



U.S.E. Research Institute
Working Paper Series 21-10

Identifying Politically Connected Firms: A Machine Learning Approach

Deni Mazrekaj¹

Vitezslav Titl^{2,3}

Fritz Schiltz⁴

¹ Department of Sociology, University of Oxford

² Utrecht University School of Economics, Utrecht University

³ Department of Economics, Faculty of Law, Charles University

⁴ Leuven Economics of Education Research, KU Leuven

June 2021

Abstract

This article introduces machine learning techniques to identify politically connected firms. By assembling information from publicly available sources and the Orbis company database, we constructed a novel firm population dataset from Czechia in which various forms of political connections can be determined. The data about firms' connections are unique and comprehensive. They include political donations by the firm, having members of managerial boards who donated to a political party, and having members of boards who ran for political office. The results indicate that over 85% of firms with political connections can be accurately identified by the proposed algorithms. The model obtains this high accuracy by using only firm-level financial and industry indicators that are widely available in most countries. We propose that machine learning algorithms should be used by public institutions to identify politically connected firms with potentially large conflicts of interests, and we provide easy to implement R code to replicate our results.