IMPRS SPCE

INTERNATIONAL MAX PLANCK RESEARCH SCHOOL
on the Social and Political Constitution of the Economy

H. Lukas R. Arndt

Linking Wealth and Power

Unity and Political Action of the World's Wealthiest Capitalist Families and the Corporate Elite

Studies on the Social and Political Constitution of the Economy

H. Lukas R. Arndt

Linking Wealth and Power

Unity and Political Action of the World's Wealthiest Capitalist Families and the Corporate Elite

Published by IMPRS-SPCE
International Max Planck Research School on the Social and Political Constitution of the Economy, Cologne
imprs.mpifg.de

ISBN 978-3-946416-25-8
DOI 10.17617/2.3523578

Studies on the Social and Political Constitution of the Economy are published online on **imprs.mpifg.de** and on **www.mpifg.de.**

Studies on the Social and Political Constitution of the Economy

Abstract

The exceptional concentration of wealth in western democracies during the past five decades has by now become common knowledge in the study of social stratification. This brings to the spotlight families and individuals who possess extraordinarily high levels of wealth. One concern about wealth concentration is that its beneficiaries might be a threat for democracy and the principle of one person, one vote. More precisely, the concern surrounds the potential concentration of their structural and instrumental power. From the perspective of class analysis this would be individuals who concentrate control over capital in capitalist democracies: wealthy individuals and families, and the powerful managers of large corporations and financial institutions. This closely relates to an old question: To what extent can the rich be understood as part of a capitalist class in itself and for itself? This project draws on ideas from the sociology of elites, corporate governance, corporate political action as well as (neo-)Marxist class analysis to develop hypotheses on the existence of capitalist classes in the 21st century. To test these hypotheses empirically, a large sample of ca. 1 million of the largest corporations (ORBIS) is combined with data on super-rich individuals and families, as well as individual and firm political action in Germany and the US. In one methodological and three empirical cumulative articles, this dissertation examines the existence and political action of capitalist classes in capitalist democracies and their political action.

About the author

H. Lukas R. Arndt was a doctoral researcher at the IMPRS-SPCE from 2018 to 2023. He participated in the Cotutelle Program, a Franco-German doctoral program run in cooperation with Sciences Po, Paris. He completed the program in conjunction with his PhD studies at the University of Cologne, Faculty of Management, Economics and Social Sciences, and at Sciences Po.

Linking Wealth and Power

Unity and Political Action of the World's Wealthiest Capitalist Families and the Corporate Elite

Inauguraldissertation
zur
Erlangung des Doktorgrades
der
Wirtschafts- und Sozialwissenschaftlichen Fakultät
der
Universität zu Köln

2023

presented
by

Hans Lukas Richard Arndt, MSc M.A.

from

Minden

First reviewer:	Prof. Jens Beckert (MPIfG and WiSo Uni Köln)
Second reviewer:	Prof. Olivier Godechot (CNRS-Sciences Po Paris-CRIS)
Third reviewer:	Prof. Clemens Kroneberg (ISS, WiSo Uni Köln)
Fourth reviewer:	Prof. André Kaiser (CCCP, WiSo Uni Köln)
Fifth reviewer:	Prof. Lucas Chancel (CNRS-Sciences Po Paris-CRIS)
First external reviewer:	Prof. Catherine Comet (Université Paris 8 Vincennes-Saint-Denis)
Second external reviewer:	Prof. Mark S. Mizruchi (University of Michigan)
Date of oral defense:	26 June 2023

This dissertation was written in accordance with a joint cotutelle agreement between the University of Cologne and Sciences Po Paris. The exact same version with a differing title page was submitted to the doctoral school of Sciences Po Paris.

Content

Introduction:

Should We Still Study the Capitalist Class?[1]

1. Motivation

In his famous article "The market as prison", US political scientist Charles E. Lindblom elaborates on the structural power of business in capitalist democracies. On which issues will those with economic power intervene? Lindblom writes: "the included aspects are all those in which businessmen [sic!] - *or any large or critical number of them* - see change as hurtful to their own prospects" (Lindblom 1982:330; emphasis by me). This sentence and especially its parenthesis is the starting point of this dissertation. One way to see this cumulative dissertation is it asks four clarifying questions to apply Lindblom's quote to the empirical reality of the early 2020s. First, who are these businesspeople and how can we understand them theoretically? (Introduction) Second, is there a reliable dataset to analyze and measure super-rich involvement in firms on a large and global scale? (1st Article) Third, what (large or critical) number of businesspeople does unite? (2nd Article) Third, how do these businesspeople approach the political process? (3rd Article) And finally, fourth, how critical is a critical number of them in which countries? (4th Article) This introduction serves to give theoretical context to the four alone-standing articles which are part of this cumulative dissertation and to evaluate their results against the theoretical background. It focuses on the first clarifying question towards Lindblom raised above. Namely, this is to ask who are the businesspeople that we should study when interested in business power. I argue that increased wealth concentration and the consequent rise of the super-rich make it necessary to give this question ennewed theoretical and empirical attention. The guiding question of this dissertation, therefore, is how to understand individuals and groups with power in the economy from the perspective of sociological theory: either as a class in the Marxist sense or more descriptively as an elite. But why ask this question today in the first place?

[1] This introduction is based on a paper that was presented in different versions at the SASE mini-conference "Mind the Wealth Gap? Interdisciplinary Perspectives on the Accumulation, Justification and (Re)Distribution of Wealth", 18-21 July 2020, a panel of the section of social inequality at the German Sociological Association congress on 24 September 2020, and the graduate conference of Hans Böckler Foundation "But some are more equal than others – Interdisciplinary Perspectives on Growing Inequality and the Need for Solidarity" on July 15-17, 2021. I am very thankful to Jens Beckert, Julia Cagé, Pierre François, and Olivier Godechot for helpful comments on the article on which the chapter is based. The article also benefitted from critical comments at the research group meeting of the research group on economic sociology at MPIfG Cologne on 17 February 2021.

Since the 1970s, several fundamental transformations reformed global capitalism and its national varieties. Neoliberalism led to the extension of the free market into many areas through privatization, Stalinism and Statism was overthrown in most countries, trade globalized, financialization and severe financial crises disrupted national and global economies, and economic growth declined to almost zero in many highly economically developed countries. These and other factors in combination led to the increase of global wealth and income inequality since about the 1980s, which was only realized by economists rather recently (e.g. Milanović 2016; Piketty 2014). Piketty's famous r>g hypothesis suggests that wealth accumulates faster than wages and economies grow, which leads to a significant concentration of wealth in the hands of those who already possess a lot of wealth. While this hypothesis was challenged since its publication (Góes 2016; King 2017), it is still an open question if the suggested relationship is actually at work but there is some good evidence especially in the long run (cf. De Donder and Roemer 2015; Ederer, Mayerhofer, and Rehm 2020; Madsen 2018; Milanović 2017).

One of the striking aspects of current forms of capitalism may be the immense concentration of wealth in the hands of very few. This is also reflected by simply tracking the number of billionaires worldwide and the wealth they hold. In 2020, there are about 2,095 billionaires worldwide (Forbes 2020). The top 1% owned 38.7% of global private wealth, and the bottom 50% of the distribution owned 1.8% of global private wealth (WID.WORLD 2018). Figure 1 shows the development of billionaires' assets in comparison to a company index average (MSCI ACWI). Billionaire-controlled firms grew much stronger - which UBS and PwC call "the billionaire effect" (UBS and PwC 2019). This dissertation goes beyond the focus on billionaires and analyzes a larger group of the wealthiest global families with an estimated net worth of 250 USD or EUR or more. According to wealth data provider Wealth-X, this group is estimated to include 25,120 individuals worldwide in 2022 (Wealth-X 2022:5).

Only after Piketty's popular contribution, we started to learn more about the group of the world's richest today. For example, Page et al. (2018:13–18) describe the 100 richest Americans as "much older (mostly in their sixties and seventies), much more often male, and much more frequently white and of Anglo-Saxon or other Western European origin than the very diverse American population as a whole" (Page et al. 2018:14). Presenting an enriched dataset on global billionaires, Freund and Oliver (2016) show similar characteristics also for the richest outside the US. Relevant differences between regions include that in Europe, more wealth is inherited (about half of all Forbes-listed fortunes) compared to the US (about one-third). Companies on which fortunes are based are younger in the US and driven by other sectors such as especially finance in the US case. Generally speaking, one can say that "one needs business assets to be rich" (Keister,

Li, and Lee 2021) and that the largest global fortunes are related to a valuable company held by a family.

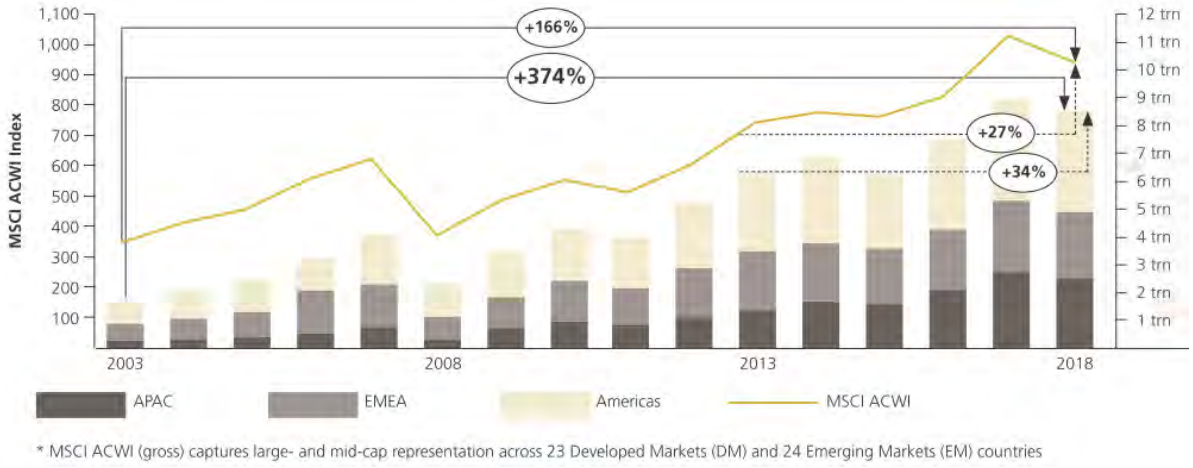


Figure 1: Asset growth of "billionaire controlled public companies" by region, in comparison to an index of average companies in developed and emerging markets (UBS and PwC 2019:19)

In the normative debate surrounding wealth and its concentration, it was claimed that capital concentration in the hands of a few might lead to the concentration of political power. Some observers even see an oligarchy emerging in the United States (Formisano 2017; Winters and Page 2009). However, it remains relatively unclear how this translation from economic to political power is supposed to take place. In this introduction and this dissertation as a whole, I will examine how sociological theory may help to understand the power of the super-rich, and other powerful actors in the economy who benefit from capital concentration. As a first step, I ask how we can understand the super-rich (and powerful managers) from a theoretical perspective. I suggest that two traditions invite themselves for inspiration. These are social class analysis, and the sociology of elites which are among the oldest traditions in sociology.

To bring this argument forward, I will begin by reviewing the literature on class analysis, especially the top part of class schemes, and economic elites. Second, I conceptualize how the super-rich can be understood to hold power through their control over capital. Third, I present the four articles which are part of this cumulative dissertation, and how they are related to the raised theoretical concerns. Finally, I present how far the empirical results of the included articles undermine the theoretical assumptions of neo-Marxist class analysis, and rather suggest that currently there is not much empirical evidence for a uniformly acting capitalist class in western capitalist democracies. To close, I briefly make the case of how more heterogenous elite networks

including the corporate elite, as well as super-rich capitalist families as distinct groups, seem to have explanatory power at this point in history.

