Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Economics and Management
Department of Economics

Thea Heiene
Oda Heggedal Longvastøl

Lobbying, Finance and Inequality

An empirical study of financial sector lobbying using the EU Transparency Register

Master's thesis in Economics
Supervisor: Costanza Biavaschi
June 2019

Photo by Maryna Yazbeck on Unsplash

NTNU
Kunnskap for en bedre verden

Thea Heiene
Oda Heggedal Longvastøl

Lobbying, Finance and Inequality

An empirical study of financial sector lobbying using the EU Transparency Register

Master's thesis in Economics
Supervisor: Costanza Biavaschi
June 2019

Norwegian University of Science and Technology
Faculty of Economics and Management
Department of Economics

NTNU
Norwegian University of
Science and Technology

Preface

Finishing this master's thesis has been a milestone. It concludes five years as students at NTNU. It has been a true joy to share this time together, not only as fellow students, but also as cabin-mates, friends and hiking companions.

Working on our thesis has been challenging, fun, interesting and sometimes frustrating. We are truly grateful to our supervisor Costanza Biavaschi for her guidance, support and good, professional discussions. Her positive spirit and enthusiasm for our work has given us the confidence and motivation we needed to create and carry out our own project.

We would also like to thank Håkon Måløy (NTNU), Kenneth Haarr and Vicky Cann (Coorporate Europe Observatory) for helping us retrieve the data we needed.

Further we want to thank our friends and family for your extra support and caring during this period. Thank you Siw Heggedal Longvastøl, Ivar Myklebust Longvastøl, Rypereiret, Hilmar Nypan Claes and Eystein Skjerve.

It can be challenging to find the motivation to complete a large project such as this. But as Oda puts it: "Berre gjer det, og drit i å klag[1]".

[1]Translation from Nynorsk: "Just do it, and stop complaining"

Abstract

Lobbying by special interest groups is an integrated part of EU legislation and decision making processes. This master's thesis investigates the relationship between the political influence of the financial sector and income inequality in a selection of European countries. Using data from the EU Transparency Register, we construct a unique data set on lobbying and inequality. We use lobbying expenses as a measure of financial sector influence and the Gini index to measure income inequality. Pooled OLS and Fixed effects estimation show that an increase in lobbying expenses is associated with a higher level of income inequality. We also attempt to use an IV strategy to correct for possible endogeneity issues.

As robustness checks, we estimate the model using different measures of inequality and a reduced sample. The small, positive effect is qualitatively the same across different estimation methods and subsamples, implying that successful lobbying might result in a transfer of wealth towards influential interest groups and lead to increased income inequality.

Samandrag

Lobbyisme frå interessegrupper er ein integrert del av EU si lovgjeving- og avgjerdsprosess. I denne mastergradoppgåva undersøkjer vi forholdet mellom den politiske påverknaden til finansielle sektorar og grada av økonomisk ulikskap i eit utval europeiske land. Vi nyttar data frå EU Transparency Register til å lage eit makelaust datasett for lobbyverksemd og økonomisk ulikskap. For å måle finanssektoren sin påverknad, ser vi på kostnadar brukt på lobbyisme. For å måle grada av ulikskap i inntekter, nyttar vi Gini-indeksen. Estimering med minste kvadrats metode med og utan faste effektar syner at ei auke i kostnadar til lobbyverksemd er knytt til større grad av ulikskap i inntekter. Vi nyttar også IV-estimering i eit forsøk på å justere for moglege endogenitetsproblem.

For å teste for robustheit estimerar vi den fulle modellen ved å nytte forskjellige mål på ulikskap og eit minska utval. Vi finn ei lita, positiv effekt som er kvalitativt den same på tvers av ulike estimeringsmetodar og utval. Dette tyder at vellukka lobbyverksemd kan føre til overføring av verdiar til innflytelsesrike interessegrupper, og såleis auka økonomisk ulikskap.

Contents

1 Introduction

Lobbying, or the representation of special interest groups plays a central part in the legislative process in the European Union (Klüver, 2012). It brings together the interests of member states, business, NGOs, trade unions and other organizations in the decision making process. As stated by the European Council, business interest constitutes about 80% of interest groups that lobby in the EU (The General Secretariat of the Council, 2013). When it comes to lobbying on financial sector issues, Wolf, Haar, and Hoedeman (2014) find that financial sector organizations largely outnumber public authorities, NGOs and other groups. They are disproportionately represented in terms of lobbying expenses, number of lobbyists and number of lobbying organizations (Wolf et al., 2014, pp. 11-12).

In this thesis we aim to quantify lobbying effort of the financial sector in a selection of European countries. The goal is to study whether the influence of the financial sector affects the income distribution within a country. We combine insights from the research on finance, inequality and rent seeking with a unique data set on lobbying and income inequality to study the aggregate economic consequences of financial sector lobbying.

We have constructed a data set combining lobbying data from the EU Transparency Register and inequality data from the Eurostat database. As a proxy variable for influence, we use lobbying expenses of financial sector organizations by country and year. We use the Gini index to measure income inequality. Our panel data set contains a sample of 29 European countries over the period 2010-2017. When controlling for other factors that might affect inequality, we find a small, positive effect of financial sector lobbying on income inequality. The causal impact is estimated using pooled OLS and fixed effects estimation. We do robustness checks using different measures of inequality and with a limited subsample. The magnitude and sign of the estimated effect is found to be qualitatively similar across different estimation methods and model specifications. We also attempt at an instrumental variable strategy, where we use the share of population working in commercial banks in 1995 to construct an instrument.

According to McCormick and Tollison (1981) (as cited in Mueller, 2003, p. 347), all

legislations consists of a transfer of wealth. Lobbying on financial regulation can be seen as a way of special interest groups to capture the wealth, or "rents", following a certain legislation. This suggests that there are distributional consequences of lobbying activity. If rents are transferred to a certain group as a result of successful lobbying effort, it should be possible to observe the impact it has on the economy. Recent research shows that wealth and income inequality has increased in Europe since 1970 (Piketty & Zucman, 2014), and inequality has been put on the political agenda. Our hypothesis is that the degree of political influence of the financial sector can help explain change in the income distribution within a country.

Our research complements the literature on finance, growth and inequality, combining it with insights from the economics of special interest groups and rent seeking theory. With this thesis, we hope to shed new light on one aspect of how the financial sector affects distribution of income.

The contribution of our thesis is twofold. Firstly, it represents an approach to studying the functioning of the financial sector in the economy that differs from similar research. While most literature regarding finance, growth and inequality looks at financial development measured as private credit to GDP (see Levine, 2005), we attempt to capture a different part of financial sector activity. We try to measure the influence, rather than the size of the financial sector.

Secondly, it contributes to the research on lobbying and rent seeking. By constructing our own data set, we exploit the information in the Transparency Register to study the effects of lobbying empirically. Similar use of lobby registers can be found in research from the US, as done by Bertrand, Bombardini, and Trebbi (2014) and Igan, Mishra, and Tressel (2012). To the best of our knowledge, there has not been much economic research doing the same for European countries. The decisions made in the EU affect the lives of all its inhabitants, and to assure democratic legitimacy, the participation of interest groups has been encouraged lately (Klüver, 2012). There has been an increased focus on transparency in the decision making process, resulting in for instance the Transparency Register (European Commission, 2016). The Register is still in its early years, and we hope that our attempt to quantify lobbying activity can inspire future use of the information available in the register.

The thesis is organized in the following way: In Chapter 2 we review relevant literature

in the areas of finance, growth and inequality and the economics of special interest groups. In Chapter 3 we describe the data collection process and our key variables. In Chapter 4 we explain the estimation methods and the identifying assumptions for pinning down the causal relationship between lobbying and inequality. The main findings are presented in Chapter 5. In Chapter 6 we present robustness checks and discuss possible limitations of our method and data set. In Chapter 7 we come with suggestions for further research on the topic. A conclusion is provided in Chapter 8.

2 Literature Review

To answer our research question, we combine insights from studies on finance, growth and inequality with research on the economics of special interest groups and rent seeking. In this chapter, we give a short review of the existing literature and empirical research that we build our analysis upon.

2.1 Finance, Growth and Inequality

There is a large body of literature studying the role of the financial sector in the economy. The financial sector provides important functions for the functioning of the modern economy[2] (Levine, 2005). Having a well-functioning financial sector can be growth enhancing and be important for the economy as a whole (Arcand, Berkes, & Panizza, 2015). Levine (2005) presents a meta-analysis of the research on the connection between financial sector development and growth. The main findings of the literature is that increased financial development has a positive effect on economic growth. After summarizing research on the finance-growth nexus, he presents theory and evidence on the link between financial sector development and inequality. Two contradictory effects are discussed: on one hand, a well-functioning financial system might allocate capital or credit to its most efficient use. This could allow credit to be allocated to its best uses, regardless of economic background. Credit can be given to people who would else be without access to finance. This can create a larger benefit for the poor than for the rich, and lead to a decrease in inequality (Levine, 2005, p. 887). On the other hand, the poor might face credit restrictions, leading to a productive part of society being excluded from the benefits of financial development (Levine, 2005, p. 887), and hence increased inequality.

Empirical evidence of the tendency of financial development to decrease income inequality is presented in a working paper by Beck, Demirguc-Kunt, and Levine (2004). Using the average of different inequality measures[3] and private credit to GDP ratio

[2]Such as providing information about investments and allocate capital efficiently, to provide finance and monitor investments risk trading and management and risk allocation, savings pooling and mobilizing and easing the exchange of goods and services (Levine, 2005, p. 869)

[3]Such as the Gini coefficient, per capita income of the poor and growth in population living

as a measure of financial development, they find evidence that income inequality decreases as financial development increases. Levine (2005, p.921) points out that this analysis might be methodologically weak and calls for more diverse methodology and data sets to investigate the issue further.

An improvement over previous work is more recently provided by Jauch and Watzka (2016). Using data on 168 countries over 48 years they find a positive effect of financial development on the Gini coefficient. This implies that increased financial development increases income inequality within a country.

We build upon the literature on finance, growth and inequality with regard to the econometric methods and what variables we control for. Our research differs from other literature because we use lobbying expenses as a measure of financial sector influence, instead of measuring the size of the financial sector. The contribution of this thesis lies closer to the finance and inequality literature. Because this literature is limited, we have also leaned on the research on finance and growth. The relationship between finance and growth has been more thoroughly studied than the link between finance and inequality, and it provides more robust and well-tested results. The literature on finance and growth provides a more extensive theoretical and methodological framework than the studies on finance and inequality.

