Political Connections and Business Performance: Evidence from Turkish Firms

Mehmet Ataberk Atan[1]

[1]*Department of Economics and Business, Central European University*

In partial fulfillment of the requirements for the degree of Masters of Arts

Supervisor: Assistant Professor Tomy Lee

Vienna, Austria

2023

Abstract

I investigate the impact of political connections on the economic performance of Turkish industrial firms. I propose a novel approach to measure connections using text data, considering both textual content and associated sentiment. By collecting online news articles on firms and politicians and matching them to respective firms, I examine the effect of shocks to connections between politicians and companies, constructed through sentiment analysis. Utilizing financial data from the top 1000 industrial companies over a decade, I employ a fixed effects model to analyze the influence of these connections on profitability. My hypothesis suggests that companies with connections outperform those without connections, both in terms of average performance and additional profits. Confirming this hypothesis, the analysis demonstrates that connected companies exhibit better average performance and experience higher profits compared to non-connected companies. However, when clustering standard errors at different levels, the second result, indicating a positive influence of connections on firms' performance, does not maintain statistical significance. These findings highlight potential limitations in the conservative methodology used to construct connections, suggesting it may not effectively capture actual connections between companies and politicians. I argue that while the measured connections show a positive impact on performance, the methodology may not capture signals strong enough due to challenges in uncovering underlying business-politics ties in Turkey. Consequently, the study emphasizes the need for more sophisticated methods to unveil and measure political connections and reaffirms the role of political institutions in shaping business and economic outcomes.

Acknowledgements

I would like to express my deepest gratitude to my parents, Şehnaz and Fahri, for their enduring support and belief in me throughout my academic journey. Their constant encouragement, love, and sacrifices have been the foundation of my success, and I will forever cherish their presence in my life.

I want to extend a sincere thanks to Yiğit Tahmisoğlu, whose inspiration and support have played a significant role in shaping my ideas and keeping me motivated. His guidance, patience, and unwavering belief in me have been invaluable, and I am grateful for the role he has played in my academic growth.

A special word of thanks goes to my supervisor, Professor Tomy Lee, for providing invaluable feedback, guidance, and mentorship. His expertise, insights, and support have been instrumental in refining my thesis, and I am truly appreciative of his contributions.

I would also like to acknowledge and express my gratitude to Professor Arieda Muço for her exceptional courses on Political Economy and Natural Language Processing. Her teaching has equipped me with a solid foundation and inspired the ideas that I have incorporated into my thesis.

Lastly, I want to extend my heartfelt appreciation to my classmates in the department and my friends at the university. Your friendship, intellectual discussions, and support have created an enriching environment that has helped me grow as a scholar. The late-night study sessions and shared experiences have been a source of motivation and inspiration, and I am genuinely grateful for the bond we have formed on this challenging journey.

Contents

List of Figures

List of Tables

List of Abbreviations

AKP Adalet ve Kalkınma Partisi (Justice and Development Party)

NLP Natural Language Processing

RNN Recurrent Neural Network

LSTM Long Short-term Memory

EBITDA Earnings Before Interest, Taxes, Depreciation, and Amortization

FMC-1 Alternative Connection Definition 1

FMC-2 Alternative Connection Definition 2

İSO İstanbul Sanayi Odası (İstanbul Chamber of Industry)

MÜSİAD Müstakil Sanayici ve İşadamları Derneği (Independent Industrialists and Businessmen Association)

TUSKON Türkiye İşadamları ve Sanayiciler Konfederasyonu (Turkish Confederation of Businessmen and Industrialists)

1 Introduction

Political connections have been recognized as a significant factor influencing the economic performance of firms, particularly in emerging market economies. The intricate interplay between business and politics in these contexts necessitates a thorough examination of the impact of political connections on firm outcomes. In this thesis, I investigate the influence of political connections on the economic performance of Turkish industrial firms.

The relationship between business and politics is an age-old phenomenon, and its impact on the economic returns of firms, individuals and electorates has been widely studied (Fisman 2001, Labonne et al. 2017, Ferguson 2008, Li et al. 2008, Muttakin et al. 2015). On the other side of the equation, the social costs in the form of corruption, lack of regulatory monitoring, higher labor related mortality (Xu 2018, Fisman and Wang 2015) brought upon have been scrutinized. With rise of social media, and online media in general, the mechanisms through which the social burdens of politically connected entities can be tamed in the context of imperfect democracies and emerging economies have also been explored (Sonin et al. 2018). Turkey, being no exception to such countries, follows a similar pattern (Gürakar et al. 2016, Önder et al. 2011).

The Turkish economy is widely known to be steered heavily by the ruling government, which has reigned uninterrupted since 2002, and made way for many current and ex-ministers and members of parliament to be well-known business persons. Using media data and building on Özcan and Gündüz (2015), I analyze a panel that covers the top 1000 industrial companies in Turkey over a period of 10 years, which constitutes roughly 10% of GDP and 40% of exports, to assess the impact of such connections on the performance of companies.

I have formulated two hypotheses to test the impact of political connections on the performance of firms. The first hypothesis posits that connected firms perform better on average, as politicians are more likely to build and maintain connections with larger firms that can extract more rents and build more political capital while firms profitable and large enough are the ones with enough resources to sustain connections and make significant benefits out of them in the forms of easier access to government contracts, better access to credit and financing, and a higher likelihood of receiving favorable regulatory decisions. The second hypothesis suggests that political connections have a positive impact on firms' performance. Meaning, not only firms and politicians believe that such connections will result in economic benefits but firms actually reap additional profits because they are connected, compared to their unconnected counterparts.

I use a fixed-effects strategy to analyze the effects of shocks to connections on the business outcomes of industrial companies. This helps to control for the firm level fixed effects which absorb the personal ability of the managers or owners of the firms to build connections of all sorts which is beneficial in any business setting and not only for politics. I also develop and use an algorithm to automate the search process and collect news articles on politically connected firms. I combine the collected news articles with the data provided in the ”Turkey's 500 Biggest Industrial Companies” survey (İstanbul Chamber of Industry, 2021). The survey covers top organizations operating in the industrial sector within the borders of Turkey, and among the key variables are sales from production, net sales, number of employees, ownership structure, EBITDA, and aggregate balance sheet items.