2. The agents of capital: a class or an elite?

Karl Marx developed outlines of classes in his study of capitalist political economy and especially in his *18th Brumaire of Louis Bonaparte* from 1852 (Marx 1965). Max Weber's concept of class, status, and party in *Economy and Society* – published 100 years ago – is the second important reference among the classics (Weber 1922). The study of elites, or powerful positions within organizations or societies as a whole, was at least implicitly already a topic since the ancient Greeks (cf. Korom 2015). In modern sociology, it is mostly traced back to Italian writers Alfredo Pareto (1936) and Gaetano Mosca (1939) originally written as a critique of Marx's concept of class (cf. Korom 2015). The two traditions are united in the ambition to explain how economic positions affect individuals or groups that obtain them. In the last century, sequences of updated versions of class schemes were proposed to apply Marx's and Weber's class analysis to large-scale empirical cases such as national panel surveys. Such updates were often justified by fundamental economic changes that required adjustments over time to better represent the empirical reality.[2] One could argue that a revision of existing class schemes is once more urgently necessary. Recent economic developments have especially affected a group that was quite absent from class analysis in sociology after WWII: those at the very top of the wealth distributions, and those at the very top of income distributions. In terms of class and elites, these are the owners of capital (top wealth, and top capital incomes), and economic elites (top labor incomes) which constitute the very top economic positions in capitalist societies. Individuals with power over the economy seem to be a blind spot in some of the most frequently used theoretical concepts in sociology. However, as I aim to present in this chapter, drilling into this aspect of class analysis challenges the theoretical foundations and empirical explanatory power of existing class schemes further, rather than contributes to its reformulation.

Many of the predictions Marx made about the development of capitalism proved wrong in the course of history so far (cf. e.g. Dahrendorf 1959; Przeworski 2021). Wealth accumulated by top holders of productive capital relative to the rest, however, does not seem to be too far apart today from what he had in mind - judging from current empirical evidence. But what do

[2] A prominent example in the neo-Weberian tradition is the history of the long prominent neo-Weberian EGP class scheme in its first (Erikson, Goldthorpe, and Portocarrero 1979) and second (Erikson and Goldthorpe 1992) version, to the updated version by Oesch (2006). In the original article on which this introduction is based, I included a more detailed comparison and discussion of neo-Weberian vs. neo-Marxian class schemes and their theoretical foundations. However, in terms of space, I skip this discussion and focus on the neo-Marxist scheme for this introduction which has devoted more attention to the capitalist class and is, therefore, more applicable to the super-rich.

recent changes in capitalist dynamics and the resulting distribution of economic resources have to do with the study of class and elites today? One of the many consequences of capital concentration is that it naturally raises interest in those individuals and families that possess this wealth for the first time in possibly a century. As Suzanne Keller (1963:5) strikingly wrote: "in periods of rapid social change, the outline of […] elites against a shifting background is sharply visible. In periods of relative stability, these elites are merged with the objects, habits, and the manners of their age." When re-reading literature on class analysis from the past 70 years, it is striking how little developed the top of the capitalist class is theoretically as well as empirically to this day (cf. Manza and McCarthy 2011:163; Scott 2002:25), with only a few exceptions (e.g. Giddens 1972; Useem 1978; Wright 1976). To start looking for reasons for this, a relevant debate took place between so-called managerialists and Marxists in the 1970s and 1980s (cf. e.g. Scott 1990 for an overview). They surrounded the question of whether in the 20th century, corporations were more and more disconnected from the power of their owners. Other powerful actors within the firm gained relevance such as managers and financial shareholders, and more conflicts over power on firm decisions arose (cf. Allen 1981; Berle and Means 1991; Godechot 2005; Jensen and Meckling 1976; Mizruchi 2004; Mizruchi and Bey 2003; Scott 1997; Zeitlin 1989). Obviously, "A theory of class based on the division of society into owners and nonowners of means of production loses its analytical value as soon as legal ownership and factual control are separated […] any effective supersedure of Marx's theory of class has to start at this point." (Dahrendorf 1959:136). If one wanted to conclude these debates in some way, it seems clear that owners often – depending on firm governance and size - do not have full control of their corporations, and therefore capital, anymore as this may have been the case during Marx's lifetime (cf. Dahrendorf 1959:20–23; Offe 2006:52–53). This needs to be considered in an analysis of the modern bourgeoisie and their firms.

As one answer to this, Wright (1976, 1998), drawing from Poulantzas (1975), suggested that individuals can hold several and even contradictory class positions. His reasoning is based on a detailed description of how the control and ownership function differentiated in the 20th century based on three processes: "control over the physical means of production; control over labour power; control over investments and resource allocation. […] Again, it must be stressed that these three processes are the real stuff of class relations in capitalist society; they are not merely analytic dimensions derived from a priori reasoning." (Wright 1976:30) To be a member of the capitalist class, you need to have full control. This implies, that for Wright (1976:33), top managers belong to the bourgeoisie and hold contradictory class locations between being employed and at the same time controlling capital. However, the important distinction coming from this debate that I would like to point out here is ownership vs. control of capital. The

difference between top managers (the economic elite), and traditional capitalist families (the capitalist class) is that the former *only* may have full control over capital, while the latter owns significant shares of a company *and* may have full control over it. While this might seem like a small difference in terms of categorization, it becomes a larger problem about the question of who holds economic power, and especially who can transform it into political power. This question of who holds what kind of power is more central in the study of economic elites, and it is thereby where the study of class and elites intersect.

CHART 5
Levels within Ownership Relations

	RELATIONS OF ECONOMIC OWNERSHIP	RELATIONS OF POSSESSION	
		Control of means of production	*Control of labour power*
Full control	Control over the overall investment and accumulation process	Control over the entire apparatus of production	Control over the entire supervisory hierarchy
Partial control	Participation in decisions concerning either sub-units of the total production process or partial aspects of the entire investment process	Control over one segment of the total production process	Control over one segment of the supervisory hierarchy
Minimal control	Participation in decisions concerning narrow aspects of sub-units of production	Control over one's immediate instruments of production; some autonomy in the immediate labour process	Control over the direct producers, over immediate subordinates but not part of the hierarchy as such
No control	Complete exclusion from participation in investment and accumulation decisions	Negligible control over any aspect of the means of production	No ability to invoke sanctions on other workers

Figure 2: The three processes of control over capital taken from Wright (1976:31).

The study of elites is concerned with individuals in top positions in their field. This mostly implies privileged access to one or more resources such as power (Khan 2012). Elite sociology developed as an early challenge to Marx's idea of a classless society which claimed that

elites may be a necessity in all forms of societies regardless of capitalist development (Mosca 1939; Pareto 1936). Within the study of these elites since then, one of the cleavages is the level of skepticism toward their existence. One side claims that pluralist elite groups compete democratically for influence in capitalist democracies and that this is not a problem per se (Dahl 1958; Mannheim 1940). The other side of the spectrum claims that there is one powerful and possibly closed elite in many capitalist societies that concentrates disproportional amounts of power in favor of capital and reproduces itself (Domhoff 1967; Hunter 1953; Mills 1956; Scott 1990; Useem 1984).

An open question remains about how the super-rich are described from an elite perspective. There are two ways of understanding them. If they hold business assets and own and/or control capital through firms, they are traditionally described as economic elites. [3] These can be owners, managers, or both (owner-managers) who compete for power over a firm with other actors (Gourevitch and Shinn 2005; Scott 1990). A large body of literature also describes how these individuals connect across firms, e.g., by sitting on the same corporate boards. Through mechanisms such as these, it was argued that an "inner circle" of central managers acts according to the interests of business in general (Dreiling 2000; Mizruchi 1996; Stokman, Ziegler, and Scott 1985; Useem 1984). It was further argued that such networks have fractured since the beginning of the 2000s (Chu and Davis 2016; Mizruchi 2013), however, the persistent effects of such ties on corporate political action for example seem to remain (Banerjee and Murray 2020; Heerwig and Murray 2019a; Murray 2017). With the study of such cohesive elites and their political behavior, which can be traced back for example to the Mills vs. Dahl debate (Dahl 1958, 1961; Mills 1956), the theoretical focus turns away from describing which position groups hold within the economy and the effects of this position. The interest now is which group holds power in modern states.

So far, I have tried to show a link between increasing wealth concentration and the rise of the super-rich and the classic sociological literature of class analysis and elite sociology. As

[3] More recently, some scholars have also used the term "wealth elite" to describe the very top of the wealth distribution (e.g. Fernandez, Hofman, and Aalbers 2016; Keister and Lee 2017). This use of the term does not require any relation to ownership or control over capital such as business assets. It includes the idea of pure rentier families who are disconnected from all entrepreneurial activity. I find the term "wealth elites" misleading for two reasons. First, on the empirical side, the number and relevance of such families are questionable (Keister et al. 2021; Smith et al. 2019). Second, on the theoretical side, while this definition includes more individuals and families, in my view the term "wealth elite" merely remains a descriptive, and rather tautological, category which yet would have to be filled theoretically. I, therefore, concentrate on the former understanding of economic elites because it closely relates the study of the super-rich to questions of ownership and control of capital, i.e., business assets. At least when interested in power, I argue that this is at least one, if not the most relevant, channel that should not be obscured (as it would be when using the term "wealth elite"). This is at the same time the theoretical reason why the scope of this dissertation is on those families with an identifiable connection to business assets only.

outlined above referring to Lindblom, the concept of power[4] and how it relates to the super-rich, and the economic elite are the center of this dissertation. To evaluate this question, a missing link is how economic power stemming from control over capital may be the source of political power in capitalist democracies. Among the most interesting debates on capital and the state - from a Marxist perspective - were probably held around the 1970s. Gold et al. (1975) for example give a good overview of Marxist theories of the state at the time.[5] Especially the distinction between the instrumentalist power of the capitalist class, and the structural power of this class are interesting and helpful. The former suggests that members of the capitalist class intentionally try to capture the state by connecting to the political elite (Domhoff 1967; Miliband 1969; Sweezy 1951; Zeitlin 1989). The latter claims that the state is dependent on the economy in capitalism and that therefore the state is fundamentally acting in favor of capital without being instrumentally captured (e.g. Block 1977).[6]

Block (1977) made an argument challenging the instrumentalist view that the state reduces to a political "ruling class" simply governing in favor of capital, Block argues that in fact, two sets of mechanisms translate business interests into policies in capitalism. First, and less importantly, there is the direct influence of capital, i.e. through lobbying, legal action, ties to politicians such as favors, as well as the formation of policy groups and the like (Block 1977:12–14). Second, Block argues that the state and politicians depend on economic success, again for two reasons. First, the state needs economic prosperity to finance itself. Second, public support and resistance against a regime strongly depend on the state of the economy. In short, it is this structural dependence that leads to policies that favor capital. Moreover, the state has an interest in a positive outlook of capitalists to keep investments at a high level and secure accumulation.[7] This increases the likelihood of policies in the general interest of capital even more (Block 1977:14–15). Following this suggestion from Marxist debates on the power of the capitalist class, the next section develops two working hypotheses to evaluate the results produced in the individual parts of this dissertation.