2.2 Lobbying and rent seeking

Given our focus on financial sector lobbying rather than financial development, we present a review of the literature on the economics of special interest groups and lobbying. We consider two main purposes of lobbying activity. One view is that the role of lobbyists is to provide information and expertise to decision-makers who do not have the time to specialise in all legislation processes they take part in (Bertrand et al., 2014). The other view is that of lobbying as a rent seeking activity. McCormick and Tollison (1981) (as cited in Mueller, 2003, p. 347) consider all legislation processes as a transfer of wealth. The redistribution of this wealth depends on which interest group that is most successful in influencing the legislative process (Mueller, 2003, p. 347). If so, one can consider lobbying as a way of a special interest group to

under the poverty line.

try and receive part of the wealth that is to be redistributed. Igan et al. (2012, p. 198) define lobbying as "legal activity aiming at changing existing rules of policy or procuring individual benefits".

Igan et al. (2012) find that mortgage lenders in the US that lobbied on mortgage-market-specific issues had higher loan-to-income ratios, higher degree of securitization and more rapidly growing portfolios in the period prior to the financial crisis of 2007-2008. They argue that lobbying companies that profited on laxer regulations were more prone to lobby on the same regulations. This suggests that the behaviour of lobbying lenders might have contributed to the accumulation of risk that culminated in the financial crisis. As this example suggests, the influence and actions of financial organizations might have consequences for the economy as a whole.

It is not always obvious whether lobbying activity is aimed at information and expertise sharing or at rent seeking. We are aware of this caveat, and are not trying to undermine the role that lobbyists play as information providers and experts in the EU. As we are not able to distinguish between the two functions, the focus in this paper is on lobbying as a rent seeking activity. Our choice of lobbying expenses as a measure of influence is an attempt to capture the part of financial sector activity aimed at rent seeking. We study whether lobbying from the financial sector leads to an observable distortion of the income distribution within countries.

2.3 Summary

Earlier research on financial sector development discusses the functioning of the financial sector in the economy. The literature suggests that there is a positive effect of increased financial development on economic growth, while the effect on inequality is more indecisive. It is an area where more research is needed in order to discover the possible causal effect.

This thesis complements the literature on finance, growth and inequality and combines it with insights from the literature on the economics of special interest groups. We focus on lobbying as a rent seeking activity and investigate the causal link between the influence of the financial sector and income inequality.

3 Data Description

A key contribution of this thesis is the construction of a new data set. Using the EU Transparency Register and the Eurostat database we constructed an unbalanced panel data set on lobbying and inequality. It consists of 29 countries from 2010 to 2017, with a total of 169 country-year observations. In the following chapter, we describe the data on lobbying activity, income inequality and control variables.

3.1 Influence

We measure financial sector influence by country and year as lobbying expenses of the financial sector. The lobbying expenses were obtained using the EU Transparency Register. It is our attempt to find a measure that captures the part of financial activity aimed at rent seeking activity through influencing laws and politics. The register contains organizations that lobby the European Commission and the European Parliament. It provides information about each organization and their lobbying activity. Registration in the Transparency Register is required to get access to certain advantages as a lobbyist at the EU level, such as access to meetings with high-level decision makers and speaking at public hearings (European Commission, 2019). However, it is not mandatory if one does not seek these advantages. A description of the process of retrieving the data and our key explanatory variable follows next.

The first step was to select a list of financial sector companies and organizations. We used a web-scraping algorithm[4] which provided detailed information about all organizations in the Transparency Register [5]. Using string analysis we were able to identify all organizations that had names, goals or issues of interest that contained specific finance-related keywords[6]. We manually went through these organizations to make sure that we only had financial companies and interest organizations[7] in our sample. To extend our data set, we supplemented this list with companies that were

[4]We are grateful to Håkon Måløy for help with the web-scraping.

[5]It includes 11,337 registrations in the Transparency Register as they were on 28th January 2019

[6]We searched for "bank"/"banc"/"banq", "financ", "MiFID" (Markets in Financial Instruments Directive"), "CRD" (Capital Requirement Directives", "securities" and "credit"

[7]Mainly banks, insurance companies, wealth management firms, interest organizations and law- and consultancy firms that have financial issues as their main area of interest

used in a 2014 report about financial sector lobbying in the EU (see Wolf et al., 2014).

The next step was to trace the lobbying expenses of these organizations over time. This was done by downloading yearly data from Lobbyfacts (Corporate Europe Observatory & LobbyControl, n.d.), a database containing historical data from the Transparency Register[8]. This enabled us to construct a data set containing the lobbying expenses of all 617 companies as far back as 2010[9].

The final step was to create a measure of lobbying expenses for each country and year. We summed up lobbying expenses of organizations with head quarters in the same countries for each year. We use the lower reported expenses[10], meaning that the numbers are likely to be underestimating the actual lobbying expenses. The average lobbying expenses by country and year is €4.85 million, as can be seen in Table 1. The between standard deviation is 8.2, while the within standard deviation is 3.84. This indicates that the lobbying variable has more variation between countries, than over time within the same countries. We also keep track of the number of lobbying organizations for each country-year observation. On average, there are 15.09 lobbying organizations for each country each year, with an overall standard deviation of 23.09.

[8]We downloaded the registrations of all the organizations from the Lobbyfacts.eu from the 17th of February of every year from 2012-2019.

[9]Because of the nature of the registration process, the lobbying expenses for a year is reported with a 2-year lag. For this reason, we have lagged all observations by 2 years, meaning that 2012 entries are counted as 2010 expenses in our data set. 2013 is counted as 2011 and so forth.

[10]If an organization has reported that they spend €50, 000-€100,000 per year on lobbying in the EU, we have used €50,000.

Table 1: Summary statistics of lobbying variables

Variable	Panel	Mean	Sd	Min	Max	Obs.
Lobbying expenses	Overall	4.85	9.92	0	48.47	N = 169
	Between		8.20	0	38.09	n = 29
	Within		3.84	-13.62	20.66	$\overline{\text{T}}$ = 5.83
Organizations	Overall	15.09	23.09	1	126	N = 169
	Between		18.93	1	69.38	n = 29
	Within		10.16	-34.28	71.72	$\overline{\text{T}}$ = 5.83

Notes: Lobbying expenses is the total yearly spending on lobbying (in €1,000,000) by financial sector organizations from the same country. *Organizations* is the number of financial sector organizations in the EU transparency register for each country-year observation.
N is the total number of observations. n is the number of countries. $\overline{T}$ is the average number of years observed for each country.

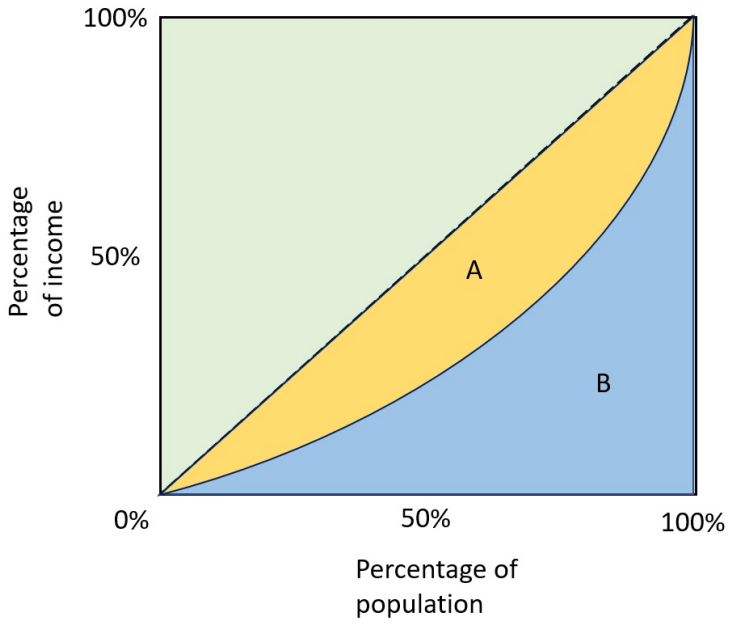
3.2 Inequality

To measure inequality, we use the Gini index of equivalised disposable income[11] from the EU Survey of Income and Living Condition (Eurostat, 2017). The Gini index is a number between 0 and 100, where 0 is perfect income equality, and 100 is perfect inequality.

Figure 1 shows the different elements of the Gini index. The convex line in the figure is the Lorenz curve and it represents the distribution of income within a country. Each point along the Lorenz curve represents the percentage of total income earned by a certain percentage of the population. The diagonal line represents a (hypothetical) line of perfect equality, where 10% of the population earns 10% of the income, 50% earns 50% and so forth. In other words, in a country with a diagonal Lorenz curve, income is identically distributed within the population. The Gini index is computed

[11] Total disposable income (after taxes and transfers) of a household divided by number of household members.

Figure 1: Graphic presentation of the Gini index



Notes: The convex line is the Lorenz curve, representing actual distribution of income. The diagonal line represents a line of perfect income equality. The Gini index is equal to the area A, divided by the total area beneath the equality line, A+B.

by dividing the area between the Lorenz curve and equality line, here seen as area A, by the total area beneath the equality line, area $A + B$, and multiplied by 100.

$$Gini = \frac{A}{A + B} \cdot 100 \tag{1}$$

A higher Gini index indicates higher degree of income inequality.

From Table 2 we can see that the average (overall) Gini index is 29.21, with an overall standard deviation of 3.20 and within standard deviation of 0.69.

Table 2: Summary statistics of inequality variables

Variable	Panel	Mean	Sd	Min	Max	Obs.
Gini index	Overall	29.21	3.20	23.2	40.2	N = 167
(Eurostat estimate)	Between		3.93	24.1	40.2	n = 29
	Within		.69	27.18	31.18	$\overline{\text{T}}$ = 5.76
Gini index	Overall	31.17	3.07	25.4	36.3	N = 112
(World Bank estimate)	Between		3.28	25.55	35.92	n = 26
	Within		.55	29.35	33.15	$\overline{\text{T}}$ = 4.31
S80/S20 ratio	Overall	4.63	.91	3.3	8.2	N = 167
	Between		1.15	3.3	8.2	n = 29
	Within		.18	4.09	5.19	$\overline{\text{T}}$ = 5.76

Notes: The Gini coefficient is a measure of income inequality. It is a number between 0 and 100, where 0 represents perfect income equality and 100 is perfect inequality.
S80/s20 is the ratio between the income of the 20% with highest income and the 20% with lowest income.
N is the total number of observations. n is the number of countries. $\overline{T}$ is the average number of years observed for each country.