The methodology employed in this thesis to measure political connections involves automated searches on two target news websites, using a combination of the name of a firm and a number of political signifiers. All articles retrieved from these searches are then subjected to sentiment analysis using an LSTM-based model to assign sentiment scores. The matching process links each article to its corresponding firm-year observation based on the firm name and publication year. The sentiment scores are then aggregated in three progressively more restrictive ways to define political connections. This includes considering all matched articles, articles where the first word of the firm's name is mentioned, and articles where the first two words of the firm's name occur together. By employing this methodology, I derive connection measures that capture different levels of connectedness, enabling an investigation into the impact of political connections on firm performance.

In the following subsections, I give an institutional background for the Turkish context and present and in-depth review of the literature on the effects of political connections. Then, I provide a detailed account of the methodology and empirical framework employed. Building on this framework, I discuss the results and test whether they hold in alternative settings. Finally, I analyze the limitations of the study, offer suggestions for future research, and conclude.

1.1 Institutional Background

Justice and Development Party (Adalet ve Kalkınma Partisi, from now on will be referred as AKP), led de-facto by President Recep Tayyip Erdoğan has been in power continuously since 2002. Since the day it was founded, AKP has come the first party in the parliament in all of the elections it has participated in and has been in power alone in five of the six general elections (2002, 2007, 2011, November 2015 and 2018). Although started as a right-wing, conservative-democrat, pro-market liberal party in the beginning, the party has come to be characterized as Neo-Ottomanist, Nationalist and Far-Right, tightening their authoritarian grip on the economy, media and the social realm, increasingly so in the last decade. Most significantly, the party has changed the regime in Turkey from parliamentarian democracy to presidential democracy, in doing so effectively merging executive and legislative powers in President (formerly Prime Minister) Erdoğan's hands.

Turkey's economic landscape is deeply intertwined with its political system, where political connections play a prominent role in shaping business dynamics and outcomes. The country's unique institutional context, characterized by a blend of illiberal democracy and political patronage, creates an environment where firms' relationships with politicians can significantly impact their performance.

Political parties often rely on support from business elites and vice versa, leading to the formation of intricate networks of patronage and clientelism. These networks serve as conduits for various forms of political influence and access to resources, which can greatly shape business opportunities and outcomes.

Moreover, Turkey's economic structure is dominated by the industrial sector, accounting for the 29% of employment and 36% of value of goods and services produced, encompasses a wide range of manufacturing and production activities (TurkStat 2022). Industrial firms, therefore, operate within a complex network of relationships with political actors, including politicians, government officials, and regulatory bodies. These connections can provide firms with advantages such as preferential access to contracts, subsidies, licenses, and regulatory leniency, influencing their economic performance.

In countries where political decision-making is mostly non-transparent, as in Turkey under AKP, figuring out political connections is not straightforward. To study a company's political links in Turkey, one needs to know about its relationships with many government groups and have a way to aggregate these connections. Additionally, getting the needed information is hard because business-political relationships are sensitive, often covert, and fluid.

Against the backdrop of Turkey's political landscape dominated by the Justice and Development Party (AKP) and its consolidation of power under President Recep Tayyip Erdoğan, the intertwined nature of the country's political and economic systems becomes evident. The AKP's shift towards a more authoritarian approach and the merging of executive and legislative powers have intensified the role of political connections in shaping business dynamics and outcomes. I explain the approach I take in attempting to bypass such challenges in measuring and evaluating connections in the Methodology section.

1.2 Literature

A crucial building block of this study, Özcan and Gündüz (2015) investigates the impact of direct personal and indirect network connections on business performance in Turkey. They distinguished between direct personal ties and indirect institutional ties, aiming to understand their differential effects on business outcomes. The authors utilized the membership databases of MÜSİAD and TUSKON, which were business organizations with close government ties. However, it is important to note that the database of TUSKON is no longer available as the organization dissolved following its affiliation with the Gülen Movement, which was designated as a terrorist organization after the 2016 coup attempt in Turkey. The unavailability of the necessary databases do not allow this study to replicate a similar procedure to reveal institutional ties and analyze their distinct effect from personal ties.

Two types of political connections were distinguished: direct personal connections and network connections. To identify direct personal connections, manual search engine queries were performed. For the indirect network approach, the researchers examined firms' affiliations with business associations known for their government connections. While some of these associations' ties could be verified through official websites, additional manual internet searches were conducted for further information.

In this study, I build upon the work of Özcan and Gündüz (2015), who examined the impact of direct personal and indirect network connections on business performance in Turkey. Utilizing the same source of data for firms' financial information and using a similar algorithm to form search queries for news articles, this study takes a step further by adopting a more rigorous approach. By employing econometric methods, I aim to quantitatively assess the impact of political connections on firm outcomes. By expanding upon the foundation laid by Özcan and Gündüz, this study contributes to a deeper understanding of the relationship between political connections and business performance in Turkey.

Fisman (2001) is a significant contribution to the field of political economy. This seminal work examines the relationship between a firm's level of political connectedness and the returns on its stock price in Indonesia. By utilizing an index developed by a private business consulting firm and incorporating expert anecdotal knowledge, Fisman establishes a robust and positive correlation between a firm's political connections and the financial performance reflected in its stock returns. This study lays a foundation for understanding the value of political connections in the realm of business and provides valuable insights into the dynamics of the Indonesian context. Li et al. (2008) conducted a study in China to examine the impact of affiliations with the ruling Communist Party on the performance of private enterprises. Using a comprehensive nationwide survey of private firms, the study found a positive correlation between party membership of private entrepreneurs and firm performance. This relationship remained significant even after controlling for factors like human capital and other relevant variables.

In a separate study, Muttakin et al. (2015) compared the performance of family firms to non-family firms and explored the role of political connections. The findings revealed that family firms, in general, outperformed non-family firms. Additionally, politically connected family firms demonstrated even better performance compared to family firms without such connections.

These studies collectively emphasize the significant influence of political connections on firm performance, particularly within the context of family businesses and private enterprises. They highlight the importance of considering political affiliations when examining the factors that contribute to the success of firms.

Sonin et al. (2018) conducted a study to explore the role of new media, specifically blog posts published by Russian opposition figure Aleksei Navalny, in promoting accountability in non-democratic countries with limited media freedom. The authors focused on corruption in Russian state-controlled companies and investigated the impact of blog posts exposing such corruption on market returns. The study found a negative causal effect on market returns following the publication of these blog posts. The authors utilized precise timing and within-day analysis with company-day fixed effects to establish causality. Furthermore, the study revealed that these blog posts were associated with higher management turnover and fewer conflicts among minority shareholders. These findings suggest that social media, despite censorship and limited political competition, can play a role in disciplining corruption.