---

[4] I do not discuss the concept of power in theoretical detail here in terms of scope. The mentioned notions of influence on politics as described by Lindblom (1982) or Block (1977) but also Mills (1956) and Dahl (1958) is in my view sufficient for the scope of this dissertation. In short, power is therefore defined, or rather reduced to if you prefer, individuals or groups in the top of distributions of different economic resources (wealth, income, and control over capital) successfully using these to realize or maintain their interests in any political area, in cases of conflicts of interest, in the long- or short-run. I refer the reader to Lukes (2004) for a succinct while in-depth theoretical discussion of the underlying concept of power.

[5] For a more recent overview see Barrow (1993).

[6] The two traditions were later also taken up but redeveloped in Marxist analyses of the state and its mechanisms (Offe 2006; Therborn 1978).

[7] A very similar argument about the interdepence of the state and entrepreneurs was already made by Kalecki (1943:328): "The entrepreneur remains the medium through which the intervention is conducted. If he does not feel confidence in the political situation, he will not be bribed into investment."

3. A critical number of businesspeople: The "old families", the "inner circle" or both?

As outlined above, two competing theoretical concepts can be used to understand those individuals with power in the economy. Either it is super-rich capitalists, the "old families" who might hold dynastic wealth over generations, who own and control their corporations and may use them as vehicles to influence the political process. Their goals could be in line with the economic interests of accumulation of the specific fraction of capital they hold and wealth dynasty management (cf. e.g. Gilding 2005), but they may also be after their ideological agendas on non-economic issues (e.g. Pierson 2017; Skocpol and Hertel-Fernandez 2016). Alternatively, there might be social networks of economic elites, i.e., employed managers of large corporations, densely connected who act in the interest of capital more generally (e.g. Comet 2019; Dreiling 2000; Heerwig and Murray 2019b; Useem 1984). Broken down to one sentence, the empirical question asked here is whether ownership of capital *in addition* to controlling capital makes a difference in terms of (coordinated) corporate political action in capitalist democracies, and how relevant the two groups are. To evaluate this empirical question with relevant theoretical implications, I develop working hypotheses for this chapter. The articles presented in this cumulative dissertation will serve as an empirical basis to evaluate these hypotheses. I present them in the next section before a discussion evaluates how their results speak to the developed hypotheses and the concepts of classes and elites for the study of individuals with economic power today.

The first working hypothesis is most informed by the structuralist Marxist side but considers both - instrumental and structural power. To describe the capitalist class appropriately, it is key to consider its structural economic positions in terms of (conscious or unconscious) integration, and potential for structural power. It is these positions that shape actors' interests in the original meaning of class. The idea of structural power rather obviously implies that some actors within the capitalist class may be more important than others. To recall Lindblom's (1982:330; emphasised by me) quote: "In just what aspects of political/economic life that mechanism operates I have not yet said, except to note that the included aspects are all those in which businessmen - *or any large or critical number of them* - see change as hurtful to their own prospects." His distinction between large and critical may imply what I would like to bring forward here too: naturally, it is not so much their number but how much relative corporate capital they represent.[8] Actors within the capitalist class differ in their structural leverage. In other

[8] That is in sum – and with consistent interests. Which is easier the smaller the number or, in other words, the higher the concentration of capital.

words, the dimension of control over a relative share of capital represented by a member of the capitalist class is stressed as bridging between economic position and potential for structural power. From my view, this is what is needed to assign Warren Buffet and the owner of a bookstore in town their more appropriate class positions. This view has implications for both sets of mechanisms of structural power brought forward by Fred Block. Concerning direct influence, this means that those individuals and groups from the capitalist class should also have a higher propensity to be heard by politicians and the state. A member of parliament will more likely pick up the phone or make space in her timetable when a manager or owner weighing billions worth of capital calls for it, than for the average request. What follows from this is that I expect this group, the classical capitalists who control *and* own the largest amounts of capital, to be especially important for corporate political action:

H1: Super-rich capitalists are more relevant political actors in favor of capital than the inner circle.

The alternative idea could be that the separation of ownership and control led to the decline of the family firm and the classic model of owning families in line with Berle and Means (1991). Managers may be relatively independent of owners and unite in inner circle communities, subject to business scan and other mechanisms of assimilation and coordination. They are the agents of capital and more politically active, and politically more relevant in general, than the classic capitalists. Empirically, this finding would only to a limited extent be new or surprising. As mentioned above, some observers claim the decline of the inner circle, while others show their continued relevance for corporate political action. The contribution here would be to show whether this group is more or less relevant than economic elites who also *own* significant amounts of capital, i.e. super-rich capitalists. The hypothesis following would be:

H2: The inner circle is the more relevant political actor in favor of capital, compared to super-rich capitalists.

The so-far developed hypotheses work with an antagonism between ownership and control, between managers and owners, and between an economic elite of managers and traditional capitalists. However, the story could of course also be more complex or less complex. It could be more complex if the two groups might act differently but fulfill a division of labor both acting politically in the interest of capital, but in different ways (Lunding, Ellersgaard, and Larsen 2021; Soref 1976; Useem 1978). From this it would follow:

H3: The inner circle and super-rich capitalists both act politically in favor of capital but in different ways.

Of course, the story could also be less complex. This would follow the categorization of Wright (1976, 1998) and both, top managers and traditional super-rich capitalists have the same class positions and act in the same way politically in the name of the specific form of capital that they represent.

H4: Top managers and super-rich capitalists act politically in favor of capital in the same way.

Finally, there is of course also the large alternative hypothesis that the two groups should be seen as rather independent actors with separate interests and behavior. This would follow mostly the corporate governance literature and the idea of managers as agents of owners:

H5: The inner circle and the super-rich are two distinct groups with separate interests and diverging (political) behavior and can therefore not be seen as a class.

This dissertation offers the first attempt to test these ideas on a large-scale database and for the contemporary period around 2020. While the analyses cannot fully answer these questions and only offer a start to rule out some of these hypotheses, they still are the larger theoretical context and motivation in which the thesis must be seen. The next section briefly presents the four articles making up the dissertation, and what their results tell us about the presented hypotheses.

4. Measuring political leverage and action of the super-rich, the corporate elite, and their firms in 2020

The cumulative dissertation is made up of one article describing the data source, what necessary adjustments needed to be made, a critique of its wide range of shortcomings, and some examples of applied solutions, as well as open problems. This technical article is followed by three substantial articles on the issues outlined above. All four articles are outlined in the following.

1st Article: When should we believe research using the ORBIS firm data?

This article is a critical examination of the largest and most complex data source used for this dissertation. One of its purposes is to collaborate and present the numerous challenges when

using the data source for research, as well as some suggested solutions and not least the preconditions necessary to derive valid results from the data. The article picks up a discussion initiated by scholars engaging with using big (network) data for social scientific research (Heemskerk et al. 2018) and takes the position that the social science research community also needs standards when a data source is too prone to error to the inexperienced user. The article, therefore, offers both, a concrete hands-on manual for researchers and reviewers working with the ORBIS data, and it also raises and discusses more meta-methodological questions. Although the first article does not directly contribute to the theoretical concerns outlined above, it nevertheless presents the fundamental basis on which the other articles give answers to them. After all, if I would have to put numbers to it, about 60-80% of the workload devoted to this dissertation project was spent on downloading, exploring, preparing, and describing the ORBIS data, and matching it to the used rich list provided by Bornefeld (2019; Neßhöver and Bornefeld 2018).[9] The project is overall based on more than 30,000 lines of Python code. This first article, therefore in summary explains the data characteristics on which the following articles are based, including some necessary details to fully evaluate the limitations of the results presented in the articles.

2nd Article: The Peak of Capital? Super-Rich Capitalist Families in Global Interlocking Directorate Networks

The second and first substantial article focuses on the coordination between the two groups of interest: the corporate elite and the capitalist class. It analyzes probably the most classic and widely studied channel through which coordination among powerful individuals has taken place. Namely, this is shared board membership of individuals on management boards of corporations, and the network between firms and individuals that are spanned in this way (David and Westerhuis 2014; Jeidels 1905; Koenig and Gogel 1981; Mizruchi 1996; Stokman et al. 1985). A reformulation of the theoretical frame of this dissertation from a network perspective would be the question of connectedness *within* and *across* the two groups of the corporate elite and the super-rich. This question is analyzed based on networks of interlocks.

The idea of a capitalist class in a Marxist sense presupposes that class members should be more likely to coordinate via class formation to take political action because they have the same economic interests due to their shared economic positions (e.g. Wright 1998:10). The idea of an inner circle by definition requires a highly interlocked circle of managers at the center of the network (Useem 1984). For the most part, the paper locates the super-rich within global, i.e.,

[9] These important steps of the analysis are presented in some detail in the appendices of the 2nd and 3rd articles. The matched rich list is described in the main texts of these articles.

national and transnational, corporate elites, and therefore ties between the two groups. Findings suggest that a relevant share, especially the largest and oldest fortunes, is part of the largest component of the global network of interlocking directorates. This speaks for the entanglement of the two groups. The most relevant empirical test is to focus on super-rich capitalist family members and to ask whether these are more likely to be connected among each other than to other managers. This is an easy though strict test to the question of unity among the capitalist class. Wright suggests that the capitalist class consists of all individuals with control over capital, i.e., top managers and owners of large corporations. However, traditional capitalists in a way have the purer class position because they fully own the capital they control, while managers have a contradictory class location having control on the one side but being employed on the other (Wright 1998:42–51). In my interpretation, this implies that the two subgroups are homogenous in terms of their class positions and consequently, due to class formation, more likely to be connected among each other. Several scholars suggested cohesion among the richest capitalist families as a class (e.g. Sweezy 1951; Zeitlin 1989). In contrast to this, results from regression analyses suggest that super-rich family members are not more likely to sit on the same boards with super-rich family members from other rich families than with other managers in the network. While this result can be interpreted in different ways, it speaks against coordination between the owning families of large firms through this channel. One could object that coordination might not take place between super-rich family members directly but through their employees. This claim would predict that firms with super-rich family members on board should be closer to each other. Yet, this is also not the case as tested in a separate regression analysis on the firm level.

Concluding these findings, results challenge some of the hypotheses developed above while some would need further analysis to be ruled out. More precisely, there is not much evidence that the super-rich are the central group in corporate elite interlock networks. While they are more likely to be part of the network, they are not more likely to be connected among each other. Hence, it rather seems the primary characteristic is having *control* over capital, i.e., being embedded in the corporate elite, rather than *owning* capital. This implies that at least no relevant coordination or class formation among super-rich families takes place through this channel, and hypothesis H1, that the super-rich are the most relevant group of individuals with power over capital coordinating and acting in favor of it, must be rejected (for this channel). There is some reason to keep believing in hypothesis H2, that corporate elite networks exist, and that there is relevant coordination among the managers of the largest corporations, especially at the national level. The super-rich are only part of this network but do not dominate it. H2 can therefore not be rejected. It remains relatively unclear, however, what the embeddedness of the

super-rich in this network tells us. It could either be that the super-rich fulfill a specific function in the division of labor among the corporate elite (H3), or it could be that those super-rich individuals who are embedded in the corporate elite network act the same as other elite managers (H4). This question is investigated and discussed to some extent based on the next article. Results do rather not speak in favor of hypothesis H5 that both groups are separate from each other. The conclusion from the empirical results so far is a primacy of the elite perspective and the relevance of control, or rather the irrelevance of ownership for cohesion through networks of interlocking directorates. It should be stressed here once more that this article only speaks about the channel of interlocking directorates as one possible channel of coordination among several. Limitations and problems with this channel as raised in the literature, as well as possible alternative channels, are discussed in more detail in the paper.