For robustness checks, we estimate the model using different measures of inequality. We use the Gini index from the World Development Indicators (World Bank, n.d.) and the S80/S20 ratio from Eurostat (2017).

The World Bank estimates the Gini index using different sources and computational method, hence there is some discrepancy between the two measurement. The World Bank estimate is generally larger than the Eurostat estimate. It has a mean of 31.17 in our sample.

The S80/S20 ratio is the total income earned by the 20% of the population with highest income divided by the total income earned by the 20% with lowest income. The overall average in our sample is 4.63, meaning that on average, the 20% with

highest income earn almost five times as much as the 20% with lowest income. The S80/S20 ratio ranges from 3.3 to 8.2 in our sample[12].

For both Eurostat indicators we have 29 countries in our sample, running from 2010-2017 with some missing observations. The number of countries and average number of time periods ($\overline{T}$) is smaller for the World Bank Gini. There is more variability between countries than within for all variables, as can be seen by comparing the between and within standard deviations showed in Table 2.

3.3 Control Variables

Following the literature (see Arcand et al., 2015; Jauch & Watzka, 2016) we control for financial development, government consumption as a share of GDP and inflation.

Financial development is measured as the private credit credit to GDP ratio. Private credit includes all financial resources provided to the private sector by financial corporations. From Table 3 we see that on average, private credit makes up 104.26% of GDP in our sample.

To control for other economic factors, we use government consumption as a share of GDP and inflation. Government consumption includes all public expenses on goods and services, excluding military expenses. It has an average of 20.20% in the sample. Inflation is measured as the yearly percentage change in the price of buying a specific "basket" of goods and has an average of 1.11%.

In addition to this, we include population size as a control variable, to account for the fact that a country with a larger population might have a larger finance sector and hence larger lobbying expenses in absolute numbers. The average population size is 21.75 million inhabitants.

For more detailed information about our control variables, see Appendix A.1 and A.2.

[12] 3.3 for Iceland in 2014, 8.2 for Bulgaria 2017.

Table 3: Summary statistics of control variables

Variable	Panel	Mean	Sd	Min	Max	Obs.
Private credit	Overall	104.26	43.44	33.37	250.01	N = 168
(% of GDP)	Between		45.41	34.43	229.49	n = 29
	Within		11.95	62.41	151.20	$\overline{\text{T}}$ = 5.79
Gov. consumption	Overall	20.20	3.69	11.6	27.4	N = 169
(% of GDP)	Between		3.39	11.88	26.04	n = 29
	Within		.83	16.66	23.46	$\overline{\text{T}}$ = 5.83
Inflation	Overall	1.11	1.27	-2.10	4.24	N = 169
(%)	Between		.91	-1.15	2.93	n = 29
	Within		1.02	-1.26	3.85	$\overline{\text{T}}$ = 5.83
Population	Overall	21.76	25.24	.34	82.70	N = 169
(In mill.)	Between		23.23	.34	81.35	n = 29
	Within		.37	20.12	23.37	$\overline{\text{T}}$ = 5.83

Notes: Private Credit is the credit to the private sector relative to GDP. *Gov. Spending* is government consumption as a share of GDP. *Inflation* is the yearly inflation, measured as the percentage change in the price of buying a specific "basket" of goods. *Population* is given in millions.
N is the total number of observations. n is the number of countries. $\overline{T}$ is the average number of years observed for each country.

4 Empirical Framework

To answer whether having an influential finance sector affects income inequality within a country, we estimate the following model:

$$gini_{it} = \alpha + \beta le_{it} + \boldsymbol{X'_{it}}\boldsymbol{\gamma} + \boldsymbol{\delta_t} + \eta_i + u_{it} \tag{2}$$

Our dependent variable is the Gini index in country i at time t, denoted by $gini_{it}$. le_{it} is the amount spent on lobbying from financial organizations in country i at time t, and α is an intercept. $\boldsymbol{X'_{it}}$ is a vector of control variables and $\boldsymbol{\gamma}$ is a vector of coefficients. All variables are measured in country i at time t. $\boldsymbol{\delta_t}$ is a set of year dummies, containing the years from 2011-2017, with 2010 as a base year. They capture Europe-specific shocks that affect all countries at the same time. η_i is an unobserved country specific error component that does not vary over time, while u_{it} is an idiosyncratic error component, varying both between countries and over time. We define $v_{it} = \eta_i + u_{it}$ as the composite error term.

The coefficient of interest is β, which is the point increase in the Gini coefficient following a €1,000,000 increase in lobbying expenses. In this chapter we present the different estimation methods that we use to identify β and discuss some issues with these methods. The estimation results, discussion and robustness checks will be provided in later chapters.

Based upon the view of legislation as a transfer of wealth, we expect that having an influential financial sector leads to distortions in the distribution of income. If a special group succeeds in affecting decisions, it will lead to increased profitability for that group. Therefore our hypothesis is that β is positive, meaning we expect that larger lobbying expenses leads to an increase in the Gini index.

4.1 Estimating β

To pin down the causal relationship between lobbying and inequality, we use pooled ordinary least squares (POLS) and fixed effects (FE) estimation. Both are methods for panel data regression and lie within the ordinary least squares (OLS) framework. The unbiasedness of the estimator and validity of statistical inference relies on different

assumptions regarding the underlying data and correlation structure of the composite error term v_{it}.

Identifying Assumptions

Some assumptions are the same for both POLS and FE estimators (see Wooldridge, 2009). Linearity, random sampling, and no perfect collinearity among explanatory variables are necessary for both estimators to be unbiased. The idiosyncratic error term, u_{it} must have a mean of zero in the population and it has to be uncorrelated with all explanatory variables. Which estimation method we prefer depends on the correlation structure between the unobserved heterogeneity, η_i, and the explanatory variables.

If statistical inference is to be valid, it is necessary with homoskedastic and serially uncorrelated error terms. With homoskedasticity we mean that the variance of the idiosyncratic error term must be constant. By serially uncorrelated errors we mean that there is no correlation between the idiosyncratic error terms in different time periods.

POLS

If there is no correlation between the unobserved heterogeneity , η_i, and the explanatory variables, the POLS estimator will be unbiased. It is estimated by applying ordinary least squares to equation (2), where each country-year observation is treated as an independent observation . An advantage of POLS is that it uses all the information in the data and exploits both the variation between countries and over time (within countries).

Fixed Effects:

However, the assumption of no unobserved heterogeneity is very strict. If there is correlation between the country fixed effects and the explanatory variables, the POLS estimator will be biased. With fixed effects, we construct the time-demeaned version

of the variables so that the country fixed effects are removed from the model. We transform the data to allow for country-specific unobserved effects. This is done by taking the time average of all variables and subtracting them from equation (2). We define the time average of a variable x_{it} as:

$$\overline{x_i} = \frac{\sum_{t=1}^{T_i} x_{it}}{T_i} \tag{3}$$

$\overline{x_i}$ is the average of the variable for all observartions of the same country. T_i is the number of years with non-missing observations for country i. The next step is to subtract the time mean from equation (2):

$$\begin{aligned} gini_{it} - \overline{gini}_i &= \beta(le_{it} - \overline{le}_i) + \boldsymbol{\gamma}(\boldsymbol{X_{it}} - \overline{\boldsymbol{X}}_i) + \boldsymbol{\delta}_t + u_{it} - \overline{u}_i \\ \ddot{gini}_{it} &= \beta\ddot{le}_{it} + \boldsymbol{\gamma}\ddot{\boldsymbol{X}}_{it} + \boldsymbol{\delta}_t + \ddot{u}_{it} \end{aligned} \tag{4}$$

Equation (4) is called the within transformation and the variables denoted with two dots are the time-demeaned variables. The important thing to note about this transformation is that we have eliminated the unobserved country heterogeneity η_i[13].

We get the FE estimator by doing POLS estimation on equation (4). It is unbiased even if the unobserved heterogeneity is correlated with the explanatory variables, provided that the rest of the assumptions for unbiasedness are satisfied.

Valid Inference

If nothing else is specified, standard errors are estimated under the assumption of homoskedasticity and serially uncorrelated errors. If one or both assumptions are violated it will result in incorrect standard errors, test statistics and P-values. This invalidates the tests for statistical significance and might lead us to wrongly reject (or fail to reject) a null hypothesis that should not be (or should be) rejected.

To correct for heteroskedasticity and serial correlation, we do all estimations using robust standard errors. In the case of fixed effects estimation, they are clustered at the country level. Robust and clustered standard errors allows for arbitrary forms of het-

[13] $\overline{\eta_i} = \eta_i$, and it falls out of the model in the within transformation, since $\ddot{v}_{it} = v_{it} - \overline{v}_i = \eta_i + u_{it} - \overline{\eta_i} - \overline{u}_i = u_{it} - \overline{u}_i = \ddot{u}_{it}$

eroskedasticity and serial correlation within countries[14] (see Verbeek, 2017, pp. 398-400).

4.2 Summary

In this chapter we have presented the computational methods and identifying assumptions for the POLS and FE estimators. These are the two estimation methods we use to investigate the causal relationship between financial sector influence and inequality. Throughout our analysis, we use cluster robust standard errors to correct for heteroskedasticity and serial correlation in the idiosyncratic error term.

[14] It does not allow for correlation between error terms across countries.

5 Results

In this section we present the results from pooled OLS (POLS) and fixed effects (FE) estimation of different specifications of the model in equation (2). The Gini coefficient is the dependent variable in all regressions. We are interested in the estimated coefficient on lobbying expenses, representing the expected change in the Gini coefficient following a €1,000,000 increase in lobbying expenses, when holding all other variables fixed.

We estimate three specifications of the model, using different sets of control variables. The results from all estimations can be seen in Table 4. We start with a simple model with only a few explanatory variables, and expand the model step by step. All estimations include year fixed effects to control for factors that affect all European countries at the same time[15]. The estimation sample consists of 29 countries and 167 country-year observations.

