Data, Technology, and Firm Dynamics: An LLM-Based Analysis

February 28, 2025 Tianxiang Zheng¹

Extended Abstract

This paper examines the interplay between data usage and data technology adoption and their joint impact on firm dynamics using a panel of 3,031 publicly traded Chinese firms over the period 2016–2022. In the context of the ongoing digital transformation associated with the Fourth Industrial Revolution, firms increasingly rely on data not only as a byproduct of operations but also as a critical input in production and strategic decision-making. However, traditional approaches to measuring these intangible assets have proven inadequate. To address this gap, we introduce a novel measurement framework that leverages large language models (LLMs) to extract quantitative assessments from the Managements Discussion and Analysis sections of annual reports. This approach yields multidimensional scores that capture various facets of data usage and data technology adoption–including data infrastructure, collection, processing techniques, and data architecture.

Our empirical strategy first identifies the determinants of changes in data usage and data technology adoption. We document a convergence effect, whereby firms with higher initial levels of either data usage or technology adoption exhibit smaller subsequent improvements. This result suggests diminishing marginal returns as firms advance in their digital integration. Conditional on initial scores, our analysis further reveals that higher levels of data usage facilitate more rapid improvements in technology adoption, indicating complementarities between the two dimensions.

Second, we investigate the economic consequences of changes in data usage and technology adoption for firm performance. Employing a long-difference specification and controlling for a comprehensive set of firm characteristics and fixed effects, our findings indicate that improvements in both data usage and technology adoption are significantly associated with higher sales, increased total factor productivity, greater capital investment, and expanded employment. In particular, disaggregated analysis reveals that the adoption of data technologies

¹Department of Economics, University College London. Tianxiang.Zheng.22@ucl.ac.uk.

aimed at enhancing production efficiency yields the most pronounced benefits for employment growth.

A further contribution of this study is the demonstration that conventional estimates of the impact of artificial intelligence (AI) on labor outcomes may be substantially biased if they neglect the complementary effects of broader data usage and data technology dimensions. Our evidence suggests that such omission may lead to an overestimation of AI's effect on employment by as much as 60%.

This research makes three key contributions to the existing literature. First, it offers an innovative LLM-based method to measure firm-level data usage and technology adoption, overcoming the limitations of traditional proxies. Second, by integrating data usage and data technology within a unified analytical framework, the study provides a more holistic understanding of digital transformation in firms. Third, by focusing on an emerging market context, it broadens the empirical evidence on digital transformation beyond the predominantly studied advanced economies. The findings have important implications for both managers and policymakers, underscoring the need for strategies that foster comprehensive digital integration to sustain firm competitiveness and promote broader economic growth.