In addition to employing a similar methodology for identification purposes, I utilize online news data to thoroughly analyze the contextual information within texts. However, a notable distinction arises as the research focuses on a broader range of firms, examining their balance sheet and income statement items instead of solely relying on stock market returns. This approach enables a thorough assessment of the long-term economic value associated with political connections, extending beyond the immediate response of investors in the stock markets to news events.

Another divergence lies in the methodology employed. While Sonin et al. (2018) utilized negative coverage of a specific set of firms, identified based on their ownership structures, I employ positive coverage of firms from broader mass media channels to establish connections, treating the publications as naive signals of connections. These positive signals are treated as indicators of political ties, contributing to the overall analysis.

2 Data

I use data from three sources; the Istanbul Chamber of Industry Survey of Top 1000 Ranking Industrial Companies (2021) and news article texts from Hürriyet and CNNTürk, two Turkish publications.

Founded in 1952 with the written request of nearly 750 industrialists, the Istanbul Chamber of Industry (ISO) continues to be the largest chamber of industry in Turkey with more than 22 thousand members (2022). The survey by İSO originally included the top 500 industrial companies by sales from production, later extended to the next 500 largest companies.

Hürriyet, established in 1948, was initially known for its relatively oppositional stance. However, following its acquisition by Demirören Holding in 2018, the publication experienced extensive layoffs and a substantial number of its editorial staff resigned in protest. Consequently, there was a noticeable shift towards a publishing approach that aligned more closely with the government.

CNN Türk, primarily operating as a news channel, maintains a highly active online newspaper platform. Established in 1999, the ownership of CNN Türk is divided between Demirören Holding and TBS Networks, an American television network that also holds ownership of CNN.

I chose the newspapers based on their political stance and overall news coverage. Moreover, both are frequently followed mainstream outlets. Although they are both relatively close to the government at the moment, the aim here is to estimate a lower bound by creating a more conservative political connection variable.

2.1 Firms - Turkey's 500 Biggest Industrial Companies

I utilize data provided by the Istanbul Chamber of Industry, specifically from the ”Turkey's 500 Biggest Industrial Companies” survey. This annual survey is complemented by a similar survey for the subsequent 500 largest companies, yielding a total of 1000 observations per year.

The surveys encompass organizations operating in the industrial sector—including mining and quarrying, manufacturing, and electricity generation—within the borders of Turkey. It is important to note that the surveys exclude organizations primarily engaged in agriculture, livestock, wholesale or retail trade, services, construction, finance, insurance, and related fields.

Eligibility for participation in the survey is determined based on each company's net sales from production, with a minimum threshold that is revised annually. Companies exceeding this threshold are invited to complete the survey questionnaire. Companies from eligible fields of work are highly incentivized to participate in the survey since participation brings prestige and promotion of the company's name. Therefore, ranking in the top 1000 and abstaining is expected to be low.

The survey collects data on several key variables, such as net sales, sales from production, average number of employees, ownership structure (state/public/private), EBITDA, and asset size.

Variable	EBITDA	Avg. No. of Employees	Exports (thousand $s)	Leverage Ratio	Sales from Production	Predicted Sentiment	FMC-1	FMC-2
count	3800	6127	6451	5892	10000	10000	10000	10000
mean	1.63821e+08	1135.72	94445.4	2.65406	9.16091e+08	0.1794	0.1512	0.0278
std	6.73987e+08	1431.65	329466	15.7129	3.09273e+09	0.686923	0.607104	0.209837
min	-7.97219e+09	6	0.39	-369.615	7.43637e+07	0	0	0
25%	1.74868e+07	390	10576	0.675482	1.91088e+08	0	0	0
50%	4.36541e+07	694	33072	1.55477	3.44842e+08	0	0	0
75%	1.08028e+08	1308.5	70782.5	2.71525	7.40327e+08	0	0	0
max	1.70351e+10	18110	5.86988e+06	403.094	1.36793e+11	16	12	5

Table 1: Descriptive Statistics of Selected Variables

The Istanbul Chamber of Industry has been conducting the aforementioned survey since 1968, providing readily accessible data on the largest 1000 firms for the years 2009-2021 on the official survey website. For the purpose of the following analyses, a period of 10 years, spanning from 2012 to 2021, is utilized. However, it is important to note that attrition and the availability of the main outcome variable only from 2013 onwards restrict the analysis to a 9-year window, wherein less than 60% of the firms can be traced throughout the entire period.

2.2 News Articles - Hürriyet and CNN Türk

I develop an algorithm that streamlines the search process while simultaneously documenting the steps involved in collection of the news articles. The analysis of media articles obtained through automated search queries involves the application of natural language processing (NLP) techniques, with a specific focus on sentiment analysis. It is important to note that sentiment analysis is a fundamental tool within the broader NLP toolkit, enabling the extraction of sentiment or emotional information from textual data. In this study, the sentiment analysis is performed using Recurrent Neural Network (RNN) models, which take into account the sequential ordering of words in the text when learning the sentiment values. By quantifying the information contained within the texts and operationalizing a variable based on the sentiment values, this approach facilitates a robust foundation for subsequent econometric analysis.

The algorithm employed for collecting news articles follows a systematic procedure. To ensure a balanced and focused search, the official names of each company are truncated to the first two words. This approach aims to prevent an excessive inclusion of keywords that could overly restrict the search. Subsequently, these truncated names are combined with specific political keywords, namely ”bakan” (minister), ”başbakan” (prime minister), and ”cumhurbaşkanı” (president).

It is important to note that until 2014 Erdoğan held the office of prime minister. Since the 2018 general elections following the 2017 regime change referendum, the office of prime minister has been abolished and the powers the office have centralized in the presidency. Additionally, it should be acknowledged that these keywords primarily retrieve articles related to ministers, the prime minister, or the president and the office of prime minister remained until 2018.

The algorithm is executed on two news websites, namely Hürriyet and CNN Türk. It records several key elements from the retrieved articles, including the combination of firm name and politician keywords, the main text of the article, headline, author, publication date, description, and word count.