5.1 POLS Results

The results from the POLS regressions are presented in column 1-3 in Table 4. In column 1 we estimate the effect of increased lobbying expenses controlling only for number of organizations and population size. The estimated coefficient is 0.058, meaning that the Gini index is expected to increase by 0.058 points when lobbying expenses increase by €1,000,000, holding population and number of lobbying organizations fixed. For comparison, the average change in the Gini coefficient within the countries in our sample is 0.066. The largest decrease in the sample period occured in Cyprus between 2014 and 2017, where the Gini index decreased by 4 points. The largest increase was in Luxembourg between 2010 and 2017, where the Gini index increased by 3 points. Although it is small compared to the extreme cases, an expected

[15]We also estimated the model using country-specific time trends (See Table 9 in Appendix **??**). The estimated coefficients are similar in magnitude when compared to the case with year fixed effects, but the are no longer found statistically significant at conventional levels. This is to be expected, as including the country-specific time trend uses up an additional 28 degrees of freedom. It is not surprising that this affects the statistical significance of the coefficients in a relatively small sample.

Table 4: Pooled OLS and fixed effects estimation of different specifications of the model.

	OLS			FE		
	(1)	(2)	(3)	(4)	(5)	(6)
Lobbying Expenses	0.058	0.137***	0.105***	0.020	0.034	0.046**
	(0.044)	(0.044)	(0.031)	(0.020)	(0.020)	(0.022)
Reg. Organizaitons	–0.064**	–0.101***	–0.071***	–0.005	–0.005	–0.007
	(0.027)	(0.026)	(0.019)	(0.013)	(0.017)	(0.016)
Population	0.072***	0.083***	0.074***	–0.153	–0.277	–0.283
	(0.009)	(0.009)	(0.008)	(0.254)	(0.319)	(0.319)
Private Credit		0.023***	0.023***		0.017*	0.027*
		(0.005)	(0.005)		(0.010)	(0.015)
Gov. Consumption			–0.340***			–0.210
			(0.053)			(0.172)
Inflation			–0.112			–0.104
			(0.298)			(0.141)
Constant	27.452***	24.442***	31.961***	32.305***	32.997***	36.531***
	(0.616)	(0.877)	(1.324)	(5.474)	(6.628)	(7.422)
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R-Squared	0.14	0.22	0.37	0.01	0.07	0.09
N	167	167	167	167	167	167

Notes: The dependent variable in all columns is the Gini coefficient. It is a number between 0 and 100, where 0 represents perfect income equality and 100 is perfect inequality.
Lobbying Expenses is the total spending on lobbying (in €1,000,000) by financial sector organizations. *Reg. organizations* is the number of financial sector organizations in the EU transparency register. *Private Credit* is the credit to the private sector relative to GDP. *Gov. Spending* is government consumption as a share of GDP. *Inflation* is the yearly inflation, measured as the percentage change in the price of buying a specific "basket" of goods See Chapter 3 for detailed description of variables. Robust standard errors are reported in parenthesis, clustered at the country level when fixed effects estimation is used.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

0.058 point increase cannot be neglected. It is close to the average change in the Gini index and would account for a large increase compared to this. It should however be noted that the effect is not found to be statistically significant at conventional

significance levels.

The basic correlation just found is however plagued by the omission of relevant explanatory variables. If there are factors that affect inequality and are correlated with lobbying expenses, the POLS estimator will suffer from an omitted variable bias. To control for omission of relevant variables, we include additional control variables. This is to make sure that we are ruling out the effect on inequality that is caused by other variables.

In column 2 we include private credit to GDP as a control variable. The estimated coefficient on lobbying expenses is 0.137, meaning that a €1,000,000 in lobbying expenses is expected to lead to a 0.137 points increase in the Gini index. The effect is found to be statistically significant at a 10% significance level. In column 3 we control for private credit to GDP, government spending and inflation. The expected change in the Gini index following an increase in lobbying expenses is now found to be 0.105. The effect is statistically significant at the 5% level. When we add control variables to the model, we get a change in both the magnitude and statistical significance of the estimated coefficient.

Including a set of control variables should reduce the problem with omitted variables. However, this might not be enough to solve the problem. As mentioned in Chapter 4, it is strict to assume that there are no country-specific characteristics correlated with lobbying expenses. We are concerned that we are failing to include such variables in our model. An example of a such country-specific factor is a country's geographical position. Organizations in countries closer to the EU head-quarter will probably spend more on lobbying there than identical companies that are located further away. The effort might not be worth it for a bank in Finland, because of high transport and time costs, while a similar bank in the Netherlands would choose to lobby. The geographical position of a country might also affect the income distribution in a country, due to historical traits of an area or other socio-geographic characteristics. If we are failing to include such country characteristics in the model, the POLS estimator will be biased. FE estimation will allow us to control for this unobserved heterogeneity.

5.2 FE Results

In column 4-6 of Table 4, we show the FE estimation results on the same model specifications as used in the POLS estimation. The estimated coefficient on lobbying expenses in the simplest model specification is 0.02, and not statistically significant at conventional levels. When we control for private credit to GDP, the estimated coefficient is 0.034, and not statistically significant at conventional levels. In the model with a full set of control variables the expected increase in the Gini index following a €1,000,000 rise in lobbying expenses is 0.046 points. The effect is statistically significant at the 5% level.

FE estimation yields coefficients that are positive, yet smaller in absolute value than the corresponding POLS estimates. As expected, the FE and POLS differ in magnitude, which indicates that the POLS estimates are suffering from a positive bias. The coefficient is found statistically significant in the full specification of the model (column 6), and statistically insignificant in the two simpler models. Nonetheless, the magnitude of the coefficients is statistically the same across models. This is in line with the POLS results showing a small, positive effect of lobbying expenses on the Gini coefficient.

Except population size, the sign of the coefficients on all control variables are the same across different estimations. In general, most variables are found to have a statistically significant effect in the POLS estimation. However, when using POLS the error term will by definition be serially correlated unless the country fixed effects are constant across countries. This means that the assumptions for valid inference are violated and the apparent statistical significance of the coefficients in columns 1-3 cannot be trusted.

An advantage of the FE estimator is that it computes standard errors clustered by country, allowing for serial correlation in the idiosyncratic error term (Verbeek, 2017, pp. 398-400). In general the control variables are not found to be statistically significant when using FE estimation. The only control variable that is found statistically significant in the FE estimation, is private credit to GDP. From columns 5 and 6 we see that it is found to have a small, positive effect that is statistically significant at the 10% level. This is consistent with the findings of Jauch and Watzka (2016), who

find that the degree of the financial sector development increases the degree of income inequality.

In all model specifications, an increase in the number of registered organizations is associated with lower income inequality, when holding all other variables constant. We include it to correct for the change in lobbying costs that is caused by an exogenous increase in registered organizations[16]. This means that if two countries has the same lobbying expenses, but one has a larger number of lobbying organizations, the latter is expected to have lower inequality. An explanation for this could be that in a country with many small financial organizations, the market power and hence the influence of the sector can be smaller than in a country with fewer, but larger financial organizations.

Government consumption is expected to have a negative effect on inequality. High government consumption can be a sign of a country having an active redistributional policy using taxes and transfers to distribute wealth and income. Therefore, it is not surprising that increased government spending is associated with lower income inequality. The effect of inflation is found to be negative, and is the only control variable that is statistically insignificant in both POLS and FE estimation. This is contradictory to Jauch and Watzka (2016), who actually find inflation to be the only variable that is statistically significant in the FE estimation.

To summarize, we find evidence of a small, positive effect of lobbying expenses on inequality. The effect is qualitatively the same for POLS and FE estimation when controlling for other economic factors. This suggests that the financial sector contributes to increased income inequality, which is consistent with the findings of Jauch and Watzka (2016). They find that it is the degree of the financial sector development that affects distribution of income. We control for this by including private credit to GDP as a control variable. Our findings suggest that there are other mechanisms driving the increase in the Gini index. The distortions of income distribution can be accredited to the ability of the financial sector to exert its influence on the legislation process and capture the associated rents. This said, one should be careful in drawing conclusions from this analysis alone. We have studied one small part of the lobbying activity of a sector, and focused on a single measure of inequality. Fur-

[16]The register opened in 2011, and there has been a steady increase in number of registered organizations during the sample period.

ther research with diverse methodology and data sets will be necessary to draw more robust conclusions on the matter.

Our preferred model specification is the one in column 6 of Table 4, where we control for country and year fixed effects in addition to a full set of control variables. In this model, a €1,000,000 increase in lobbying expenses is expected to increase the Gini coefficient by 0.46 points. It is close to the average change in our sample, which is 0.066, meaning that the economic significance of the effect is substantial. Even though most control variables are not found statistically significant, we choose to include these time-varying factors, building upon earlier research on finance, growth and inequality. The FE estimates are smaller in magnitude than the POLS estimates, indicating that there are some unobserved country-specific characteristics causing a positive bias in the POLS estimates. We argue that FE estimation provides more credible results than POLS estimation. However, there are still issues with the model and the underlying data that should be addressed to improve the empirical analysis. These issues and possible solutions are discussed in the following chapters.

6 Robustness Analysis

In this section we perform a number of robustness checks to supplement the results from Chapter 5. We start by estimating the full specification of the model using different measures of inequality. We also do an estimation where we exclude interest organizations from our sample. This is done to correct for possible measurement errors.

Next, we discuss possible endogeneity issues. Our main concerns are measurement error, omitted variables and simultaneity. Finally, we present an instrumental variable approach using a Bartik instrument to address these issues.

6.1 Alternative Model Specifications

We do all estimations using FE estimations on a full specification of the model. The estimation results are presented in Table 5.

This far we have used the Gini index of income as our dependent variable. In this section we investigate whether our findings change when using different measures of inequality. In column 2 we use the Gini index from the World Bank and find that the estimated coefficient on lobbying expenses is 0.012, and it is not statistically significant at conventional significance levels. To compare it with the results from our Eurostat Gini index, we estimate it using the same subsample. The results can be seen in column 1. This yields a coefficient of 0.037, statistically significant at the 10% level.

The estimated coefficient is slightly smaller when we use the World Bank Gini, but the sign of the coefficient does not change. Two reasons could explain such difference: 1) the two measures differ in methodology 2) the underlying data is collected from different sources. Still, the two measures move together[17] and both measure income inequality as a Gini index. We loose 1/3 of our sample when using the World Bank estimate, which can explain why the estimates are less precise.

[17]The correlation coefficient is 0.9342.

Table 5: Estimation with alternative measures of inequality and excluding interest organizations from the sample.

