. Quarterly Journal of Economics, 87(3), 355-374.
[16] Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling theory: A review and assessment. Journal of Management, 37(1), 39-67.
[17] Suchman, M. C. (1995). Managing legitimacy: Strategic and institutional approaches. Academy of Management Review, 20(3), 571-610.
[18] Campbell, J. L. (2007). Why would corporations behave in socially responsible ways? An institutional theory of corporate social responsibility. Academy of Management Review, 32(3), 946-967.
[19] Drempetic, S., Klein, C., & Zwergel, B. (2019). The influence of firm size on the ESG score: Corporate sustainability ratings under review. Journal of Business Ethics, 167(2), 333-360.
[20] Barney, J. (1991). Firm resources and sustained competitive advantage. Journal of Management, 17(1), 99-120.
[21] Beck, J., Steinberg, A., Dimmelmeier, A., et al. (2025). Addressing data gaps in sustainability reporting: A benchmark dataset for greenhouse gas emission extraction. Scientific Data, 12, 1497.
[22] Li, Q., & Song, J. (2022). Performance expectation gaps and corporate greenwashing behavior. Journal of Nanjing Audit University, 19(3), 51-61. (in Chinese)
[23] Lilo, R., Paemurru, E., & Pärl, Ü. (2025). The textual representation of double materiality in ESG reports: Developing a content analysis codebook within a business model perspective. Journal of Business Models, 13(2).

Appendix: Robustness Check for Missing Financing Amount Values Using Multiple Imputation

A.1 Missing Data Description

As noted in the main text, 192 out of 622 firms (30.9%) had missing or range-based (e.g., “tens of millions”) financing amounts that could not be accurately assigned a midpoint, leaving 428 valid observations for Ln_Amount. To test whether this missingness biases estimates, we apply multiple imputation (MI) and re-estimate the baseline OLS model.

A.2 Imputation Model Specification

Method: Chained equations multiple imputation (MICE), predictive mean matching (PMM, k=5), 20 imputed datasets.

Variables in imputation model: Financing amount (original amount, then log-transformed after imputation); predictors: Disclose, Ln_Age, green, region dummies, year dummies, Stage. Cases with missing Ln_Age (40 observations, 6.4%) are listwise deleted before imputation, final sample N = 582.

A.3 Regression Results After Multiple Imputation

Table A-1: OLS Regression on Financing Amount (Multiple Imputation)

Variable	Model 1 (Complete case, N=428)	Model 2 (MI full sample, N=582)
Disclose	0.244 (0.168)	0.208 (0.171)
Ln_Age	-0.543*(0.293)	-0.628**(0.301)

green	-0.472(0.285)*	-0.470(0.288)
Region & Year FE	Yes	Yes
R^2 / Pooled R^2	0.111	0.107
observations	428	582

* p<0.1, ** p<0.05, *** p<0.01; robust standard errors in parentheses.

The coefficient for Disclose is 0.208 (p=0.227), highly consistent with the baseline (0.244, p=0.168) in sign, magnitude, and statistical significance. The conclusion that environmental disclosure is not significantly associated with financing amount is robust to missing data.

A.4 Comparison of Methods

Table A-2: Comparison of Estimation Methods

Method	Disclose coefficient	Significance	Conclusion
Baseline OLS (complete case)	0.244	p=0.168	Not significant
Multiple imputation (full sample)	0.208	p=0.227	Not significant
PSM matched sample	0.211	p=0.211	Not significant
Quantile regression (median)	0.198	p=0.201	Not significant

Thus, the main conclusion remains robust.In summary, after handling missing values in financing amount using multiple imputation, none of the key conclusions changed materially. Therefore, the finding in the main text that “there is no significant association between environmental disclosure and financing amount” is robust, and no estimation bias due to missing data is present. The analysis in the appendix provides strong statistical support for the main conclusions.

