Copyright and Competition: Estimating Supply and Demand with Unstructured Data

SUKJIN HAN, University of Bristol, UK

KYUNGHO LEE, Yale University, USA

Copyright policies play a pivotal role in protecting the intellectual property of creators and companies in creative industries. The advent of cost-reducing technologies, such as generative AI, in these industries calls for renewed attention to the role of these policies. This paper studies competition in a market of creatively differentiated products and the competitive and welfare effects of copyright protection. A common feature of products with creative elements is that their key attributes (e.g., images and text) are *unstructured* and thus high-dimensional. We focus on a stylized design product, fonts, and use data from the world's largest online marketplace for fonts. We construct neural network embeddings to quantify unstructured attributes and measure the visual similarity. We show that this measure closely aligns with actual human perception. Based on this measure, we empirically find in a descriptive analysis that competitions occur locally in the visual characteristics space. We then develop a structural model for supply and demand that integrate the embeddings. On the supply side, our model describes firms' location choices within the visual characteristics space as well as pricing and entry decisions. A copyright policy is modeled as imposing restrictions on the area of possible choices in the characteristics space, providing local protection to right holders. On the demand side, we characterize consumers' heterogeneous preferences for visual attributes, focusing on recovering substitution patterns across different designs. Overall, the estimated demand model reveals that the substitution patterns are effectively explained by the visual similarity. The estimated supply-side model indicates that the firm's development costs are low when mimicking close competitors and increase as products become more visually differentiated. This suggests the existence of a mimicking externality: the presence of a design product reduces the fixed costs of visually similar entrants. Through counterfactual analyses, we find that local copyright protection can enhance consumer welfare when products are relocated, and the interplay between copyright and cost-reducing technologies is essential in determining an optimal policy for social welfare. We believe that the embedding analysis and empirical models introduced in this paper can be applicable to a range of industries where unstructured data captures essential features of products and markets.

CCS Concepts: • **Social and professional topics** → **Copyrights**; • **General and reference** → **Empirical studies**; **Estimation**; • **Applied computing** → **Economics**; • **Computing methodologies** → **Neural networks**; **Dimensionality reduction and manifold learning**.

Additional Key Words and Phrases: Creative industries, Unstructured data, Embeddings, Visual similarity, Consumer demand, Product positioning

Authors' Contact Information: Sukjin Han, sukjin.han@gmail.com, University of Bristol, Bristol, UK; Kyungho Lee, kyungho.lee@yale.edu, Economics, Yale University, New Haven, USA.