	Alternative dependent variable			Without Interest Org.
	FE	FE	FE	FE
	Gini (Lim. samp.)	WB	S80/S20	Gini (W/o int. org.)
	(1)	(2)	(3)	(4)
Lobbying Expenses	0.037*	0.012	0.014*	
	(0.021)	(0.030)	(0.007)	
Expenses w/o Int. Org.				0.067***
				(0.023)
Reg Organizaitons	–0.021*	–0.001	–0.003	–0.004
	(0.012)	(0.009)	(0.004)	(0.015)
Population	–0.322	–0.384	–0.067	–0.450
	(0.271)	(0.312)	(0.095)	(0.349)
Private Credit	0.006	–0.004	0.004	0.033**
	(0.013)	(0.012)	(0.005)	(0.014)
Government Consumption	0.001	0.277*	–0.014	–0.147
	(0.166)	(0.161)	(0.049)	(0.160)
Inflation	0.053	0.268***	–0.055	–0.172
	(0.104)	(0.080)	(0.037)	(0.121)
Constant	35.561***	33.911***	5.865***	38.016***
	(7.226)	(7.933)	(2.005)	(7.889)
Year Dummies	Yes	Yes	Yes	Yes
Adj. R-Squared	0.05	0.13	0.11	0.18
N	112	112	167	154

Notes: WB is the Gini index from the World Bank DataBank. Its interpretation is the same as the Gini variable. *S80/S20* is the ratio between the income of the 20% with highest income and the 20% with lowest income. *Expenses w/o Int. Org.* are the total lobbying expenses when interest organizations have been taken out of the sample.
See footnotes of Table 4 or Chapter 3 for description of all variables.
Column 1 shows estimation with the Eurostat Gini index as dependent variable when using the same subsample as in column 2.
Robust standard errors are reported in parenthesis, clustered at the country level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Column 3 shows the results when the S80/S20 ratio is the dependent variable[18]. The estimated coefficient is 0.014 and is statistically significant at 10% significance level. This indicates that a €1,000,000 change leads to a 0.014 points increase in the s80/20 index. The average change in the S80/S20 ratio in the sample period is 0.033 so an expected change of 0.014 is not neglectable.

The S80/S20 ratio captures a different aspect of income inequality, focusing more on the extremes of the distribution. The S80/S20 ratio is highly correlated with the Gini coefficient[19], the two measures move in the same direction. Therefore it is not surprising that the effect of lobbying expenses has the same sign when estimating the model using the S80/S20 ratio as dependent variable.

Changing our measure of inequality does not change our main findings. An increase in lobbying expenses seems to have a small, positive effect on income inequality across different measures of inequality. Still, we need to be cautious in interpreting the results, as they are not always found to be statistically different from zero.

The estimation results when we exclude interest groups from our sample are reported in column 4 in Table 5. We do this to address potential measurement error caused by double-counting of lobbying expenses. Many companies that employ their own lobbyists are also members of local branch organizations or other interest groups that lobby on their behalf[20]. If a company is incorporating the expenses of lobbying through interest groups into their own expenses, the expenses of that company would be counted twice. To correct for such measurement error, we try excluding interest organizations from our sample. In column 5 we see the results from the FE estimation when we exclude interest groups[21]. The estimated coefficient on lobbying expenses is 0.067 and is statistically significant at a 1% significance level. This is a marginally

[18] As described in Chapter 3, the S80/S20 ratio is the income earned by the 20% with highest income divided by the income of the 20% with lowest income.

[19] The correlation coefficient is 0.96

[20] For instance, DNB ASA is a Norwegian bank, and has their own register ID in the transparency register. They are also a member of Finance Norway ("Finans Norge"), an interest organization for Norwegian financial companies. Finance Norway has their own registration ID and entry in the register. Thus, DNB is represented both through their own lobbyists and though the lobbying of Finance Norway.

[21] Summary statistics for lobbying expenses without interest groups can be found in Appendix A.1

bigger effect than what we get when including interest groups, which gives a coefficient of 0.059[22], but the effects are statistically the same. Excluding interest groups does not seem to affect our results.

6.2 Endogeneity Issues

There might be additional issues that are not solved by the robustness checks described above. In this section, we discuss three possible endogeneity issues that might be affecting our results. These are measurement error, omitted variable bias and simultaneity.

Measurement Error

Since lobbying expenses are self-reported by the lobbying organizations, we are worried about measurement error. It will vary between organizations how they report and count their lobbying expenses, and some might type the wrong number when registering. If there is measurement error, the observed lobbying expenses can be written as:

$$le_{it} = le^*_{it} + \epsilon_{it} \tag{5}$$

Where le_{it} denotes the observed lobbying expenses and depends on the true lobbying expenses, le^*_{it}, and a measurement error, ϵ_{it}[23]. Measurement error in an explanatory variable can be serious, as it causes an attenuation bias in the POLS estimator (Wooldridge, 2009), meaning that the absolute value of the estimator will have a downward bias.

To correct for measurement error we either have to improve the measure of lobbying or use instrumental variable estimation.

[22]The coefficient on lobbying expenses when using the same subsample as in column 4. It has a standard error of 0.02 and is statistically significant at the 1% level.

[23]ϵ_{it} can include both fixed country unobservable characteristics and time varying factors

Omitted Variables

If we fail to include a variable that affects both lobbying expenses and inequality, we have an omitted variable bias. We have already discussed FE as a solution to omitted time-constant variables. Our concern is that we are leaving out some factors that change over time. If this is the case, and they are not captured by the other control variables we use, the FE estimator will also be biased.

An example could be that we are not measuring lobbying activity directed at a country's own government. The amount spent on lobbying by a German bank in Germany can be correlated with lobbying expenses in the EU and with the income distribution in Germany. If they are positively correlated, and we are not including lobbying expenses at home in our model, the estimated effect of lobbying expenses will be overstated. Some of the change in inequality that is caused by lobbying at home, is assigned to the observed lobbying expenses in the EU, causing a positive bias in the FE and POLS estimators.

A solution to the problem of omitted variables is to include the omitted variables in the model or find a good proxy variable. We have included control variables to reduce the chances of omitting relevant variables, but we are still concerned that there are some unobserved factors that should be included in the model. However, they can be unknown to us or difficult to measure. If this is the case, instrumental variable estimation can be a solution.

Simultaneity

A possible caveat is that the causal relationship between lobbying expenses and inequality go both ways. As Alstadsaeter, Johannesen, and Zucman (2017, p. 6) find, the part of income earned as capital income increases as wealth increases. This indicates that wealth accumulation at the top can lead to increased activity in the financial sector and result in higher lobbying expenses.

The leading method for correcting for simultaneity is IV estimation. We need at least one instrument that helps determine lobbying expenses, but that does not affect the income inequality (Wooldridge, 2009, p. 554). This is necessary to identify the causal

relationship between the two when we have simultaneity.

6.3 Instrumental Variable Estimation

The endogeneity issues described above can all be addressed by using an IV. In this section we describe the construction of, and theory behind the Bartik instrument and show the results from the two-stage least-squares (2SLS) estimation when using a Bartik instrument. The instrument is named after Bartik (1991) (as cited in Goldsmith-Pinkham, Sorkin, & Swift, 2018, p. 1), and the approach was initially used in predicting local employment growth. More recently, it has been used in other areas of economic research, such as migration, development, macroeconomics and finance (Goldsmith-Pinkham et al., 2018).

Constructing the Barik Instrument

The intuition behind the Bartik instrument approach is that we redistribute the total lobbying expenses based on an exogenous country characteristic. We assign different shares, or weights, to each country than what we observe in our sample. More clearly: because we are worried that the observed lobbying expenses are endogenous, we take the sum of all expenses for each year, and re-distribute them to the countries in the sample. The variable we use to re-assign lobbying expenses must be exogenous if the instrument is to be valid. At the same time, it has to explain enough of the variation in the observed lobbying expenses, making sure the instrument is relevant.

To assure the exogeneity of the instrument we tried to find a variable that was far enough into the past to be independent of the situation today. At the same time we needed a variable with enough power in explaining the variation in lobbying expenses. Detailed statistics gets harder to obtain the further back you go, so we had to choose between going far enough back in time and getting observations for many countries[24]. There was a trade-off between validity and relevance.

We use the share of population that were employed in commercial banks in 1995.

[24]Other candidates were using the value of stocks traded relative to GDP in 1980 or the number of financial institutions in 1995. However, this would severely limit the size of our sample.

The data was obtained from the Structure of the Financial System database (OECD, 2019). We found data for 19 of the countries in our sample [25], allowing us to construct the instrument only for these 19 countries. For each country $i \in [1, 19]$, we got a value ecb_i, which is the share of population in country i working in commercial banks in 1995. From this we constructed the weights assigned to each country by computing the relative share of commercial bank employees. The weights, z_i, were computed as follows:

$$z_i = \frac{ecb_i}{\sum_{i=1}^{29} ecb_i} \tag{6}$$

A country with a large share of the population working in commercial banks in 1995 would be assigned a larger part of the total expenses than a country with a small share. The next step was to re-distribute the total lobbying expenses for each year, tle_t [26] by using the country weights, z_i. The Bartik instrument was then constructed as follows:

$$B_{it} = z_i tle_t \tag{7}$$

B_{it} is the Bartik instrument for each country-year observation. It represents lobbying expenses redistributed in the sample based on a historical feature of the financial sector. We instrument the potentially endogenous lobbying expenses, le_{it}, using B_{it} as an instrumental variable. Because tle_t changes over time we get variability in the instrumental variable, even though the weights are constant. Necessary conditions for validity and relevance are that B_{it} does not correlate with the idiosyncratic error component, but that it correlates with lobbying expenses.

[25] Summary statistics of the variables used in the IV estimation can be found in Appendix A.1

[26] $tle_t = \sum_{i=1}^{19} le_{it}$

Results of IV Estimation

Table 6: Instrumental variable estimation with Bartik instrument.