The algorithm operates in two stages. In the initial stage, search queries are automatically generated based on the aforementioned logic. As an illustration, for a firm such as ”Ford Otomotiv Sanayi A.Ş.,” an example query would be ”Ford Otomotiv” + ”başbakan” (Ford Otomotiv + prime minister). Consequently, for each firm, three politician signifiers are combined and queried in two news websites, yielding a total of 870 firms (considering only those appearing in at least two time periods within the original dataset for the second top 500 companies) multiplied by three politician keywords and two news websites. This results in 5220 queries solely for approximately half of the entire dataset of top 1000 companies.

After filtering out articles falling outside the time range of the original dataset (2012-2021), a total of 5673 unique matches were retained, representing the intersection of firm-year-article combinations.

3 Methodology and Empirical Framework

In order to investigate the effects of unexpected changes in connection signals on the performance of industrial firms, I employ a fixed effects approach. I consider EBITDA (Earnings Before Interest, Taxes, Depreciation and Amortization) as the measure of a firm's economic performance. This methodology is particularly advantageous as it takes into account unobservable factors that remain constant over time and may influence the relationship between political connections and profitability. By utilizing a fixed effects model, I isolate the impact of connection shocks on business outcomes while controlling for these underlying factors. This approach allows for a more accurate assessment of the causal relationship between shifts in connections and their implications for industrial companies.

3.1 EBITDA as Outcome Variable

The choice of using EBITDA as the main measure of firm performance in the context of investigating the impact of political connections is supported by several academic reasons.

First, EBITDA is a widely recognized financial measure that provides a comprehensive assessment of a firm's operating performance. By excluding non-operating expenses such as interest, taxes, depreciation and amortization, EBITDA provides a clearer focus on a company's underlying operational efficiency and profitability. This measure is particularly relevant to the study of political connections, as it provides a more direct assessment of how such connections affect a firm's underlying financial performance.

Second, EBITDA serves as a valuable indicator for comparing and benchmarking firms across different sectors and regions. Its use in empirical research facilitates comparability of results and enables researchers to draw meaningful conclusions about the impact of political linkages on firm performance within and across industries.

EBITDA is also less sensitive to accounting practices and financial reporting biases than other performance measures such as net income or earnings per share. This makes it a robust measure to capture the true economic performance of firms, which is essential in assessing the impact of political connections on financial results.

Finally, the use of EBITDA is consistent with the objectives of many previous studies examining the relationship between political connections and firm performance. To study its impact on EBITDA, I consider three progressively more restrictive definitions of political connections which are discussed in the following subsection.

3.2 Sentiment Analysis Model

To construct the connection variable, I compiled a list of firms and political signifiers. This involved constructing a set of search queries for each entity on each news website and its search engine. News papers are clearly not politically objective and relying on one news paper alone might bias the measurement of connections in favor or against the government.

After retrieving the articles, I removed duplicates and calculated the sentiment score. The sentiment scores are calculated using a Recurrent Neural Network model. What makes it distinct is the fact that it can take into account both word embeddings and the ordering of the words when learning. I use a generic publicly available dataset in Turkish to train the sentiment analysis model.

I run a sentiment analysis on unlabelled news articles on Turkish firms and politicians. The purpose of this analysis is to label the news articles based on sentiment for the construction of a ”political connectedness” index which will depend on the co-occurrence of politician and firm names in news articles with positive sentiment.

I have two sources of data for this analysis. The first dataset is used for the training of the sentiment analysis model. It is a generic open source dataset that has been used in training similar models for Turkish language. The second dataset is created by me and contains information mainly on; when the article is published, from which politician and firm name combination it came from, word count, headline of the article, author, in addition to the article text itself.

I automatically generate and queried combinations of politican and firm names on Hürriyet and CNN Türk, websites of two large mainstream news agencies in Turkey. I then save the news articles that are returned from individual queries to a dataset.

The data to be used in model development has only two variables. The text and its sentiment score.

It consists of roughly 19 thousand short texts and their sentiment scores. Texts with negative or neutral sentiments are labelled ”0”, whereas the ones with positive sentiment are labelled ”1”. Although the texts in the training dataset are significantly shorter than target news articles, the properties of the RNN model and using a common vocabulary allows it to predict efficiently on unlabeled data. In the training data, around 70% of the texts are labelled positively.

I create a vocabulary from both the data for model development and the data that the model is later applied to and preprocess the news articles the same way I did with development data - Remove punctuation marks (commas, periods, exclamation marks, colons, question marks, parentheses, and double quotes) from the text, convert the text in the to lowercase characters only, remove any leading or trailing whitespace from the text, remove Turkish stopwords, tokenize into words, return the processed text.

An example text from the training dataset is as follows:

”gayet güzel ürün tavsiye ederim : 1”

The above sentence in Turkish reads ”Such a nice product. I recommend it.”, and its sentiment value is ”1”. It is properly preprocessed: No punctuation, no upper case letters, no stop words, no white space.

I continue with padding. Padding is necessary when working with sequence data in machine learning models, particularly deep learning models like RNNs and LSTMs (Long short-term memory) because these models require input sequences to have a consistent and fixed length. LSTM is a type of RNN that I employ in this analysis.

Padding is generally considered a necessary step when working with sequence data in deep learning models. It's crucial to choose an appropriate padding length (e.g., by using a length close to the mean sequence length or a length that covers most sequences in the dataset) to minimize the negative impacts (such as creating noise or loss of information) of padding on the model's performance.

I decide the padding length by adding to the mean number of tokens in a text two standard deviations of the distribution of all token lengths. Here the average number of tokens was 13.89, while the highest number was 207. However, with above logic I calculated the desired padding length at 44 which covers 96% of all tokenized texts.

Once, the text data is all processed and ready to train the sentiment analysis model, I initialize a sequential learner.

Long Short-Term Memory (LSTM) is a type of recurrent neural network (RNN) architecture designed to address the limitations of traditional RNNs, such as the vanishing gradient problem. LSTMs have been widely used in natural language processing, time series analysis, and other tasks that involve sequential data.

An LSTM layer can learn and remember patterns in sequences of data over long time steps, making it well-suited for handling tasks where the temporal dependencies (in my case word order dependencies) in the data are crucial. I add in total 3 LSTM layers, with 16, 8 and 4 neurons respectively. The last layer returns only the last hidden state in the output sequence. This is done to connect the LSTM layer to a Dense layer, as the Dense layer expects a single output for each input instance, not a sequence of outputs.