3rd Article: Linking Wealth and Power. Direct Political Action of Corporate Elites and the Wealthiest Capitalist Families in the US and Germany.

The third article offers a second empirical contribution to evaluate the relevance of, and relationships between, the capitalist class and the economic elite. It attempts to explain the variation in political action of corporate elites, the capitalist class, as well as their firms in the US and Germany. Three "avenues of influence" (Bonica 2016) are analyzed: corporate party contributions, corporate lobbying, and individual party contributions. Case selection used data sources, and other aspects are discussed further in the article. The relevant empirical question for the theoretical concerns raised above is how similar the political action of the two groups is in the US and Germany through these avenues. The developed results make clear that national variation, or institutional variation, matters. The two groups seem to work differently and have different magnitudes in the two countries. The political action of the inner circle for example seems irrelevant in Germany while it is more relevant in the US. The super-rich do however play a role in both countries – in Germany more through their firms, and in the US especially as individual donors.

Possibly the most relevant finding is that political participation of both groups is relatively low, and, in comparison to the total sum of all money of business in politics, its amount is negligible in almost all three avenues in the two countries. The only exception is donations by super-rich individuals in the US. On top of that, the odds estimated from regression analysis suggest that different forms of super-rich involvement increase the

odds of a firm to lobby, making a party contribution, or having a corporate PAC. The coefficients for the super-rich seem to be larger than those for the inner circle in the US, while in Germany the inner circle, operationalized through interlocks, does not play any role in political action through these channels. There seems to be a tendency in both countries, that the more control a super-rich family has over a firm, the more likely this firm is to take political action through either of the two channels. All of this rather clearly speaks for the higher relevance of the super-rich as political actors in the name of capital and therefore supports H1. Again, it is difficult from these analyses to interpret whether the super-rich and the inner circle are members of one class and act similarly, or in different ways. But I think more evidence speaks for hypothesis H5, that the two groups act very differently and there is not much explanatory power in treating them as one class. On top of the different effects, there are also pressing differences in ideology when measured as partisan donations. The inner circle in the US seems a lot more liberal than the super-rich. In Germany, the inner circle does not engage at all. All of this shows clear differences, between the two groups' behaviors, interests, and their relevance for the political process, and questions the analytical value of the concept of an inner circle in Germany, and that of a capitalist class as a whole in both countries.

To sum up, the study of direct political action of the capitalist class and the corporate elite shows relevant differences between the two groups. It is especially the differences in partisan contributions that in my view speak most clearly against the idea that the two groups are part of the same class. This is also not implausible since those who control *and* own wealth and especially business assets, politically are more dependent on conservative parties' economic policies which tend to be strictly against exorbitant inheritance taxes, not to speak of wealth taxes. Also as Burris (2000) argued before, super-rich capitalists will also be more likely to come from upper-class backgrounds and therefore from conservative families. The corporate elite as top labor income earners however might be more in line with socially liberal parties and also their social values. Previous literature suggests that this may vary by sector (Broockman, Ferenstein, and Malhotra 2019). Overall, not much evidence presented here speaks in favor of keeping the concept of capitalist class action, but rather to see owners and the corporate elite as separate sociological groups.

4th Article: Finance Capital 2.0? Shareholder Structure and Institutional Capital Pools. *With Benjamin Braun and Sebastian Kohl.*

This article was written together with Benjamin Braun and Sebastian Kohl as coauthors. My contribution was mainly the construction of the underlying dataset and I provided cleaned versions of the ORBIS data, and continuously fixed identified remaining data issues in steady conversation with the coauthors. I also contributed the network illustration of capital relations between regions. The regression analyses were conducted by Sebastian Kohl, and most of the text and figures based on the prepared datasets were produced by Benjamin Braun. The paper uses firm-level data to describe the shareholder structures of OECD countries. It engages with literature in comparative political economy to show that shareholder structures and concentration are relatively independent of "national developmental paths and the composition and financing needs of the non-financial sector" (Arndt, Braun, and Kohl 2023:2).

As a final contribution to this cumulative dissertation, the main function this paper serves is to locate the magnitude of capital held by super-rich families at the macro level, in comparison to other shareholder types. This article, therefore, switches perspectives away from the wealth inequality and concentration literature which centers around the wealth of all individuals within a national population. Instead, it asks: "who owns the economy" and which groups are relevant owners of the world's largest corporations? Following the argument on structural power made above, the share of capital controlled is a relevant dimension to understand how much leverage for structural power individual capitalists, as well as the whole group, have in a country. While this paper focuses on ownership and disregards control, and also only analyzes listed firms, the share of capital owned does give a proxy of how much capital can be controlled by a group of shareholders. Results challenge the idea of an overall dominance of the capitalist class in the traditional sense. Even when accounting for indirect channels such as foundations or more complex holding structures, the share of listed capital owned by individuals and families directly is relatively low in most, and especially in the largest OECD countries. Even in Germany, a country with a relatively large share of super-rich families among shareholders and a large *Mittelstand* (Berghoff 2006; Lehrer and Celo 2016), the share owned by super-rich families is less than 10% of listed firms. Other more important shareholder types have evolved and in combination with the results of the other articles, this challenges the overall relevance of the super-rich in terms of potential leverage of structural power. This is also because it implies that most of the listed, and therefore largest, companies are not owned by super-rich families. It speaks in favor of newer conceptualizations of power over large corporations by finance, for example, developed by Fligstein (1987), or more specifically institutional investors, as promoted

in newer contributions by Useem (e.g. Useem and Gager 1996), or as the incredible rise of asset managers argued very recently by Braun (2021).

5. Conclusion

The theoretical frame of this dissertation surrounds the question of how to understand individuals with power in the economy from the perspective of sociological theory, and which of these groups can we assume to have power over politics. Two classic concepts were suggested to describe the groups of the corporate elite with control over the economy as hired top managers of large corporations and, second, super-rich capitalist families as traditional capitalists owning *and* controlling capital through owned firms. The theoretical discussion suggested that while it is control over capital that could be used to translate economic to political power, ownership of capital might play a distinct role. The four alone-standing articles that make up this dissertation were presented, and their relation and contribution to this theoretical frame were discussed. Overall, the evidence developed and presented here challenges the explanatory power of the idea of a capitalist class comprising managers and owners in 2020. Article 2 shows that while both groups are connected in networks of interlocking directorates, it does not seem to be the case that super-rich capitalist families are more likely to be connected among each other. It hence rather seems that parts of the super-rich are embedded in the corporate elite, than the other way around. Article 3 then shows that super-rich families are overall more relevant in terms of political action, while their behavior and magnitude vary by national context. Results speak against the idea of similar political action due to the shared class interests of the two groups. Rather, there are ideological differences, especially in the US. In Germany, the super-rich seem to act politically at least to some degree, while the inner circle is entirely inactive – at least through the analyzed channels.

The idea of capitalist class political action is more than one hundred years old but still regularly used in social science research. This dissertation started with a classic Marxist concept to describe a contemporary case but had to conclude with the insufficiency of the concepts' explanatory power today. The evidence does not give support for the idea of a capitalist class uniting owning super-rich families, and top managers due to their shared class position and economic interests. In my view, the gathered empirical findings and general state of the literature put future research at a crossroads of one path which was started by Robert Dahl and what Offe (1972:74) calls the "integration theorists", and another following the tracks of Colin Wright Mills

and the "conflict theorists".[10] In other words, a decision future research needs to make is whether it wants to stick with, further adapt, and defend the idea of groups whose power is based on economic resources that somehow exert disproportionate power in modern capitalist democracies and thereby are the agents of capitalist reproduction (Domhoff 1967; Mills 1956; Sweezy 1951). Alternatively, it could start from the idea of plural elites competing for influence in a relatively equal competition (Dahl 1958, 1961; Dahrendorf 1959).

Those researchers deciding to stick with ideas from the ruling elite model, or conflict theorists, could either analyze other channels of capitalist class coordination and political action and influence that were not studied here as suggested in the individual articles. There are also more theoretical questions to be asked, for example by picking up Marx's idea of a class-for-itself vs. a class-in-itself. Are there any indications that the super-rich and the inner circle simply do not have to work together at this point in history, but maybe they would if more fundamental issues were at stake, e.g., what Block (1977:14) calls "anti-capitalist policies"? In Germany, the broad public debates surrounding the ideas of dispossessing car maker BMW (Bittner and Hildebrandt 2019) or real estate companies in Berlin (dwenteignen.de 2023) and potential reactions of the capitalist class – however defined - could be interesting cases for this. This could also be one way to understand the observed disinterest of both groups in politics that was described for the US by Mizruchi (2013), and that also resembles some of the results presented here for Germany. Alternatively, one could work out on a broader empirical basis, e.g. by exploiting other data than interlocking directorates (cf. e.g. Larsen and Ellersgaard 2017), that the super-rich and the corporate elite are two distinct classes in themselves and for themselves. This would require more theoretical groundwork in developing a contemporary class scheme that does pay enough attention to these top classes with control over a relevant share of capital, following the work of Wright and others (Manza and McCarthy 2011; Wright 1998, 2015).

However, after finishing this project, I expect more explanatory power in analyzing the potential power of capitalist families theoretically and empirically as part of corporate elites – at least at this point in history. This would follow more the perspective offered by the pluralists and analyze the conditions under which capitalist families or the corporate elite do engage in politics at all – and what they want to achieve. The interesting question to me seems to establish when elite unity and composition matters and for which macro-outcomes, rather than assuming it does and only describe tautologically defined elite groups in all imaginable kind of ways. In other words, current elite sociology should be confronted with Dahl's (1958) critique of the ruling elite

[10] For the sake of the argument, I simplify the different traditions here and subsume the idea of a capitalist class in the Marxist sense and the possibility to stick to it under the ruling elite model albeit they do have different theoretical foundations but similar interests (cf. e.g. Offe 1972; Sweezy 1956).

model once again, and I think the super-rich and the corporate elite today would offer a very interesting setting to put his critique to the test. Understanding power of super-rich capitalists and top managers requires a more systematic current observation of their political behavior and interests. Who among the super-rich participates in elite networks and who does not and why? Who is more related to politics, managers, or owners? And through which channels? One way to approach this is to change directions of study and interview top political office holders from the regional to the federal level and hear their impressions of if and how individuals representing business – the agents of capital – approach the political process. Either way, there is a lot of work to be done to provide economic elite studies with a theoretical basis, and to fully understand elites' power potential. The central question seems to me: what are the sociological micro-foundations of capitalist reproduction on the side of the economy, if not driven by behavior of the traditional owners and managers of the large corporations?

References

Allen, Michael Patrick. 1981. "Power and Privilege in the Large Corporation: Corporate Control and Managerial Compensation." *American Journal of Sociology* 86(5):1112–23.

Arndt, Lukas, Benjamin Braun, and Sebastian Kohl. 2023. ""Finance Capital 2.0? Global Varieties of Shareholder Concentration."" Cologne and Berlin.

Banerjee, Tarun, and Joshua Murray. 2020. "Class Dominance or Fracturing? Sources of Broad Interest in Lobbying by Fortune 500 Corporations." *Sociological Perspectives* 64(2):157–75. doi: 10.1177/0731121420921883.