	FE	2SLS (Second stage)	FE (First stage)
	Gini	Gini	Lobbying Expenses
	(1)	(2)	(3)
Lobbying Expenses	0.037	0.501	
	(0.031)	(0.359)	
Bartik Expenses			0.169
			(0.108)
Reg Organizaitons	–0.002	–0.116	0.249***
	(0.016)	(0.091)	(0.046)
Population	–0.281	–1.462	2.638**
	(0.311)	(1.073)	(1.005)
Private Credit	0.018*	0.073	–0.137***
	(0.011)	(0.047)	(0.034)
Government Consumption	–0.027	–0.753	1.736**
	(0.213)	(0.683)	(0.694)
Inflation	–0.102	–0.167	0.141
	(0.132)	(0.252)	(0.437)
Constant	36.036***	78.752**	–96.239***
	(9.243)	(36.950)	(29.268)
Year Dummies	Yes	Yes	Yes
Adj. R-Squared	–0.22		0.65
N	119	119	120

Notes: Bartik Expenses is the instrumental variable. It is constructed using the share of population working in commercial banks in 1995 and the total lobbying expenses. Column 1 is the FE estimation on the same subsample as used in column 2.
See footnotes of Table 4 and Chapter 3 for further variable description.
Robust standard errors are reported in parenthesis, clustered at the country level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

We estimate the model using 2SLS with country- and year fixed effects, the results can be seen in Table 6. In column 1 we can see the regular FE estimation done on the same sample as the IV regression, while column 2 shows the results from the second stage of the 2SLS estimation. Column 3 shows the first stage FE regression.

The estimated coefficient on lobbying expenses when using the Baritk instrument is 0.5, and it is not statistically significant at conventional levels. This can be seen in column 1. The coefficient is larger than the FE estimate in column 1, where the estimated coefficient is 0.037[27]. The IV estimate also has larger standard errors, meaning that the estimate is less accurate. Using the Bartik instrument approach, we find a large(r) positive, but statistically insignificant effect of lobbying expenses on the Gini coefficient.

When interpreting the results from this IV regression, a word of caution is necessary. We have a problem with weak instruments and a small sample size, which might severely bias the 2SLS estimation results. Looking at the first-stage regression in column 3, it is clear that we should be concerned about weak instruments. When we regress lobbying expenses on the Bartik instrument and the full set of control variables, we get a coefficient of 0.169, with a standard error of 0.108. The t-statistic[28] is 1.55, which is well below the rule-of thumb for IV relevance, which states that the instrumental variable should have a t-statistic at least as large as 3.3 (see Wooldridge, 2009). The reason for using a larger critical value is that the validity of our instrumental variable is not testable. If our instrument is endogenous, and only weakly correlated with lobbying expenses, the bias in the 2SLS estimator can be larger than the bias using FE with an endogenous explanatory variable (Wooldridge, 2009). Since the coefficient on B_{it}, the Bartik expenses, is not statistically significant even at the 10% significance level, we should be concerned about a weak instrument. IV estimation is also shown to have poor small-sample properties (Wooldridge, 2009). The IV estimator is potentially biased in small samples, and is only justified in large samples. In other words, we should be very careful to draw conclusions from this IV estimation.

[27]Note that it is not found to be statistically different from zero here either. This might be because we are using a limited sample.

[28]For testing the null hypothesis that the $\beta = 0$ against the alternative that $\beta \neq 0$.

6.4 Summary

In this section we have showed that changing our measure of inequality and excluding interest groups from our estimations does not change our main conclusion. We find that an increase in lobbying expenses seems to have an effect that is either statistically zero, or a positive effect on income inequality across different measures of inequality.

Since we are concerned with endogeneity of our key explanatory variable, we attempt an IV approach with a Bartik instrument. The IV estimation shows a coefficient of 0.5, which is larger than all the other model specifications and estimation methods used before. This finding should, however, be interpreted carefully. The effect is not found to be statistically significant and we are worried about a weak instrument problem and the bad small-sample properties of the IV estimator.

7 Limitations and Further Research

In this section we discuss improvements that can be done to more accurately pin down the causal relationship between financial sector influence and inequality. We discuss our choice of proxy variables for influence and inequality, and potential issues with non-random sample selection. We also come with recommendations and thoughts on further research and ways to improve the analysis.

7.1 Income or Wealth Inequality?

We have used the Gini index of income inequality as our main dependent variable throughout the thesis. Using the Gini index allows us to estimate the model using panel data and not only a single cross-section. It has the advantage that it is comparable between countries and available for several years. However, the use of the of the Gini index has some drawbacks. Its main disadvantage in this context is that it measures the distribution of income, and not the distribution of wealth. The distribution of wealth is often more uneven than that of income, with thicker right tails in the distribution of wealth than that of income (Benhabib, Bisin, & Luo, 2017). Many of the gains from the financial sector are realized as yield on financial assets, and not as reported income. Therefore it can be expected that the financial sector might have an impact on the accumulation of wealth that is not necessarily reflected in the income distribution. Piketty and Saez (2014) stress the importance of looking at wealth distribution and not only income distribution. They suggest looking at decile shares, for instance the share of total wealth owned by the top decile of the population.

Ideally, we would have used a measure of wealth inequality rather than income inequality. However wealth inequality is often harder to measure. We were not able to obtain measures for all the countries in our sample. Existing measures of wealth inequality can also be more inaccurate. For example the work of Alstadsaeter et al. (2017) finds that tax evasion and use of tax havens is much more common among the top percentiles of the wealth distribution. This means that the true wealth distribution can be difficult to measure correctly.

To summarize, there are limitations of using the income Gini index as a measure of inequality. We justify the use of it as it is the best available measure for the countries and years in our sample.

7.2 Measuring Influence

We have used expenses on lobbying toward the European Parliament and Commission as a measure of influence of the financial sector. Two remarks can be made about this. Firstly, it is measured at the EU level, and does not measure how much the organizations spend on lobbying in their home country. Secondly, the lobbying expenses does not give the full picture of what lobbying effort is aimed at, or whether it is successful.

It is desirable to have a measure of influence that says more about how influential the financial sector is in each country, not only at the EU level. If lobbying expenses in the EU is not highly correlated with the actual influence in the home country, it might not be a good proxy variable for influence. One possibility could be to use countries that have their own lobby register and do similar analysis or to supplement data from the EU register. One could also try and find a different measure of influence that is available at country level.

Ideally, we would want a measure of financial sector influence for a wider range of countries and years, looking for example at all European countries. Instead, we measure lobbying in the EU, and are therefore missing observations for some countries and years. What countries we observe is probably not random, meaning we should be concerned with selection into our sample. The probability of appearing in our sample can be a function of underlying characteristics of the country itself and the financial sector in the country. There could be some latent variable, that we are not measuring, that determines whether a country is in our sample or not. It could for instance be that we are observing the lobbying expenses only for those countries with a financial sector larger than a certain threshold. The size, in turn, is determined by other country characteristics. A possibile solution would be to estimate a sample selection (Heckman) model, to where we take into account the non-random selection into the sample (see Verbeek, 2017, p. 260).

The second remark is that looking only at lobbying expenses we might not capture the full picture of what lobbying activity is aimed at. As mentioned in Chapter 2, lobbying is not only a means to rent seeking. Using only lobbying expenses might not be enough to tell how lobbyists act, or what the aim of different lobbying activities is. It would be possible to use more information about the lobbying organizations and create a more differentiated measure of lobbying activity. With more time and resources at hand, one could construct a more qualitative measure of lobbying, going deeper into the lobbying activity of each organization and what they achieve. We could look closer into the meetings between EU representatives and lobbying activity directly aimed at specific issues or whether they participate in expert panels. This could be a way to separate the part of lobbying activity aimed at rent seeking, and that aimed at providing information. This would give a more nuanced picture of lobbying activity.

7.3 Sample Selection Bias

Our data set contains missing country-year observations. If they were missing at random, this would not be a big problem. One source of sample selection has already been discussed, but there are two more possible sources of missing observations. It can be that we have not been able to identify all financial organization, or it can be that some organizations are lobbying without registering.

We might have failed to identify some financial sector organizations in the data handling process, which can affect the results. However, we would not be too concerned about this, as it is a feature of the manual data handling process. One can argue that these observations are missing at random and should not (on average) be affecting the results (Verbeek, 2017, p. 51).

If the problem is non-registered lobbying organizations, we might worry that these are not randomly missing. This could be handled by a sample selection model, as described above, or by trying to improve the measure of influence.

Our solution to missing observations, is to exclude them from our sample. We consider this the best method in our case, even though it decreases the sample size.

7.4 Sample Size and Within Variation

The sample we have been able to obtain consists of a total of 167 observations, with a maximum of 8 years for the same country. We argue that a larger sample, both in terms of number of time periods and observed countries, would be desirable to increase the quality of our work.

The Transparency Register has existed since 2011, meaning there is a natural limit to how far back our sample period can go. It is also an issue that the number of registered organizations has been growing for each year since the implementation of the register, meaning we might have lobbying organizations that have not registered in the beginning of the sample period. It would be interesting to repeat our analysis at a later point in time, as we anticipate that the Transparency Register will be more complete in the future, allowing for an increased sample size.

Another advantage of increasing the number of time periods is that it would allow for more variation in the Gini index as well. One issue with our data is that there has been little change in the Gini indexs within each country in the sample period. Piketty and Saez (2014) find that inequality has been increasing over time. While they have data available from 1910-2010, our time period is much shorter. It would be desirable with a data set including more observations and time-periods.

As discussed earlier, IV estimation has poor small-sample properties. A larger sample would therefor improve the quality of an IV strategy.

8 Conclusion

In this master's thesis we present evidence on the connection between lobbying and distribution of income in a selection of European countries. Using lobbying expenses of financial organizations as a measure of influence and the Gini index to measure income inequality, we find that increased financial sector lobbying is associated with an increase in income inequality. When controlling for year- and country fixed effects and a set of other economic variables, we find that a €1,000,000 increase in lobbying expenditures is expected to increase the Gini index by 0.046 points. This is not neglectable, as the average change in the Gini index in our sample is 0.066. The estimated effect is qualitatively the same across different estimation methods and model specifications.

By constructing our own data set using the EU Transparency Register, we provide a measure of political influence. To the best our knowledge, ours is the first economic paper to exploit the Transparency Register in this particular way. It contributes to the finance, growth and inequality literature. It also adds to the rent seeking literature and provides an empirical study of lobbying in Europe.

The scope of a master's thesis is, of course, limited. However, our findings and methods open up for additional research, both in terms of methods and the use of this particular data set. A challenge that should be addressed are the endogeneity issues that arise from measurement error, omitted variables or simultaneity. We attempt to solve this using a Bartik instrument approach, but find that our instrument has severe limitations due to a weak instrument and small sample size.

Future research should attempt to improve the measure of lobbying, and could also look for different ways to measure influence. Larger sample sizes in terms of countries and years could allow the use of more advanced estimation methods. Additional improvements can be made if using a measure of wealth inequality rather than income inequality.

