

Measuring Corporate Governance Quality in Swiss Listed Firms Using Textual Analysis and Readability Metrics

Jeta Lami-Beka ¹, Prof. Dr. Michael Burkert ², Prof. Dr. Raúl Barroso ³

^{1,2,3} University of Fribourg, Switzerland

ABSTRACT

This paper presents a novel, multi-method approach for assessing corporate governance quality (CGQ) in publicly listed Swiss firms by integrating computer-aided text analysis (CATA), natural language processing (NLP), and machine learning (ML). Drawing on a longitudinal dataset of 1'573 annual reports spanning 18 years (2002 – 2019), we develop and validate governance-specific dictionaries tailored to the Swiss regulatory and linguistic context. Our final dictionary, constructed through supervised learning, comprises 3'627 positive and 277 negative terms and word combinations, capturing nuanced dimensions of governance discourse. To ensure robustness, we conduct convergent and predictive validity tests using the fine-grained governance proxy proposed by Barroso et al. (2016), and successfully replicate key findings from their study. This confirms the reliability of our text-based governance measure and its alignment with established governance constructs.

Extending the analysis, we incorporate the Gunning Fog Index to examine the relationship between language complexity and governance quality. Our results reveal a significant inverse correlation: firms with higher governance scores tend to produce more readable and accessible disclosures. This finding supports theoretical expectations from agency theory and legitimacy theory, suggesting that well-governed firms prioritize transparency and stakeholder communication.

By combining dictionary-based scoring, readability metrics and machine learning, this study contributes to the emerging literature on automated governance assessment. It demonstrates the feasibility of scalable, data-driven approaches to evaluating CGQ and offers practical implications for regulators, investors and scholars seeking to monitor governance practices through publicly available textual data.

Keywords: Corporate Governance; Machine Learning; Natural Language Processing; Language Complexity