Barrow, Clyde W. 1993. *Critical Theories of the State. Marxist, Neo-Marxist, Post-Marxist*. Madison: The University of Wisconsin Press.

Berghoff, Hartmut. 2006. "The End of Family Business? The Mittelstand and German Capitalism in Transition, 1949-2000." *The Business History Review* 80(2):263–95.

Berle, Adolf A., and Gardiner Means. 1991. *The Modern Corporation and Private Property*. New York: Routledge.

Bittner, Jochen, and Tina Hildebrandt. 2019. "Was Heißt Sozialismus Für Sie, Kevin Kühnert?" Retrieved March 27, 2023 (https://www.zeit.de/2019/19/kevin-kuehnert-spd-jugendorganisation-sozialismus).

Block, Fred. 1977. "The Ruling Class Does Not Rule: Notes on the Marxist Theory of the State." *Socialist Revolution* 33:6–28.

Bonica, Adam. 2016. "Avenues of Influence: On the Political Expenditures of Corporations and Their Directors and Executives." *Business and Politics* 18(4):367–94. doi: 10.1515/bap-2016-0004.

Bornefeld, Andreas. 2019. "Vermögensforschung." *Https://Www.Reichsten.de/En/Index*. Retrieved April 26, 2019 (https://www.reichsten.de/en/index).

Braun, Benjamin. 2021. "Asset Manager Capitalismas a Corporate Governance Regime." Pp. 270–94 in *American Political Economy: Politics, Markets, and Power.*, edited by J. S. Hacker, A. Hertel-Fernandez, P. Pierson, and K. Thelen. New York: Cambridge University Press.

Broockman, David E., Gregory Ferenstein, and Neil Malhotra. 2019. "Predispositions and the Political Behavior of American Economic Elites: Evidence from Technology Entrepreneurs." *American Journal of Political Science* 63(1):212–33. doi: 10.1111/ajps.12408.

Burris, Val. 2000. "The Myth of Old Money Liberalism: The Politics of the Forbes 400 Richest Americans." *Social Problems* 47(3):360–78. doi: 10.2307/3097235.

Chu, Johan S. G., and Gerald F. Davis. 2016. "Who Killed the Inner Circle? The End of the Era of the Corporate Interlock Network." *American Journal of Sociology* 122(3):714–54. doi: 10.2139/ssrn.2061113.

Comet, Catherine. 2019. "How Does the Inner Circle Shape the Policy-Planning Network in France?" *Socio-Economic Review* 17(4):1021–41. doi: 10.1093/SER/MWX051.

Dahl, Robert A. 1958. "A Critique of the Ruling Elite Model." *American Political Science Review* 52(02):463–69. doi: 10.2307/1952327.

Dahl, Robert A. 1961. *Who Governs? : Democracy and Power in an American City*. New Haven, Conn.: Yale University Press.

Dahrendorf, Ralf. 1959. *Class and Class Conflict in Industrial Society*. Stanford: Stanford University Press.

David, Thomas, and Gerarda Westerhuis. 2014. *The Power of Corporate Networks : A Comparative and Historical Perspective*. London: Routledge.

Domhoff, G. William. 1967. *Who Rules America?* Englewood Cliffs N.J.: Prentice Hall.

De Donder, Philippe, and John E. Roemer. 2015. *The Dynamics of Capital Accumulation in the US: Simulations after Piketty*. Munich.

Dreiling, Michael C. 2000. "The Class Embeddedness of Corporate Political Action: Leadership in Defense of the NAFTA." *Social Problems* 47(1):21–48. doi: 10.2307/3097150.

dwenteignen.de. 2023. "Deutsche Wohnen & Co. Enteignen - Über Uns." Retrieved March 27, 2023 (https://dwenteignen.de/ueber-uns).

Ederer, Stefan, Maximilian Mayerhofer, and Miriam Rehm. 2020. "Rich and Ever Richer? Differential Returns across Socioeconomic Groups." *Journal of Post Keynesian Economics* 1–19. doi: 10.1080/01603477.2020.1794902.

Erikson, Robert, and John Goldthorpe. 1992. *The Constant Flux. A Study of Class Mobility in Industrial Societies*. Oxford: Oxford University Press.

Erikson, Robert, John H. Goldthorpe, and Portocarrero. 1979. "Intergenerational Class Mobility in Three Western European Societies: England, France and Sweden." *British Journal of Sociology* 30:415–41.

Fligstein, Neil. 1987. "The Intraorganizational Power Struggle : Rise of Finance Personnel to Top Leadership in Large Corporations , 1919-1979." *American Sociological Review* 52(1):44–58.

Forbes. 2020. "World's Billionaires List. The Richest in 2020." *Forbes Magazine*. Retrieved April 24, 2020 (https://www.forbes.com/billionaires/).

Formisano, Ron. 2017. *American Oligarchy: The Permanent Political Class*. University of Illinois Press.

Freund, Caroline, and Sarah Oliver. 2016. ""The Origins of the Superrich: The Billionaire Characteristics Database."" Peterson Institute for International Economics, Washington D.C.

Giddens, Anthony. 1972. "Elites in the British Class Structure." *The Sociological Review* 20(3):345–72. doi: 10.1111/j.1467-954X.1972.tb00214.x.

Gilding, Michael. 2005. "Families and Fortunes: Accumulation, Management Succession and Inheritance in Wealthy Families." *Journal of Sociology* 41(1):29–45. doi: 10.1177/1440783305050962.

Godechot, Olivier. 2005. "La Finance, Avant-Garde Du Prolétariat? Les Salariés de La Finance et La Structure de Classe." *Carnets de Bord* 10:55–66.

Góes, Carlos. 2016. "Testing Piketty's Hypothesis on the Drivers of Income Inequality: Evidence from Panel VARs with Heterogeneous Dynamics." *IMF Working Papers* 16(160):1. doi: 10.5089/9781475523249.001.

Gold, David A., Clarance Y. H. Lo, and Erik Olin Wright. 1975. "Recent Developments in Marxist Theories of the Capitalist State." *Monthly Review* 27:29–43, 46–51.

Gourevitch, Peter A., and James Shinn. 2005. *Political Power and Corporate Control: The New Global Politics of Corporate Governance*. Princeton, NJ: Princeton University Press.

Heemskerk, Eelke M., Kevin Young, Frank W. Takes, Bruce Cronin, Javier Garcia-Bernardo, Lasse F. Henriksen, William Kindred Winecoff, Vladimir Popov, and Audrey Laurin-Lamothe. 2018. "The Promise and Perils of Using Big Data in the Study of Corporate Networks: Problems, Diagnostics and Fixes." *Global Networks* 18(1):3–32. doi: 10.1111/glob.12183.

Heerwig, Jennifer A., and Joshua Murray. 2019a. "The Political Strategies and Unity of the American Corporate Inner Circle: Evidence from Political Donations, 1982–2000." *Social Problems* 66(4):580–608. doi: 10.1093/socpro/spy014.

Heerwig, Jennifer A., and Joshua Murray. 2019b. "The Political Strategies and Unity of the American Corporate Inner Circle: Evidence from Political Donations, 1982–2000." *Social Problems* 66(4):580–608. doi: 10.1093/socpro/spy014.

Hunter, Floyd. 1953. *Community Power Structure*. Chapel Hill: The University of North Carolina Press.

Jeidels, Otto. 1905. *Das Verhältnis Der Deutschen Grossbanken Zur Industrie Mit Besonderer Berücksichtigung Der Eisenindustrie*. Leipzig: Duncker & Humblot.

Jensen, Michael C., and William H. Meckling. 1976. "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure." *Journal of Financial Economics* 3(4):305–60. doi: 10.1016/0304-405x(76)90026-x.

Kalecki, Michal. 1943. "Political Aspects of Full Employment." *The Political Quarterly* 14(4):322–31.

Keister, Lisa A., Mingxiang Li, and Hang Young Lee. 2021. "Do You Need Business Assets to Be Rich?" *Socius: Sociological Research for a Dynamic World* 7:1–3. doi: 10.1177/23780231211031684.

Keller, Suzanne. 1963. *Beyond the Ruling Class. Strategic Elites in Modern Society*. New York: Random House.

Khan, Shamus Rahman. 2012. "The Sociology of Elites." *Annual Review of Sociology* 38:361–77. doi: 10.1146/annurev-soc-071811-145542.

King, J. E. 2017. "The Literature on Piketty." *Review of Political Economy* 29(1):1–17. doi: 10.1080/09538259.2016.1173425.

Koenig, Thomas, and Robert Gogel. 1981. "Interlocking Corporate Directorships as a Social Network." *The American Journal of Economics and Sociology* 40(1):37–50.

Korom, Philipp. 2015. "Elites: History of the Concept." Pp. 390–95 in *International Encyclopedia of the Social & Behavioral Sciences*. Amsterdam, Oxford, Waltham: Elsevier.

Larsen, Anton Grau, and Christoph Houman Ellersgaard. 2017. "Identifying Power Elites—k-Cores in Heterogeneous Affiliation Networks." *Social Networks* 50:55–69. doi: 10.1016/j.socnet.2017.03.009.

Lehrer, Mark, and Sokol Celo. 2016. "German Family Capitalism in the 21st Century: Patient Capital between Bifurcation and Symbiosis." *Socio-Economic Review* 14(4):729–50. doi: 10.1093/ser/mww023.

Lindblom, Charles E. 1982. "The Market as Prison." *Journal of Politics* 44(2):324–36.

Lukes, Steven. 2004. *Power. A Radical View*. Houndmills and New York: Palgrave Macmillan.

Lunding, Jacob Aagaard, Christoph Houman Ellersgaard, and Anton Grau Larsen. 2021. "The Established and the Delegated: The Division of Labour of Domination among Effective Agents on the Field of Power in Denmark." *Sociology* 55(1):110–28. doi: 10.1177/0038038520928220.

Madsen, Jakob B. 2018. ""Is Inequality Increasing in R-G ? The Dynamics of Capital ' s Income Share in the UK , 1210-2013."" Hitotsubashi Institute for Advanced Study, Hitotsubashi University, Tokyo.

Mannheim, Karl. 1940. *Man and Society in the Age of Reconstruction*. New York: Harcourt.

Manza, Jeff, and Michael A. McCarthy. 2011. "The Neo-Marxist Legacy in American Sociology." *Annual Review of Sociology* 37(1):155–83. doi: 10.1146/annurev-soc-081309-150145.

Marx, Karl. 1965. *Der 18. Brumaire Des Louis Bonaparte*. Frankfurt am Main: Insel Verlag.

Milanović, Branko. 2016. *Global Inequality a New Approach for the Age of Globalization*. 1. print. Cambridge, MA: Belknap Press of Harvard Univ. Press.

Milanović, Branko. 2017. "Increasing Capital Income Share and Its Effect on Personal Income Inequality." Pp. 235–58 in *After Piketty. The Agenda for Economics and Inequality.*, edited by H. Boushey, J. B. DeLong, and M. Steinbaum. Cambridge; London: Harvard University Press.

Miliband, Ralph. 1969. *The State in Capitalist Society*. London: The Camelot Press.

Mills, C. Wright. 1956. *The Power Elite*. Vol. 20. New York: Oxford University Press.

Mizruchi, Mark S. 1996. "What Do Interlocks Do? An Analysis, Critique, and Assessment of Research on Interlocking Directorates." *Annual Review of Sociology* 22(1):271–98. doi: 10.1146/annurev.soc.22.1.271.