Finally, the Dense output layer with a single neuron takes the last hidden state from the last LSTM layer and produces an output between 0 and 1, which can be interpreted as the sentiment score (e.g., 0 for negative sentiment and 1 for positive sentiment) for the given input text.

The prediction accuracy of the model on the hold-out set (a random sample of 25% the entire training data which the model has not made any use of during the training process) is achieved

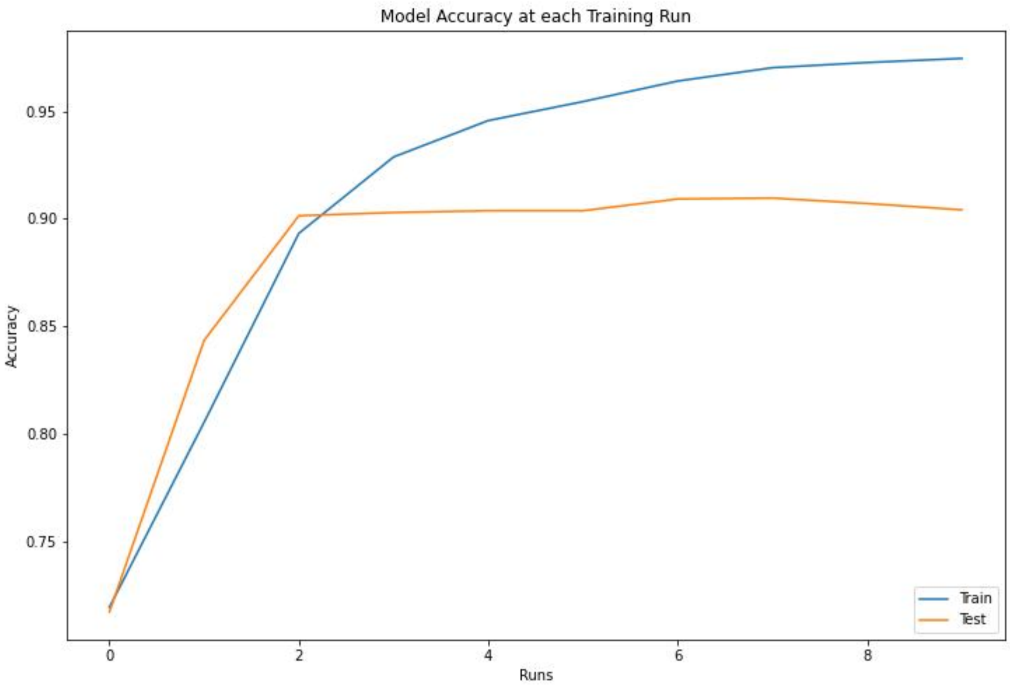

Figure 1: Sentiment Analysis Model Diagnostics

at roughly 90%.

As seen in Figure 1, the model performance on the test (hold-out) set peaks at around 90% which is an acceptable level for the use-goal of the model.

I then apply the trained model to the unlabeled news articles which it has never been trained on. Out of the original İSO dataset with top 1000 industrial firms over 10 years, 704 of them returned news articles in based on the algorithm outlined in the News Articles subsection. Out of those 704, 552 firms had at least one article matched with them with a positive sentiment value. This is the basis for connection variable definitions.

3.3 Connections

Based on the predicted sentiment values of news articles, three progressively more restrictive definitions of connections are made. The first, is the raw version. A firm-year observation can naturally be matched to not one but multiple news articles out of which some are predicted to

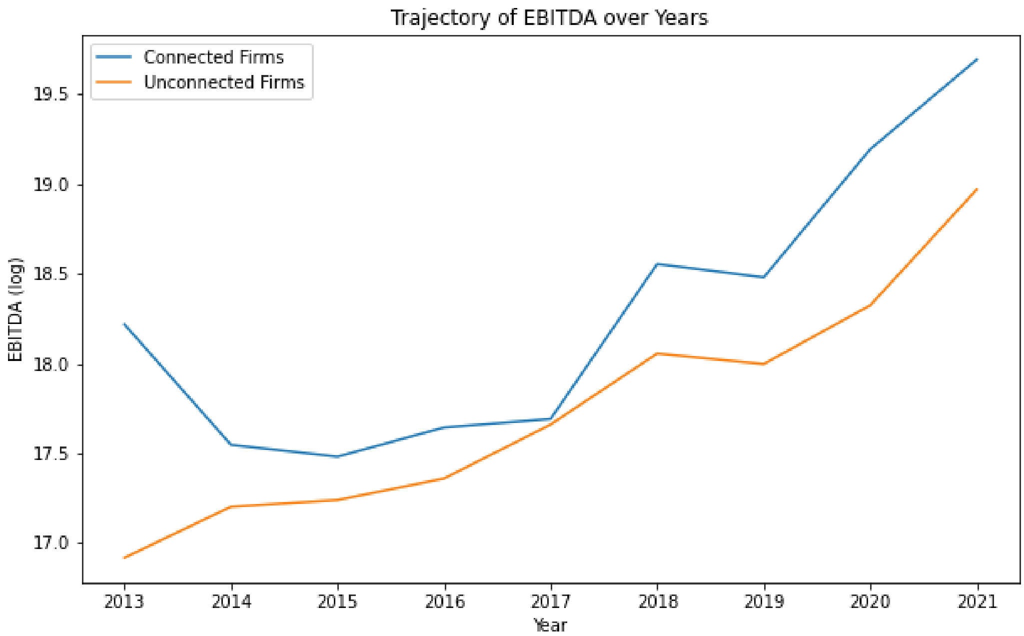

Figure 2: Trajectory of EBITDA (log) by Connectedness (FMC-1)

have a positive sentiment value. The raw version simply sums up the binary values of sentiment of articles that are matched to firm-year observation and assigns it its connection value.

The second definition, FMC-1, checks whether in each article a firm-year observation is matched with, the first word of the full firm name actually occurs in the news article (otherwise the result might just be returned due to the second, less significant word, of the full firm name). For each article the sentiment value and presence of the first word of the firm name are interacted. If either value is zero then the interaction is also zero. Finally, the results of the interactions are summed up and assigned as connection value to the firm-year observation.