Our findings imply that successful lobbying might result in a transfer of wealth towards influential interest groups. As the EU system relies on interest group organizations to represent the views of its citizens, member states and businesses, it is important to study the patterns of how interest groups affect policy.

References

Alstadsaeter, A., Johannesen, N., & Zucman, G. (2017). Tax Evasion and Inequality. *NBER Working Paper Series, 23772*.

Arcand, J. L., Berkes, E., & Panizza, U. (2015). Too Much Finance? *Journal of Economic Growth*, *20*(2), 105–148.

Beck, T., Demirguc-Kunt, A., & Levine, R. (2004). Finance, Inequality, and Poverty: Cross-Country Evidence. *NBER Working Paper Series, 10979*.

Benhabib, J., Bisin, A., & Luo, M. (2017). Earnings Inequality and Other Determinants of Wealth Inequality. *American Economic Review*, *107*(5), 593–597.

Bertrand, B. M., Bombardini, M., & Trebbi, F. (2014). Is It Whom You Know or What You Know ? An Empirical Assessment of the Lobbying Process. *American Economic Review*, *104*(12), 3885–3920.

Corporate Europe Observatory & LobbyControl. (n.d.). Statistics [Database]. Retrieved January 29, 2019, from https://lobbyfacts.eu/reports/lobby-costs/all

European Commission. (2016). *Proposal for a Interinstitutional Agreement on a mandatory Transparency Register*. European Commission. Brussels. Retrieved May 4, 2019, from http://ec.europa.eu/transparency/regdoc/rep/1/2016/EN/1-2016-627-EN-F1-1.PDF

European Commission. (2019). Who is Expected to Register? Retrieved June 1, 2019, from http://ec.europa.eu/transparencyregister/public/staticPage/displayStaticPage.do?locale=en&reference=WHOS_IS_EXPECTED_TO_REGISTER

Eurostat. (2017). Income distribution and monetary poverty [Database]. Retrieved April 30, 2019, from https://ec.europa.eu/eurostat/web/income-and-living-conditions/data/database

Goldsmith-Pinkham, P., Sorkin, I., & Swift, H. (2018). Bartik Instruments: What, When, Why, and How. *NBER Working Paper Series, 24408*.

Igan, D., Mishra, P., & Tressel, T. (2012). A Fistful of Dollars: Lobbying and the Financial Crisis. *NBER Macroeconomics Annual*, *26*(1), 195–230.

Jauch, S. & Watzka, S. (2016). Financial Development and Income Inequality: a Panel Data Approach. *Empirical Economics*, *51*(1), 291–314.

Klüver, H. (2012). Biasing Politics? Interest Group Participation in EU Policy-Making. *West European Politics*, *35*(5), 1114–1133.

Levine, R. (2005). Finance and Growth: Theory and Evidence. In *Handbook of economic growth* (Vol. 1A, pp. 865–934). Burlington: Elsevier Science

Mueller, D. C. (2003). *Public choice III*. Cambridge: Cambridge University Press.

OECD. (2019). Structure of the financial system [Database]. Retrieved May 1, 2019, from https://doi.org/10.1787/bank-data-en

Piketty, T. & Saez, E. (2014). Inequality in the Long Run. *Science*, *344*(6186), 838–843.

Piketty, T. & Zucman, G. (2014). Capital is Back: Wealth-Income Ratios in Rich Countries 1700-2010. *Quarterly Journal of Economics*, *129*(3), 1255–1310.

The General Secretariat of the Council. (2013). Interest groups in EU decision making. Retrieved May 30, 2019, from https://www.consilium.europa.eu/en/documents-publications/library/library-blog/posts/interest-groups-in-eu-decision-making/

Verbeek, M. (2017). *A Guide to Modern Econometrics* (5th ed.). United Kingdom: Wiley Custom.

Wolf, M., Haar, K., & Hoedeman, O. (2014). *The Fire Power of the Financial Lobby: A Survey of the Size of the Financial Lobby at the EU Level*. Corporate Europe Obseravatory. Retrieved March 8, 2019, from https://corporateeurope.org/sites/default/files/attachments/financial_lobby_report.pdf

Wooldridge, J. M. (2009). *Introductory Econometrics : A Modern Approach* (4th ed.). Canada: Cengage Learning.

World Bank. (n.d.). World Development Indicators [Database]. Retrieved April 30, 2019, from https://databank.worldbank.org/data/reports.aspx?source=2&series=NE.CON.GOVT.ZS&country=#

A Appendices

A.1 Variable Description

Table 7 shows a description of all variables and information about where they are retrieved from.

In Table 8 we present summary statistics for all variables that we have used in our regressions, including those that we did not refer to in the main body of the our text.

Table 7: List of all variables

Variable	Definition	Source
	Lobbying variables	
Lobbying expenses	Sum of lobbying costs of organizations with head office in the same country.	Lobbyfacts (Corporate Europe Observatory & LobbyControl, n.d.)
	Inequality variables	
Gini index (Europe)	Gini index based on the EU Statistics of Income and Living Conditions (SILC) survey.	Eurostat (2017)
S80/S20	The total income of the 20% with highest income divided by the total income of the 20% with lowest income.	SILC (Eurostat, 2017)
Gini index (World Bank)	Gini index (between 0 and 100) measures the degree of income (or consumption) inequality within a country	World Development Indicators (World Bank, n.d.)

Control variables		
Credit to the private sector (% of GDP).	The value of all loans and other credits provide to the private sector by financial corporations.	World Development Indicators (World Bank, n.d.)
Government consumption (% of GDP)	General government final consumption.	World Development Indicators (World Bank, n.d.)
Inflation	Yearly change in the consumer price index.	World Development Indicators (World Bank, n.d.)
Bartik variables		
Commercial Bank Employees (1995)	Share of population working in commercial banks in 1995.	Structure of the Financial System (OECD, 2019)

Table 8: Summary statistics of all variables used in our regressions

Variable	Mean	Sd	Min	Max	Obs.
Lobbying variables					
Lobbying Expenses	4.85	9.92	0.00	48.47	169
Reg. Organizaitons	15.09	23.09	1.00	126.00	169
Expenses w/o Int. Org.	3.34	6.93	0.00	37.02	156
Avg. Expenses	2.13	2.50	0.00	15.75	169
Inequality variables					
Gini (Eurostat)	29.21	3.20	23.20	40.20	167
Gini (World Bank)	31.17	3.07	25.40	36.30	112
S80/S20 Ratio	4.63	0.91	3.30	8.20	167
Control variables					
Private Credit	104.26	43.44	33.37	250.01	168
Government Consumption	20.20	3.69	11.60	27.40	169
Inflation	1.11	1.27	-2.10	4.24	169
Population	21.75	25.24	0.34	82.69	169
Bartik variables					
Employees in 1995 (%)	1.80	1.00	0.27	4.43	N = 121
Bartik Expenses	5.79	7.69	0.83	44.58	121

Notes: Reg. Organizations is the number of financial sector organizations in the EU transparency register *Expenses w/o Int. Org* is lobbying expenses when excluding interest groups. *Employees 1995* is the percentage of population employed in commercial banks.
See Table 7 for explanation of other variables.
N is the total number of observations.

A.2 Graphic Presentation of Lobbying and Inequality

Figure 2 and 3 shows how lobbying expenses and the Gini coefficient have evolved over time for the countries represented in our sample. Note that Figure 2 adjusts the y-axis to allow for differences between countries.

Figure 2: The evolution of lobbying expenses in the sample.

