**Measuring Technological Change with Web-Based Data and Machine Learning**

- Bearbeiter: Patrick Breithaupt
- Titel: Measuring Technological Change with Web-Based Data and Machine Learning
- Kategorie: Promotion
- Fachgebiet: Volkswirtschaftslehre
- Status: Abgeschlossen am 15.07.2024
- GutachterInnen: Prof. Dr. Irene Bertschek, Prof. Dr. Peter Winker

- Abstract: This dissertation consists of four essays and was written between July 2019 and March 2024. During that time, I was employed as a researcher at ZEW – Leibniz Centre for European Economic Research. The dissertation is submitted in partial fulfilment of the requirements for the academic degree of "doctor rerum politicarum" (Dr. rer. pol.) at the Justus Liebig University Giessen, where I was an external doctoral candidate at the Chair of Economics of Digitalisation. Chapter 1 is an introduction to the dissertation. In this part, the four essays get embedded in the economic, machine learning, and data mining literature. Furthermore, I provide definitions and present research gaps that motivate the essays. The first two essays introduce approaches for measuring technological change in firms. In both essays, we use firm website data and machine learning techniques. Chapter 2 (Essay 1) is concerned with the measurement of innovation, one of the main drivers of economic growth. In co-authorship with Janna Axenbeck, I use natural language processing and machine learning techniques to analyse firm websites. We find that firm websites contain information that can be used to detect and measure innovation activity. Compared to traditional survey-based measures, our model is cost-effective, can be updated quickly, and is available for a large number of German firms. Chapter 3 (Essay 2) deals with measuring the case of digital technologies as one prominent example of general purpose technologies (GPTs). Co-authored with Janna Axenbeck, I use natural language processing and machine learning techniques to analyse newspaper articles as well as firm websites to estimate a firm-level digitalisation score. We present a methodology for producing a digitalisation indicator that is in line with several traditional measures. Finally, we analyse the link of this indicator with firm resilience during the COVID-19 crisis as an example of a policy-relevant application. In the last two essays, we measure intangible and human capital that is linked to technological change. In both essays, we use data on firms from digital platforms. Chapter 4 (Essay 3) focuses on firm-level indicators for intangible capital. In co-authorship with Reinhold Kesler, Thomas Niebel, and Christian Rammer, I use data from the platforms Facebook and Kununu. By using this public data, we introduce two indicators for marketing and on-the-job training that are positively related to corresponding data from the Community Innovation Survey (CIS). The proposed approach is more cost-effective and can be updated more often than a survey-based measure. Chapter 5 (Essay 4) is concerned with measuring human capital. Co-authored with Hanna Hottenrott, Christian Rammer, and Konstantin Römer, I create a Linked Employer-Employee (LEE) data set using public data from the career-oriented social networking platform XING. The data contain information on employers, employees, and employee flows. By using independent data, we show that the XING data are plausible, can be used for subsequent economic research, and represent a viable alternative to official data. Chapter 6 summarises the findings of the four essays, provides concluding remarks, and points out possible directions for future research.