The third definition, FMC-2, applies the same logic as in the second definition, but takes into account the first two words of the firm name occurring together. Therefore it is more restrictive than the first two definitions and yields weaker signals for political connections.

In the following section, I first check preliminary results with alternative definitions of connections, then continue further analysis with only one of them.

4 Results

In the following sections, I first run fixed effects regression on data from 2013 to 2021 with alternative definitions of connections. Then I choose one definition and run further regression with alternative specifications. In the next section, I conduct sensitivity analysis of the results in this section. It already evident from Figure 2 that, in its raw form, firms with connections, when they are connected are performing better in terms of EBITDA throughout the sample period which confirms the first hypothesis of this study.

4.1 Fixed Effects Estimation

Potential effects that are captured by the firm fixed effects are; the skill, overall sociability and connection building ability, the name value and risk of the firm, inherent corruptness of the managers, the ease of building and maintaining connections because of the physical proximity of the firm to the capital or the financial center İstanbul, attractiveness of certain industry sub-sectors to the politicians e.g. the automotive because people hold certain products at a higher esteem an so on. To isolate away such effects and test alternative definitions of connections, I estimate the following equation (in all of the following regressions entity and time fixed effects are controlled for):

$$\begin{aligned} EBITDA_{it} = \alpha + \beta_1 Connection_{it} + \\ \beta_2 Employees_{it} + \beta_3 Leverage_{it} + \gamma_i + \delta_t + u_{it} \quad (1) \end{aligned}$$

where $Employees_{it}$ is the average number of employees for firm i in year t, $Leverage_{it}$ is the debt-to-equity ratio for firm i in year t, γ_i are entity fixed effects, δ_t are time fixed effects, and u_{it} is the error term.

The regression results in Table 2 yield large, positive and statistically significant effect for the first two definitions of connections but not the last, most restrictive one. Although not all are statistically significant, all three definitions pointing at the same direction is promising. The coefficient on FMC-1, yields 84 million Turkish Liras (less than €4,000,000 by 2023 exchange rates) in additional profits for each additional FMC-1 score, which amounts to 12.5% of the standard deviation of EBITDA for the entire sample it was recorded and less than its median. Considering the plausibility of this definition, not too restrictive like the third definition, I choose the second definition, FMC-1, for further analyses. It is often the case that firms are mentioned solely by their first name let alone the full name or the first two names. It is also not too relaxed like first definition which does not consider at all whether the actual name of the firm in any form was actually mentioned in the news article in constructing the connection signal.

For the following regressions I estimate versions of the following, progressively adding more time-varying controls:

$$\begin{aligned} EBITDA_{it} = \alpha + \beta_1 Connection_{it} + \beta_2 Employees_{it} + \beta_3 Leverage_{it} \\ + \beta_4 SalesfromProduction_{it} + \gamma_i + \delta_t + u_{it} \end{aligned} \quad (2)$$

In Table 3, I add average number of employees first as a proxy for labor intensiveness of a firm over time (keep in mind that all the firms in the data engage in manufacturing). There have been economic and other shocks to how much labor a firm can employ in the 2013-2021 time frame. Most recent of which are the 2018 currency crisis and the 2020-2021 Covid-19 lock-downs. I then add exports as a measure of firm's opennes to trade which also varied heavily over time due to similar reasons and I presume that Export level can be correlated to both profits and connections.

	Connection Definition 1	Connection Definition 2	Connection Definition 3
Dep. Variable	EBITDA	EBITDA	EBITDA
Estimator	Fixed Effects	Fixed Effects	Fixed Effects
No. Observations	2485	2485	2485
R-squared	0.0513	0.0541	0.0452
R-Squared (Within)	0.0576	0.0604	0.0486
R-Squared (Between)	-0.0562	-0.0579	-0.0748
R-Squared (Overall)	-0.0275	-0.0281	-0.0508
F-statistic	33.621	35.588	29.467
P-value (F-stat)	0.0000	0.0000	0.0000
=====================	============	============	============
Aggregated Sentiment Score	6.537e+07 (3.6390)		
Avg. Number of Employees	-3.814e+04 (-1.5378)	-3.852e+04 (-1.5557)	-3.909e+04 (-1.5712)
Leverage Ratio	7.484e+06 (9.2259)	7.497e+06 (9.2558)	7.499e+06 (9.2146)
FMC-1		8.437e+07 (4.3452)	
FMC-2			5.797e+07 (1.1560)
=======================	==============	==============	==============
Effects	Entity Time	Entity Time	Entity Time

Table 2: Comparison of Alternative Definitions of Connections
(T-stats reported in parentheses)

A firm engaging in exports more intensely might be favored in scope of the country's trade policy or might be less willing to spend resources on building domestic political connections as they are more focused outwardly. However, the statistical significance or the economic size do not change materially by its addition. I further add, leverage ratio, defined as debt-to-equity ratio, as a measure of a firm's time-varying risk-appetite which for reasons similar to above has seen several significant shocks in companies' balance sheets. Finally, and most importantly, I add sales from production as measure of a firm's productive capacity or a flow measure of its size in nominal terms. In the absence of a solid measure of company size, one can easily argue that larger sales or bigger firms bring not only higher profits but also political connections due to scale effects. Larger firms hit more regulatory barriers than smaller ones, they are more prevalent, they arguably attract politicians easily and so on. Regardless, the coefficient on FMC-1 remains economically relevant, positive, and statistically significant at almost a 99% confidence level.

One issue so far that has not been touched upon is the treatment of standard errors. The sample at hand is naturally far from random and represents large industrial firms, which is the main population make inferences on, rather than firms from any sector of any size. Therefore, clustering standard errors is called for. In the following, section I test the robustness of the statistical significance of the results of estimating equation (2) on Table 3 (Model 4).

5 Robustness

I run the exact same specification as in equation (2), only this time with clustered standard errors. I run the first with one-way clustered errors (firm-level), the second with two-way (firm-year-level), and the third intuitively with industry level clusters. See Figure 3 for the distribution of firms among industries in 2021 and Figure 4 for ISIC industry definitions.