Mizruchi, Mark S. 2004. "Berle and Means Revisited: The Governance and Power of Large U.S. Corporations." *Theory and Society* 33:579–617.

Mizruchi, Mark S. 2013. *The Fracturing of the American Corporate Elite*. Cambridge: Harvard University Press.

Mizruchi, Mark S., and Deborah M. Bey. 2003. "Corporate Control, Interfirm Relations, and Corporate Power." Pp. 310–30 in *The Handbook of Political Sociology*, edited by T. Janoski, R. R. Alford, A. M. Hicks, and M. A. Schwartz. Cambridge: Cambridge University Press.

Mosca, Gaetano. 1939. *The Ruling Class*. New York and London: McGraw-Hill.

Murray, Joshua. 2017. "Interlock Globally, Act Domestically: Corporate Political Unity in the 21st Century." *American Journal of Sociology* 122(6):1617–63. doi: 10.1086/691603.

Neßhöver, Christoph, and Andreas Bornefeld. 2018. "Die 1001 Größten Vermögen [The 1001 Largest Fortunes]." *Manager Magazin*.

Oesch, Daniel. 2006. *Redrawing the Class Map*. Basingstoke and New York: Palgrave Macmillan.

Offe, Claus. 1972. "Political Authority and Class Structures — An Analysis of Late Capitalist Societies." *International Journal of Sociology* 2(1):73–108.

Offe, Claus. 2006. *Strukturprobleme Des Kapitalistischen Staates*. Frankfurt/ New York: Campus.

Page, Benjamin I., Jason Seawright, and Matthew J. Lacombe. 2018. *Billionaires and Stealth Politics*. Chicago and London: The University of Chicago Press.

Pareto, Vilfredo. 1936. *The Mind and Society*. London: Jonathan Cape.

Pierson, Paul. 2017. "American Hybrid: Donald Trump and the Strange Merger of Populism and Plutocracy." *The British Journal of Sociology* 68(S1):S105–19. doi: 10.1111/1468-4446.12323.

Piketty, Thomas. 2014. *Capital in the Twenty-First Century*. Cambridge: Belknap Press of Harvard Univ. Press.

Poulantzas, Nicos. 1975. *Classes in Contemporary Capitalism*. London: NLB.

Przeworski, Adam. 2021. "What Have I Learned from Marx and What Still Stands?" *Politics & Society* 49(4):433–50. doi: 10.1177/0032329220958662.

Scott, John, ed. 1990. *The Sociology of Elites*. Vol. I–III. Aldershot, UK; Brookfield, USA: Edward Elgar.

Scott, John. 1997. *Corporate Business and Capitalist Classes*. Oxford; New York: Oxford University Press.

Scott, John. 2002. "Social Class and Stratification in Late Modernity." *Acta Sociologica* 45(1):23–35. doi: 10.1080/00016990252885771.

Skocpol, Theda, and Alexander Hertel-Fernandez. 2016. "The Koch Network and Republican Party Extremism." *Perspectives on Politics* 14(3):681–99. doi: 10.1017/S1537592716001122.

Soref, Michael. 1976. "Social Class and a Division of Labor within the Corporate Elite: A Note on Class, Interlocking, and Executive Committee Membership of Directors of U. S. Industrial Firms." *Sociological Quarterly* 17(3):360–68. doi: 10.1111/j.1533-8525.1976.tb00989.x.

Stokman, Frans N., Rolf Ziegler, and John Scott, eds. 1985. *Networks of Corporate Power a Comparative Analysis of Ten Countries*. 1. publ. Cambridge: Polity Press.

Sweezy, Paul M. 1951. "The American Ruling Class. Part I." *Monthly Review* 3(1):10–17.

Sweezy, Paul M. 1956. "POWER ELITE OR RULING CLASS?" *Monthly Review* 8(5):19–31.

UBS and PwC. 2019. ""The Billionaire Effect.""

Useem, Michael. 1978. "The Inner Group of the American Capitalist Class." *Social Problems* 25(3):225–40.

Useem, Michael. 1984. *The Inner Circle: Large Corporations and the Rise of Business Political Activity in the U.S. and U.K.* Oxford; New York: Oxford University Press.

Useem, Michael, and Constance Gager. 1996. "Employee Shareholders Or Institutional Investors? When Corporate Managers Replace Their Stockholders*." *Journal of Management Studies* 33(5):613–32. doi: 10.1111/j.1467-6486.1996.tb00811.x.

Wealth-X. 2022. "World Ultra Wealth Report 2022." Retrieved March 21, 2023 (https://www.wealthx.com/report/world-ultra-wealth-report-2022/).

Weber, Max. 1922. *Wirtschaft Und Gesellschaft*. Tübingen: J.C.B. Mohr.

WID.WORLD. 2018. "World Wealth and Income Database."

Winters, Jeffrey A., and Benjamin I. Page. 2009. "Oligarchy in the United States?" *Perspectives on Politics* 7(4):731–51. doi: DOI: 10.1017/S1537592709991770.

Wright, Erik Olin. 1976. "Class Boundaries in Advanced Capitalist Societies." *New Left Review* 98(1):3–41.

Wright, Erik Olin. 1998. *Classes*. London and New York: Verso.

Wright, Erik Olin. 2015. *Understanding Class*. London: Verso.

Zeitlin, Maurice. 1989. *The Large Corporation and Contemporary Classes*. New Brunswick, New Jersey: Rutgers University Press.

1st Article:

When Should We Believe Research Using ORBIS Firm Data?[1]

Abstract

Since around 2010, large off-the-shelf firm datasets covering a wide variety of variables for millions of firms are for sale by commercial providers. Such datasets are the basis for high impact scientific works, and have also beyond that been used frequently e.g. in economics, finance, political economy, economic sociology, geography, and other disciplines. One example of such an offered dataset is the ORBIS firm data, claiming to cover more than 400m firms globally in 2022. After examining and using the dataset in a four-year research project, I make the case that there are significant problems to use the data for research, depending on the question to be answered, as well as the available capacity and skills to make the data usable. Some of these are already known from the literature, others are not. But the most pressing point is that the magnitude of these problems is still understated. I gather (references to) solutions and applications to overcome a few examples of the identified problems. However, overall, I argue that the complexity of the data, existing errors, and inconsistencies, missing relevant information, the interrelation of these problems, and insufficient documentation and support provided with the data only allow for confidence in the validity of results under very specific conditions. These conditions will in most cases only apply to a much smaller sample than the claimed hundreds of millions of firms covered which undermines the main reason to work with the data in the first place. Evaluating whether conditions for valid inferences are met requires extensive knowledge of the data. ORBIS in its current state should only be used for research if sufficient time, resources, and skills are available to fully understand the data, and overcome its problems – if possible - convincingly and demonstratable. Furthermore, these issues have serious implications for the feasibility to review and evaluate research using the ORBIS data, and therefore reproducibility, unless significant additional effort is made by researchers to document the data preparation process and its implied decisions. This article gives some insights to reviewers as well as potential authors to better evaluate the use and preparation of ORBIS data. This added required effort for researchers, together with the price of getting access to the data in the first place, further speaks against its scientific use.

Submission status:

To be submitted as MPIfG discussion paper after another round of comments by colleagues.

[1] Many thanks to Benjamin Braun and Sebastian Kohl for very helpful comments and critiques of earlier versions of this manuscript.

1. Introduction

The use of large off-the-shelf datasets for research contains promises and perils (Heemskerk et al. 2018; Liu 2020). This article focusses on the perils of using the ORBIS data for research (Bureau van Dijk 2020). It is motivated by the experiences and observations I made while working with the data in a four-year Ph.D. project. It aims to raise awareness and sensitivity for the complexity and related problems of ORBIS as a frequently used big dataset for firm research. This is important because the scientific method relies on trustworthy data and critical reflection of its sources. The larger the datasets used by scientists, the more difficult it becomes to oversee the data quality and validity of the complete dataset. Depending on the size of the sample used, ORBIS is a good example of this. One purpose of this article is to share knowledge about ORBIS as a data source, to enable researchers to choose whether they should use ORBIS to answer their research question at hand. Researchers expecting yet another external (in)validation of the ORBIS data will be disappointed - but such comparisons of ORBIS' coverage with that of other sources can be found elsewhere (Bajgar et al. 2020; Garcia-Bernardo and Takes 2018; Kalemli-Ozcan et al. 2022a, 2022b). I will focus here on internal issues and observations of the ORBIS data, and my experience with it.

There are several (large) financial datasets to choose from when sketching out a new research project and its potential data sources. Liu (2020) offers a critical overview and collects common problems mentioned in the literature with some widely used financial datasets. The most extensive among the options is the ORBIS global firm database. Its claimed size of covering several hundred million firms positions it in what is sometimes called "big data". Google scholar lists 11,700 research items at least mentioning the ORBIS data, 66 of which have more than 100 citations with the top article being cited 1013 times.[1] ORBIS invites itself mostly when interested in a transnational rather than a comparative perspective, when interested in the largest firms of the globally largest economies, when interested in non-listed firms as well, and when interested in specific combinations of variables such as financial, ownership, and management data.[2] Because of its unique characteristics, the data has been used and discussed widely by research communities broadly interested in the economy. Heemskerk et al. (2018) make an important start to a conversation which I think is very necessary among the research community. Namely, the authors provide a very helpful hands-on best practice procedure for big datasets of how to tackle problems of "the five Vs": volume, velocity, variety, veracity, and variability in large firm datasets in general. They structure their overview along four questions: "Are you clear about the appropriate unit of

[1] The search was conducted on 16 March 2023. Search terms were "ORBIS AND (BvD OR Dijk)".

[2] In many other cases alternatives may be much more suitable and clean such as Compustat for globally listed firms, Refinitiv for ownership and management data in some cases, or data sources only for individual countries such as the leak of the German Handelsregister.

analysis? Is there entity ambiguity in your data? How complete are the data? How accurate are the data?" (Heemskerk et al. 2018:8)

This article draws from this work but focuses more on the pessimistic side in conversation with it. More concretely, I challenge the implicit assumption that I think could be taken from such methodological articles applied to ORBIS, that it *is* possible to make valid inferences when just following the right data preparation steps (e.g. Heemskerk et al. 2018; Kalemli-Ozcan et al. 2022a, 2022b). In other words, the question at stake is whether the respective data quality is good enough and when this may not be the case.[3] In this spirit, I take up Heemskerk et al.'s (2018:26–28) initiation to "discussions among the research community when it comes to data quality and the means of addressing it" and "a vivid, candid and critical academic debate on the merits and pitfalls of using big corporate network data". The position I would like to take in this debate is that it so far is too concerned with making data work and thereby neglects the possibility that a data source may be so problematic that, on average over all applications, it could lead to wrong, i.e., not reproducible, or not valid, scientific knowledge in too many cases. I do understand the urge in Bajgar et al. (2020:10) "not to give a definite black and white statement as to when Orbis data should be used" or the concern in Heemskerk et al. (2018:27): "we are not claiming that work that does not give an indisputable answer to the proposed set of questions should remain unpublished.". However, I think that while the problems of the dataset are spelled out on hundreds of pages by researchers already (e.g. Bajgar et al. 2020; Garcia-Bernardo and Takes 2018; Kalemli-Ozcan et al. 2022a; Kalemli-Ozcan, Fan, and Penciakova n.d.), the dangers and inaccuracies in the data are in my view still understated. I do furthermore find that our experiences with ORBIS, e.g., the way I accessed the data and its consequences, is so far not presented correctly in existing literature, and some problems (e.g., inaccuracies in the ownership data) have to my knowledge not been presented or spelled out yet.