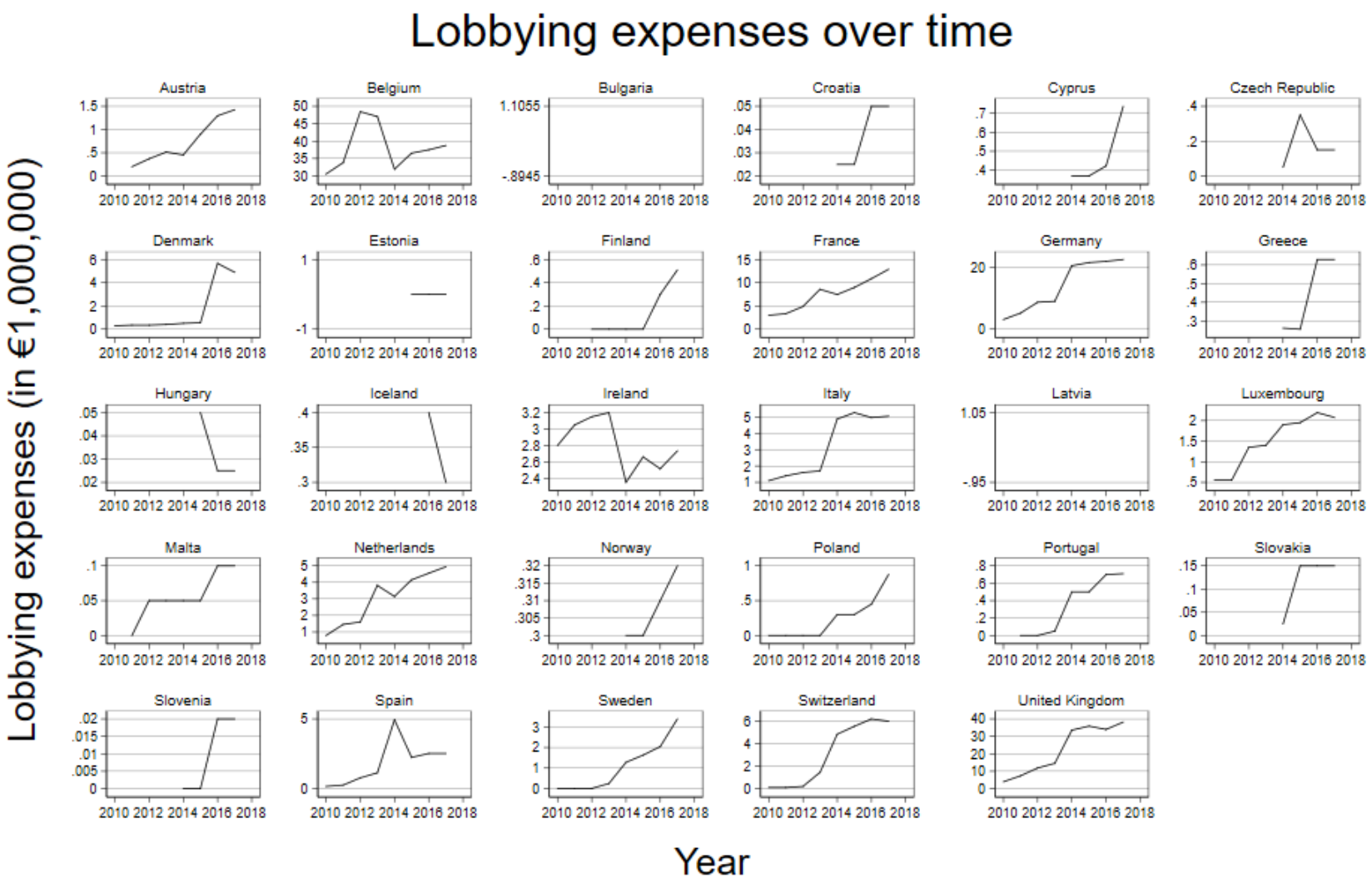
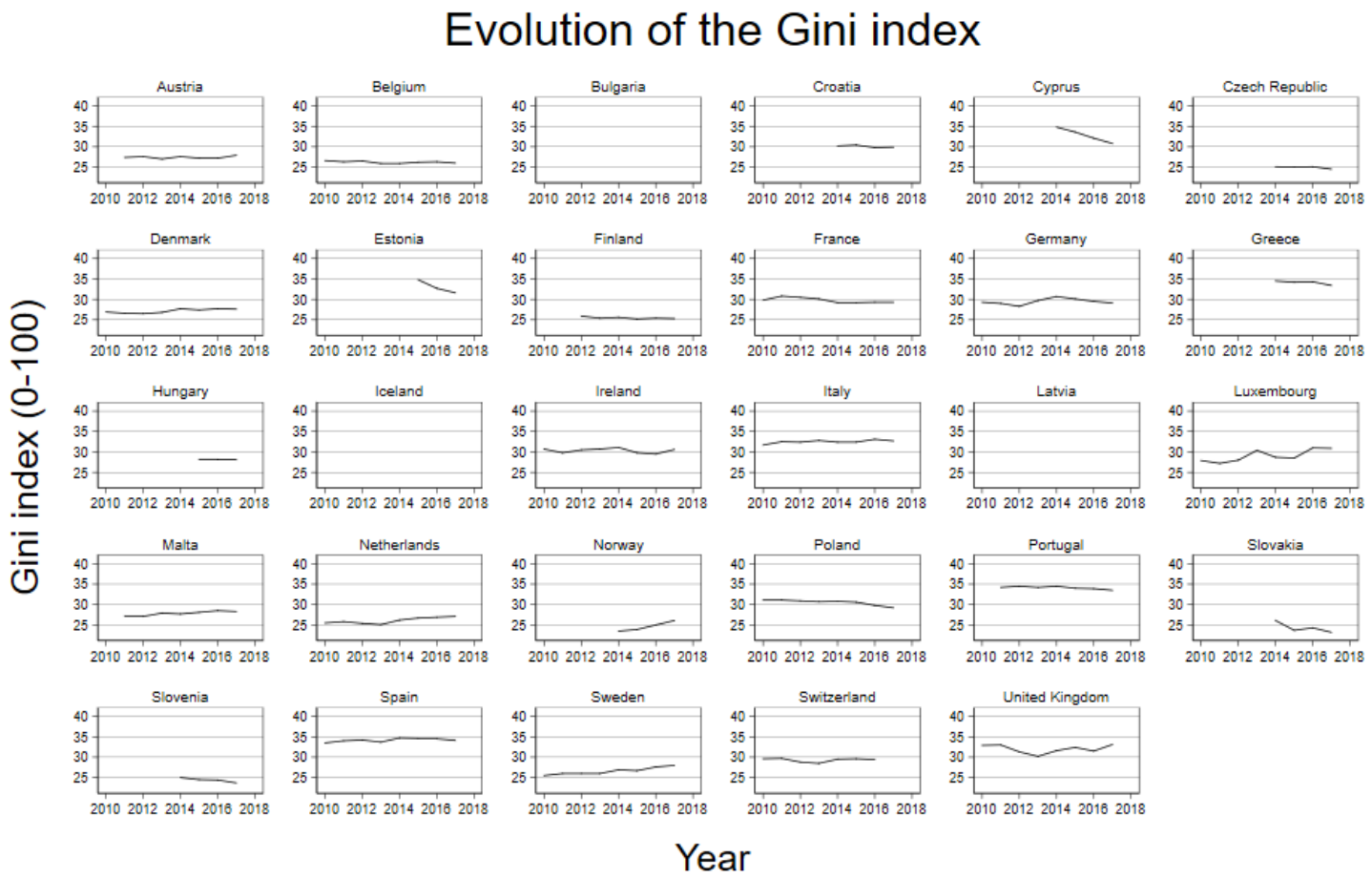


Figure 3: Evolution of the Gini index in our sample



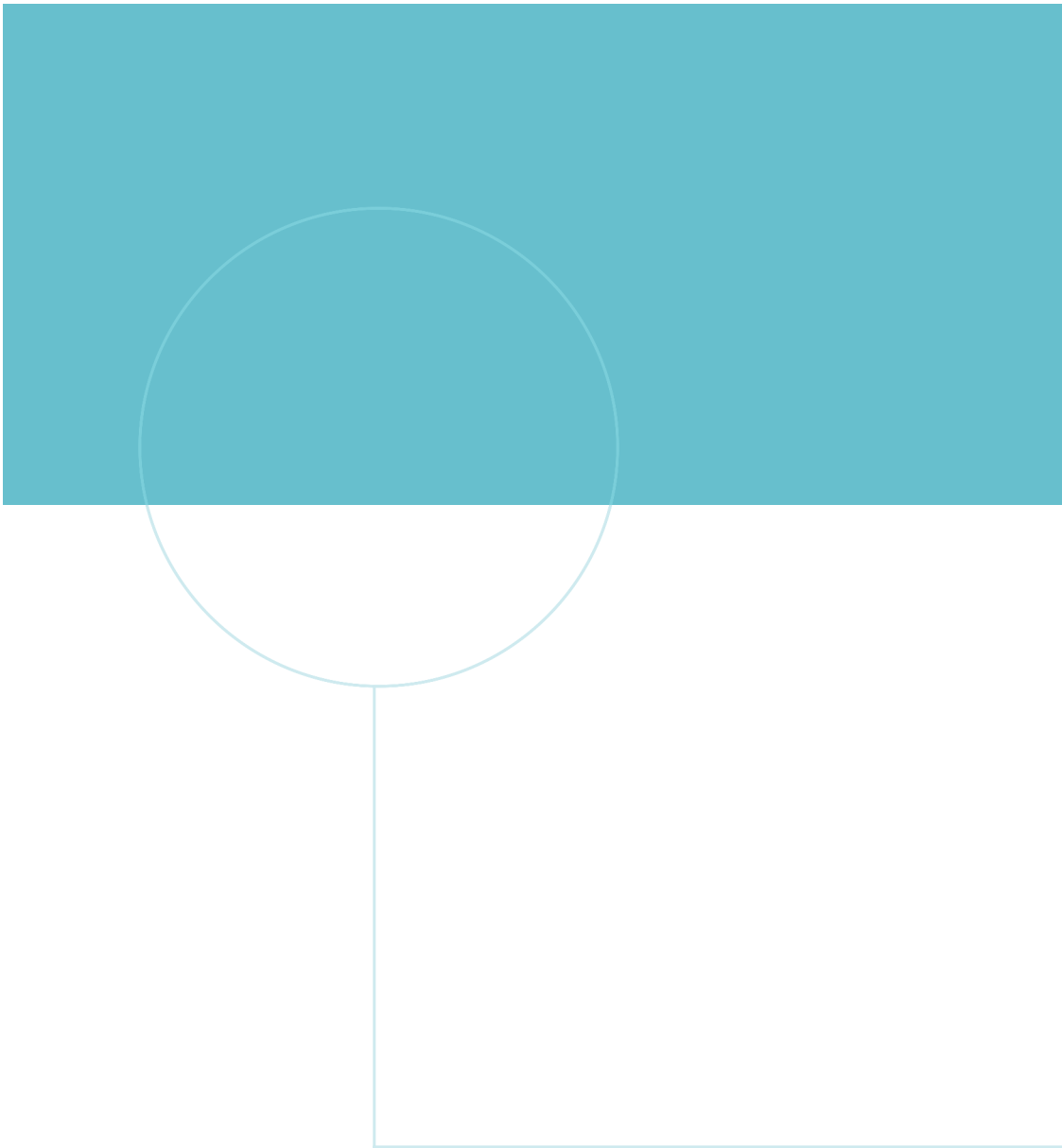
A.3 Country-specific Time Trends

Table 9 shows fixed effects estimation with country-specific time trends. The results should be compared to the results in Table 4.

Table 9: FE estimation when including country-specific time trends.

	FE	FE	FE
	(1)	(2)	(3)
Lobbying Expenses	0.022	0.044	0.048
	(0.034)	(0.038)	(0.035)
Reg Organizaitons	0.039	0.031	0.030
	(0.065)	(0.056)	(0.056)
Population	–0.159*	–0.220**	–0.233*
	(0.085)	(0.104)	(0.116)
Private Credit		0.049***	0.057***
		(0.017)	(0.018)
Government Consumption			–0.136
			(0.144)
Inflation			–0.068
			(0.045)
Constant	331.535	22.332	40.523
	(319.534)	(267.341)	(267.496)
Year Dummies	Yes	Yes	Yes
Country Trends	Yes	Yes	Yes
Adj. R-Squared	0.56	0.60	0.61
N	167	167	167

Notes: The dependent variable in all columns is the Gini coefficient. It is a number between 0 and 100, where 0 represents perfect income equality and 100 is perfect inequality.
Lobbying Expenses is the total spending on lobbying (in €1,000,000) by financial sector organizations. *Reg. organizations* is the number of financial sector organizations in the EU transparency register. *Private Credit* is the credit to the private sector relative to GDP. *Gov. Spending* is government consumption as a share of GDP. *Inflation* is the yearly inflation, measured as the percentage change in the price of buying a specific "basket" of goods See Chapter 3 for detailed description of variables. Robust standard errors are reported in parenthesis, clustered at the country level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$



NTNU

Kunnskap for en bedre verden