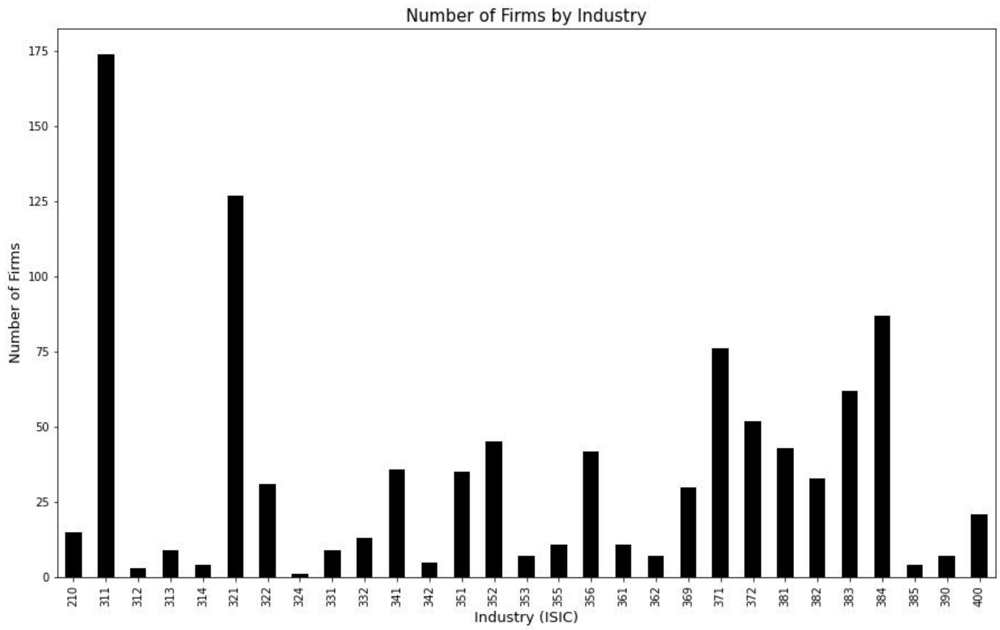

Figure 3: Number of Firms by Industry (2021)

The sample under study is naturally not one with independent and identically distributed (i.i.d) observations by definition. However, in the presence of heteroscedasticity or correlation within groups (firms, firm-years, or industries), the standard errors may be biased, inflating the statistical significance of inferences.

When clustering standard errors at the industry level, I aim to adjust standard errors to account for potential correlation of observations within the same industry. This clustering approach recognizes that firms within the same industry may share certain common characteristics or experiences that can lead to correlated errors i.e., time-varying unobserved factors that are relevant in estimating the impact of political connections on profitability.

By clustering standard errors at the industry level, my goal is to obtain more accurate standard errors that appropriately reflect the clustering structure within the data. This clustering adjustment is particularly relevant because observations within the same industry are likely to be more similar to each other than to observations in different industries.

By clustering standard errors at the industry level, my goal is to obtain more accurate standard errors that appropriately reflect the clustering structure within the data. This clustering adjustment is particularly relevant because observations within the same industry are likely to be more similar to each other than to observations in different industries.

Larger standard error estimates compared to standard errors calculated without clustering arise because the clustering adjustment accounts for the additional variability within industries. Consequently, reducing the statistical significance of the results.

Section	Divisions	Description
A	01–03	Agriculture, forestry and fishing
B	05–09	Mining and quarrying
C	10–33	Manufacturing
D	35	Electricity, gas, steam and air conditioning supply
E	36–39	Water supply; sewerage, waste management and remediation activities
F	41–43	Construction
G	45–47	Wholesale and retail trade; repair of motor vehicles and motorcycles
H	49–53	Transportation and storage

Figure 4: ISIC Industry Categories

Additionally, by clustering standard errors at the industry level, I acknowledge the potential heterogeneity across industries, which can affect the interpretation and generalizability of the findings.

	Model 1	Model 2	Model 3	Model 4
Dep. Variable	EBITDA	EBITDA	EBITDA	EBITDA
Estimator	Fixed Effects	Fixed Effects	Fixed Effects	Fixed Effects
No. Observations	2520	2272	2244	2244
R-squared	0.0106	0.1236	0.1546	0.6142
R-Squared (Within)	0.0119	0.1350	0.1683	0.6382
R-Squared (Between)	-0.0750	0.2842	0.2921	0.5873
R-Squared (Overall)	-0.0682	0.2198	0.2565	0.5839
F-statistic	10.204	79.256	75.962	528.54
P-value (F-stat)	0.0000	0.0000	0.0000	0.0000
=====================	=============	=============	=============	=============
FMC-1	8.336e+07	9.414e+07	9.921e+07	3.722e+07
	(4.2546)	(4.0447)	(4.2726)	(2.3630)
Avg. Number of Employees	-3.765e+04	-9.109e+04	-9.128e+04	-6.772e+04
	(-1.5012)	(-3.3258)	(-3.3691)	(-3.6975)
Exports (thousand $s)		1888.0	1761.8	-160.23
		(14.275)	(13.349)	(-1.6167)
Leverage Ratio			6.322e+06	-2.274e+06
			(7.7778)	(-3.9058)
Sales from Production				0.1533
				(44.467)
=======================	==============	==============	==============	==============
Effects	Entity Time	Entity Time	Entity Time	Entity Time

Table 3: Fixed Effects Models with FMC-1
(T-stats reported in parentheses)

In the last section, I acknowledge heterogeneity and sample selection related limitations, re-emphasize the place of this study in the literatures of political economy of connections, and alternative data sources and computational methods for econometric analysis (Athey & Imbens 2019, Gentzkow et al 2019, Mullainathan et al. 2017), and make suggestions for further research.

6 Conclusion

This study contributes to the literature on the role of political connections in business performance by providing new evidence and a novel methodology specific to the Turkish context. The findings suggest a robust positive correlation between political connections and business performance; however, establishing a robust causal relationship remains challenging. This limitation could be attributed to challenges in accurately measuring political ties and the lack of comprehensive data. Therefore, obtaining a larger set of observations and developing a more sophisticated methodology that captures even weaker media signals could potentially yield more robust results.

Moving forward, future research could explore alternative methods of measuring political connections and investigate the potential mechanisms through which these connections affect business performance. Additionally, testing various channels through which political connections influence firm outcomes would provide a deeper understanding of the underlying dynamics.

When examining the impact of political connections on firm performance using a sample composed of the top 1000 industrial firms based on sales from production, it is crucial to consider several issues in inference. These include the potential for sample bias, limited variability, reduced generalizability, and the risk of omitting important variables.