I make the case here that under many circumstances using ORBIS is not only a question of applying the right preparation steps, diagnostics, and fixes but the data should simply not be used for research. Ironically, this is the more the case the larger the sample in terms of number of firms, but also number of countries, which seemed the main advantage of using the ORBIS data in the first place. What is more, it is relatively hard to come to this conclusion, especially at the beginning of a research project when unfamiliar with the data. This article is therefore supposed to help researchers to decide against using the ORBIS data at an early stage of the research process if necessary. On the one hand, this could simply save a lot of time and not the least money for some

[3] This implies that they may invite researchers to overestimate themselves by thinking along the lines of "If paper X says it is possible to make the data work/representative/good enough it must be possible to do it". From my own experience, I can tell that significant data problems remain and must be tolerated without fully being able to oversee their magnitude. This aspect is not stressed enough in the existing literature.

research organizations. On the other hand, and in the worst case, investing a lot of time, effort, and money into preparing datasets as complex as ORBIS could be an incentive to publish results despite not being sure enough about their validity, i.e., bad scientific practice. This is the more dangerous in the case of ORBIS because obstacles are also exceptionally high for reviewers to evaluate data preparation and consequently results appropriately. The longer I worked with ORBIS, the surer I was that hardly any reviewer has the experience with the dataset to truly evaluate whether the sample I use, the data preparation steps, and the analyses I applied led to valid results or not. This realization can be dangerous for good scientific practice, even unconsciously.

I spell out the problems responsible for this alongside the structure suggested by Heemskerk et al. (2018). However, upfront I would like to pay attention to one important factor namely that of necessary IT skills to perform many of the necessary preparation steps and diagnostics. This is a point where I disagree with Heemskerk et al. (2018:7), who consider "it as a misperception that the integration between big data and social science is about technical capacities. Certainly, within the context of BCND, the volume is larger than before but manageable with current tools and techniques". I do have a degree in business and IT and consider myself interested in, and to some degree capable of, coding in Python in general, and data preparation specifically. Yet, I was not successful in applying several of the so far available suggested or necessary data preparation steps, at least not with reasonable time and hardware resources, to fully account for some problems in ORBIS. It is therefore important to realize that working with the ORBIS data and making it usable requires deeper knowledge of preparation of large datasets in some programming language, as well as efficient coding in terms of working memory and computation time. It may of course also be possible to hire data engineers or other people with training in this area, but this will likely not be affordable to most research projects especially given the current huge demand and low offer of these skills in the general labor market. [4] I do however speak more about the case where trained social scientists, even techy economists, attempt to make use of the data. It will in many cases not be feasible to use it as a single researcher, but a larger team of IT skilled and determined researchers or data scientists working closely together might be more promising.

Since I do not speak about big datasets in general but about ORBIS specifically, it is necessary to position this article a bit more in the ORBIS-specific existing methodological literature. Table 1 gives a succinct overview of existing literature dealing with ORBIS and how to

[4] In the business world, the tasks required to understand and clean messy big datasets are fulfilled by so called data engineers which comprises a whole fast growing and evolving field in computer science. However, this is a skill set that is to my knowledge not taught in any social science program and which requires months and years of training or experience in complex data engineering tasks. Tools commonly used in businesses to solve data problems of this magnitude are not familiar to social scientists at all.

position this contribution within it. In terms of access to the data, one relevant restriction of this article needs to be disclosed. There are several ways to access the ORBIS data which vary by the effort one needs to put into receiving the data, and not least by price. One main difference between these access options is the question of representative longitudinal vs. cross-sectional data. This is also one relevant axis along which to structure existing literature on the use of ORBIS for research, since the most extensive works are interested specifically in historical data, and many of the preparation steps only relate to the question of harmonization of different cross-sectional snapshots from ORBIS (Bajgar et al. 2020; Kalemli-Ozcan et al. 2022a). In addition to that, it should be noted that these works are mainly interested in aggregate outcomes of firm-level microdata such as market concentration, only focus on selected sectors, and mostly on European countries. Whereas my research interests in terms of time points were rather, first, to classify individual firms as either controlled by a super-rich family or not at a single point in time, to use this indicator as an independent variable to predict an outcome matched from an external data source (firm party contributions and lobbying) in a specific year. Second, to map firm networks based on the ORBIS management data in one year. Most contributions dealing with ORBIS have a specific research interest in mind and in all cases not all discovered and solved problems are relevant for all potential research projects one can pursue with ORBIS data.

We[5] accessed the ORBIS data in the following way. My institution had a paid contract with the data vendor Bureau van Dijk which allowed me to manually (only assisted by self-implemented web scraping) download the data through the ORBIS portal. This is the more affordable access option which also has some downsides in terms of data quality especially for longitudinal data, as will be spelled out in more detail in the next section.[6] It may be the case that some of the problems described here are only applicable to this download method (manual download through web scraping of the export function in the ORBIS portal). Researchers interested in using the ORBIS data should note that there are also other hands-on manuals for the ORBIS historical disks and the new "ORBIS historical product", which only relate to another method of access which is a lot more expensive to procure (Kalemli-Ozcan et al. 2022a, 2022b, n.d.). However, depending on the research questions of interest this access option is sufficient, especially when focusing on a cross-section of the most current ownership and management data – or 1-2 years before the download date for financial data to account for the reporting lag. Differences between the access options are further discussed below in section 3, and in Kalemli-Oczan et al. (2022b:2–3).

[5] More employees from my institution were involved in the communication with BvD and the selection of variables that could be relevant to us but I was responsible for the download and the only researcher really working with the downloaded sample so most of the experiences presented here are from my individual perspective.

[6] However, from our impression the products and pricing offered by BvD are not set to stone or even published somewhere but seem to vary case by case.

	Bajgar et al. (2020)	Garcia-Bernardo and Takes (2018)	Heemskerk et al. (2018)	Kalemli-Oczan et al. (2022)	This work (2023)
Focus on ORBIS as the main data source?	Yes	Yes	No	Yes	Yes
Focus on longitudinal use of the data?	Yes	No	No	Yes	No
Access method	2017 historical disk (Bajgar et al. 2020:8)	"creating a snapshot of the Orbis database in November 2015" (Garcia-Bernardo and Takes 2018:166)	Not specified/ only using ORBIS as an example	Multiple historical disks from ORBIS and Amadeus (Kalemli-Ozcan et al. 2022b:9) and "historical product" (Kalemli-Ozcan et al. n.d.)	Two downloads (2020 and 2022) from the ORBIS web portal
Research interest	"applications of Orbis to the study of productivity and business dynamism." (Bajgar et al. 2020:9)	"corporate networks, in which links represent particular relationships between corporations" (Garcia-Bernardo and Takes 2018:165)	"big corporate networks" (Heemskerk et al. 2018:6) (e.g. interlocking directorates but also others)	Mainly "macroeconomic outcomes" (Kalemli-Ozcan et al. 2022a:1), SMEs (Kalemli-Ozcan et al. 2022a:9–12), and "industry concentration" (Kalemli-Ozcan et al. 2022a:12–19)	Identifying super-rich individuals and families among shareholders and managers; describing and analyzing their firms and firm behavior

Table 1: Literature on ORBIS data quality and relevant differences.

To evaluate the claims made in this article, it may be useful to know more about how and with what purpose I approached the ORBIS data. I consider myself a computational social scientist who used the ORBIS data for three purposes: first, identifying super-rich (i.e., listed on public rich lists) individuals and their firms among the ORBIS shareholder data. In this step, I matched an external list of strings containing individual, family, and firm names to the ORBIS shareholder and manager names. Through this, I gathered in-depth knowledge of the ownership data and its problems. Second, I used other variables such as listing information, financial data, and management data as estimators in regression analysis. This way I gained some familiarity with missingness, errors and inconsistencies, entity ambiguity, and other problems with the ORBIS data in general. Finally, third, I worked with management data to analyze networks of interlocking directorates. This way I gained insights into problems with the management data, and especially its validity over time or duplicates. This article is written based on these experiences and perspectives.

One takeaway from working with ORBIS for four years is that the longer one works with the data and gets familiar with it, the less one trusts in it. Even after four years I constantly find new unexpected, incomplete, very much unintuitive, or simply erroneous data in our downloads. This led me to ask the question that drives this article and is motivated by concerns of good scientific practice: *Under what circumstances can we trust results based on the ORBIS firm data?* As a final disclaimer, I would like to stress that I published research results based on ORBIS myself. I did this because I am as confident as one could be in my results by examining them and testing them for robustness myself to a feasible extent. At the same time, I do strongly believe in peer-review and know there are probably only a few other researchers in academia with enough knowledge about ORBIS who could review my work, as well as any work using ORBIS data, to the extent it would be adequate. In the current academic system where the time available and acknowledgment for in-depth reviewing (not to speak of code review) are very restricted, the complexity of the data is a problem for the reviewing process, and also for my own research and published results of course as much as for any other research project working with the data. That is to say, I do not want to give the false impression of authority to have fully understood the fallacies of ORBIS. In the following, I begin by briefly presenting basic information about the ORBIS database in the next section. Next, I present a (likely incomplete) list of problems and shortcomings of the data, together with more and less successful solutions to some of the problems offered in the literature, or that I have tried myself. Finally, I offer some summarizing thoughts and additional reflections on the question of when we should believe research using the ORBIS firm data.

2. Background on the full ORBIS data and our used samples

The ORBIS database offered by Bureau van Dijk (BvD), founded in 1970 and since 2017 a subsidiary of Moody's, is the largest global firm database comprising data on firm financials such as revenue, EBITDA, etc. In addition, it includes ownership data including shares and shareholder information, and management data such as names of management. There is also other data included in the database that is even more sensitive, such as credit rating scores of individual companies, addresses of owners, etc.[7] The data was brought to the market by BvD already in 1987 as a CD-ROM (Group van Dijk 2013).

In January 2023, BvD claims ORBIS to cover data on almost 447m firms globally. Table 2 shows the distribution of coverage by region and available financial data. 427,399,056 of these companies are active in January 2023, 84,501 of them are listed companies, and only 4,928,754 companies are not subsidiaries, i.e., not owned by another company with 50% or more. A related issue is that companies may publish consolidated accounts for a whole corporate group, unconsolidated accounts only for the respective entity in the firm hierarchy, or both types of accounts. This fact and related problems are very relevant when interested in financial accounts and their aggregation to the market or country level and they are laid out together with potential solutions elsewhere (Bajgar et al. 2020:50–52; Kalemli-Ozcan et al. 2022a:78–81). However, related issues are also discussed below in the unit of analysis section. Bajgar et al. (2020) also present and discuss coverage by different account types. Kalemli-Oczan et al. (2022a) gather the information providers and legal regulations for European Countries in detail.

[7] According to the talks we had with representatives of the company, the only hard restriction to publish excerpts from the database, e.g., for review purposes, were the credit rating scores. We were allowed to publish everything else for scientific purposes in principle.