While the focus on large firms is justified given their relevance to the phenomenon of political connections, caution must be exercised when generalizing the findings to the broader population of industrial firms or different industries.

	Entity Level Clustering	Entity-Time Level Clustering	Industry Level Clustering
Dep. Variable	EBITDA	EBITDA	EBITDA
Estimator	Fixed Effects	Fixed Effects	Fixed Effects
No. Observations	2244	2244	2244
Cov. Est.	Clustered	Clustered	Clustered
R-squared	0.6142	0.6142	0.6142
R-Squared (Within)	0.6382	0.6382	0.6382
R-Squared (Between)	0.5873	0.5873	0.5873
R-Squared (Overall)	0.5839	0.5839	0.5839
F-statistic	528.54	528.54	528.54
P-value (F-stat)	0.0000	0.0000	0.0000
=====================	============	============	============
FMC-1	3.722e+07	3.722e+07	3.722e+07
	(0.9304)	(1.0662)	(0.8872)
Avg. Number of Employees	-6.772e+04	-6.772e+04	-6.772e+04
	(-0.7671)	(-1.0639)	(-0.7619)
Exports (thousand $s)	-160.23	-160.23	-160.23
	(-0.5266)	(-0.4529)	(-0.5216)
Leverage Ratio	-2.274e+06	-2.274e+06	-2.274e+06
	(-1.2588)	(-0.7910)	(-0.9797)
Sales from Production	0.1533	0.1533	0.1533
	(3.6793)	(3.6676)	(3.3684)
=======================	==============	==============	==============
Effects	Entity Time	Entity Time	Entity Time

Table 4: Robustness to Different Levels of Clustering
(T-stats reported in parentheses)

Despite these limitations, the study highlights a significant and economically meaningful relationship between political connections and profitability. These findings contribute to our understanding of the impact of political connections on the performance of large industrial firms. Moreover, the study extends the research agenda by showcasing the utility of alternative data formats, such as media data, in econometric analysis. The incorporation of computational methods and natural language processing techniques allows for the extraction and analysis of valuable information from textual sources, thereby enhancing the rigor of examining the relationship between political connections and business performance.

In conclusion, this research demonstrates the promising potential of incorporating media data into econometric analysis, particularly in the study of political connections and firm performance. By utilizing computational methods and exploring alternative data sources, a broader range of research questions can be addressed with greater rigor. However, it is crucial to acknowledge the limitations associated with the sample selection and the need for further research to validate and extend these conclusions to other contexts.

Overall, this study contributes valuable insights into the relationship between political connections and firm performance within the selected sample of top industrial firms. However, future research should aim to generalize these findings to a broader population and consider additional factors that may shape firm performance. By addressing these aspects, we can gain a more comprehensive understanding of the complex dynamics between political connections and business outcomes.

Bibliography

[1] Athey, S., & Imbens, G. W. (2019). Machine Learning Methods That Economists Should Know About. *Annual Review of Economics*, *11*(1), 685–725. https://-doi.org/10.1146/annurev-economics-080217-053433

[2] Cruz, C., Labonne, J., & Querubín, P. (2017). Politician Family Networks and Electoral Outcomes: Evidence from the Philippines. *American Economic Review*, *107*(10), 3006–3037. https://doi.org/10.1257/aer.20150343

[3] Enikolopov, R., Petrova, M., & Sonin, K. (2018). Social Media and Corruption. *American Economic Journal: Applied Economics*, *10*(1), 150–174. https://-doi.org/10.1257/app.20160089

[4] Ferguson, T., & Voth, H.-J. (2008). Betting on Hitler: The Value of Political Connections in Nazi Germany. *The Quarterly Journal of Economics*, *123*(1), 101–137.

[5] Fisman, R. (2001). Estimating the Value of Political Connections. *The American Economic Review*, *91*(4), 1095–1102.

[6] Fisman, R., & Wang, Y. (2015). The Mortality Cost of Political Connections. *The Review of Economic Studies*, *82*(4), 1346–1382.

[7] Gentzkow, M., Kelly, B., & Taddy, M. (2019). Text as Data. *Journal of Economic Literature*, *57*(3), 535–574. https://doi.org/10.1257/jel.20181020

[8] Gürakar, E. Ç., & Bircan, T. (2016). Political Connections and Public Procurement in Turkey: Evidence from Construction Work Contracts.

[9] İstanbul Chamber of Industry. (2022). İstanbul Sanayi Odası Tanıtım Kitapçığı.

[10] İstanbul Chamber of Industry. (2021). Türkiye'nin 500 Büyük Sanayi Kuruluşu. Retrieved from https://www.iso500.org.tr/500-buyuk-sanayi-kurulusu

[11] Li, H., Meng, L., Wang, Q., & Zhou, L. (2008). Political connections, financing and firm performance: Evidence from Chinese private firms. *Journal of Development Economics*, *87*(2), 283–299.

[12] Mullainathan, S., & Spiess, J. (2017). Machine Learning: An Applied Econometric Approach. *Journal of Economic Perspectives*, *31*(2), 87

–106. https://doi.org/10.1257/jep.31.2.87

[13] Muttakin, M. B., Monem, R. M., Khan, A., & Subramaniam, N. (2015). Family firms, firm performance and political connections: Evidence from Bangladesh. *Journal of Contemporary Accounting & Economics*, *11*(3), 215–230. https://doi.org/10.1016/j.jcae.2015.06.001

[14] Önder, Z., & Özyıldırım, S. (2011). Political Connection, Bank Credits and Growth: Evidence from Turkey. *The World Economy*, *34*(6), 1042–1065. https://doi.org/10.1111/j.1467-9701.2011.01342.x

[15] Özcan, G., & Gündüz, U. (2015). Political connectedness and business performance: Evidence from Turkish industry rankings. Business and Politics, 17(1), 41-73. doi:10.1515/bap-2013-0037

[16] TurkStat. (2022). Yıllık Sanayi ve Hizmet İstatistikleri, 2021. Retrieved from https://-data.tuik.gov.tr/Bulten/Index?p=Yillik-Sanayi-ve-Hizmet-Istatistikleri-2021-45836

[17] Xu, G. (2018). The Costs of Patronage: Evidence from the British Empire. *American Economic Review*, *108*(11), 3170–3198. DOI: 10.1257/aer.20171339