World regions and countries	Companies with detailed financials	Companies with limited financials	Companies with no recent financials	Companies without financials	Total
Total	25,657,896	119,041,365	73,526,348	228,674,634	446,900,243
North America	33,288	35,497,401	6,130,192	29,877,514	71,538,395
Western Europe	12,231,563	10,833,281	18,710,273	45,630,171	87,405,288
Eastern Europe	7,251,973	2,920,536	12,473,779	22,277,876	44,924,164
Middle East	3,220	1,681,911	1,047,266	4,689,310	7,421,707
Far East and Central Asia	3,538,960	21,193,207	28,092,054	64,443,982	117,268,203
South and Central America	2,352,097	44,122,988	2,274,382	13,630,177	62,379,644
Africa	219,200	2,782,364	521,115	17,044,118	20,566,797
Oceania	27,444	9,466	4,274,641	29,064,748	33,376,299
Supranational	48	-	-	16	64
No country specified	103	211	2,646	2,016,722	2,019,682

Table 2: Firms and financial data covered by region taken from the ORBIS portal as of 16 January 2023.

The thoughts and insights presented in this paper stem from the analysis of two samples (cf. Figure 1). The first was downloaded from the ORBIS portal by web scraping in 2020 and comprised the roughly 700,000 largest global firms in terms of operating revenue and the number of employees, as well as all their subsidiaries, and all their shareholders. In total, the sample comprised 3.2m firms. This sample was then extended by several additional downloads in 2022 tailored to our research interests, the largest of which included the 3m largest firms designated by ORBIS and their shareholders which summed up to a total of 4m firms. I also downloaded other samples such as all German firms which were not yet included in the sample. More precisely, I implemented a web scraper that clicked through the ORBIS portal in an automated manner with Selenium (2023) in Python. The script automatically logged into the ORBIS web portal, loaded a pre-defined search and a pre-defined selection of variables, and exported the results through the export function in chunks of around 80% of the maximum amount possible. This maximum amount depended on the number of variables chosen and was read in automatically by the scraper for every combination of variables I downloaded. Overall, I downloaded more than one hundred variables from the financials, ownership, management, industry, and stock sections for several million firms.

a) The first sample, which was downloaded in the first half of 2020 (own summary)

Step	Selection	Years	Number of firms
	Sample 1		
1.	The largest firms in terms of operating revenue with a minimum of US$ 50m	Any of the years 2010-2018	385,028
2.	The largest firms in terms of number of employees with a minimum of 200	Any of the years 2010-2018	425,605
3.	All shareholders of all firms from selection steps 1. and 2. included in the database	Any of the years 2010-2018	796,351
4.	All subsidiaries of any of the firms from selection steps 1. and 2. included in the database	Any of the years 2010-2018	2,406,289
		Sample 1 total excl. overlaps:	**3,207,095**

b) Summary of the second sample downloaded in the first half of 2022 (screenshot from the ORBIS portal search)

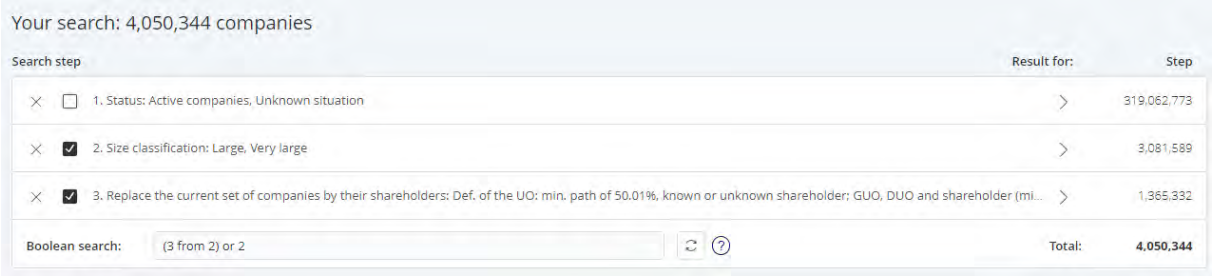


Figure 1: Summaries of samples downloaded from the ORBIS web portal a) in 2020 and b) in 2022.

3. Data access, support, and documentation

In this section, I will briefly report on the interaction our research institution had with the ORBIS data provider BvD, the different ways of data access, experiences with the documentation of the ORBIS data, and the support offered by BvD post-sale. We contacted BvD for the first time in 2019, to procure the first sample for my Ph.D. project. Two ways of access were offered to us. First, yearly snapshots of the full database since its creation (historical "disks"). Second, we got access to the portal and unlimited downloading and could download from the portal whatever we needed during an agreed period of a few months. According to BvD representatives, legal changes shortly before our request led to a steep increase in the pricing of the historical disks to a low six-figure amount. The second option could be offered much more affordable for an amount in the lower five-figure range. We chose the second option since the first was not affordable to us, as will assumably be the case for most except the best-funded research institutions in the world. We were made aware that full historical data would only be available on the historical disks but other than that there would be no limit. We could download as much as we managed to within the agreed period of around six months, and from the technical side only what is downloadable with the export function in the ORBIS web portal.

The way I accessed the data from the portal is claimed to lead to lower quality data, especially by Kalamli-Ozcan et al. (2022a, e.g. the footnote on p. 7). The main concern of the authors is the longitudinal dimension of the data. While it is true that the data from the web portal has several downsides towards the historical disks, most notably the shorter period of historical data included, I cannot confirm two of the most serious downsides of this access method presented by the authors. First, I could download historical data from the portal, e.g. operating revenue of the past 10 years, without systematically missing values ("download cap": e.g. Kalemli-Ozcan et al. 2022b:4). I am confident that there was no such cap in place for two reasons. First, no such download cap was mentioned by BvD sales representatives as our contract allowed us to download full information on millions of firms directly from the portal.[8] Second, I checked several cases by hand and each time the missing data was also missing in the web portal. Unless there is a distinct dataset distributed by BvD that has more complete information than the web portal, I think that our sample should show the same degree of missingness as the historical disks.[9] I encountered one case where I wanted to download historical ownership data and values existing in the web portal were capped during the process of our second download in 2022. However, after consultation with

[8] It is possible of course that our contacts were not aware of such a cap but in that case, I am pretty sure our organization would have legal claims against BvD because it would mean a breach of our contract.

[9] But this needs to be verified by comparing the two versions. I provide insights into our missingness structure that can be used for comparison purposes in section 4.

our BvD representative and his internal conversations with the data team, the download then worked as expected for all cases, and it was reconfirmed to us that there should not be a cap.[10] A second point I cannot confirm which is raised in Kalamli-Ozcan et al. (2022a:6) is that downloading historical ownership information was not possible for a large number of firms from the web portal. I downloaded historical ownership data successfully even for longer periods than 10 years by simply choosing direct and total shares at the respective past date as columns in the view search results view, and then exporting the output. However, indeed, the extent of the survivorship bias when retrieving historical ownership information in this way remains unclear. While we have no external data source to check attrition systematically, our download includes more than 100,000 shareholders that only held a share in some company in 2010 and no share after. Shareholders, therefore, do not seem to be excluded systematically even if there is no current share held by them in the data for the past 10 years. Disregarding these two points where I have to object to Kalamli-Ozcan et al. (2022a), it seems plausible that, as the authors report, firms are not included e.g. if they were inactive for five years or more. I can also confirm there is a lag in reporting of about 1 to 2 years on average which depends on attributes such as country, type of firm, size of the firm, etc.

The problem of available historical data in the web portal seems to vary significantly by the kind of variables used and downloaded. The non-availability of historical information in the portal is a lot more pressing for the management data. Figure 2 shows the number of management positions by year (how many managers were in office in the respective year) when using historical management data (dates of appointment and end dates) from the web portal. This data is the basis of the vast research on interlocking directorates. I already applied several steps of imputation in case of missing begin or end dates of the management appointment. A historical analysis of management positions seems impossible based on data from this download, most likely because historical management positions are not included in the portal anymore – or only in comparably few cases. The development mostly reflects which of the current directors were already in office in the respective years. The selection criteria for which historical positions are kept, and which ones are dropped are unknown to me and are also not included in the documentation. It also remains unclear whether the data is more complete in the historical vintages, and it would be worthwhile for researchers with access to the historical vintages of the management data could test and report this.

[10] This anecdote does however suggest that a download cap might depend on the exact agreement with BvD.

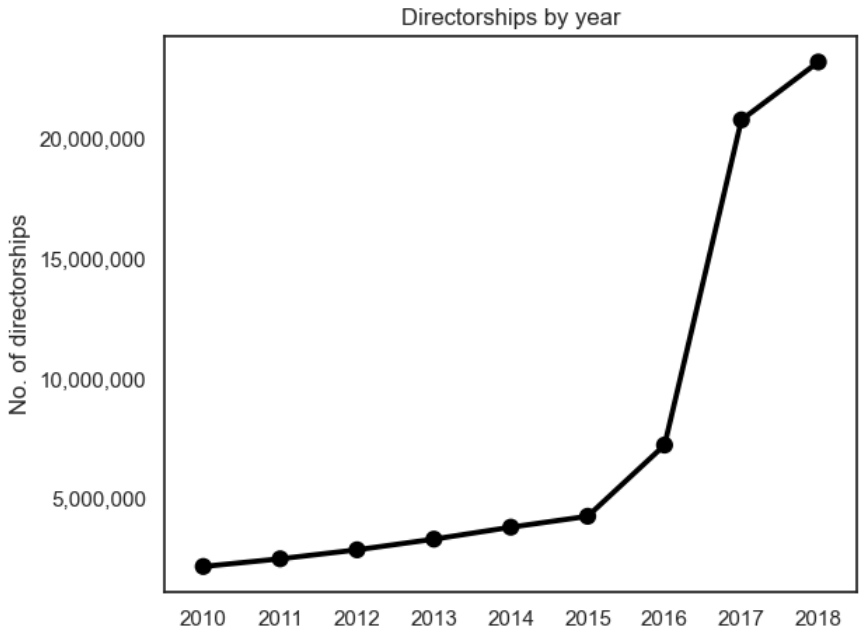


Figure 2: How many managers are in office according to the sample? The number of management positions by year in the ORBIS data from our first download in 2020. The curve suggests either attrition of management positions or a strongly increasing sample of firms for which data is available.

Changing BvD numbers over time is another shortcoming raised by Kalamli-Ozcan et al. (2022a). This may for example be the case if companies changed addresses or for other legal reasons depending on the country (Kalemli-Ozcan et al. 2022b:5). I would like to add to this that the problem needs to be segregated into two distinct problems: changing BvD IDs of companies and changing BvD IDs of individuals as shareholders. As concerns the former, we were made aware by a colleague from another institution that this can be overcome by an integrated tool in the ORBIS portal through which a list of company BvD numbers can be uploaded, and which then provides a list of old and new BvD IDs as a result which can be downloaded.[11] However, in our research when combining two downloads of ownership information, the most pressing problem was that of changing BvD ID numbers of shareholders which, in combination with our extrapolation of ownership shares, leads to a legacy of duplicate shareholders with old and new BvD IDs. Often, these are individuals or families as shareholders. One cannot simply search and download changed BvD IDs for individuals in the ORBIS portal, and therefore the only solution for this problem we see is to receive a correspondence table of changed BvD IDs of individual shareholders from a BvD representative. But this piece of information and how we retrieved it brings us to our experience with the ORBIS documentation and support. For a lot of questions, it is not possible to find information in the documentation.

[11] To find the tool just choose to search for companies by BvD ID numbers in the portal and upload e.g. a CSV file with BvD numbers.

A fundamental problem with the BvD support that I faced during our download phase and figuring out what exactly I needed was that the sales team does not seem to be closely connected to the data team. In summary, I was not successful in receiving support or additional information from the BvD Orbis support – or at least not the correct information. I was especially persistent in asking about two things. One is the question of whether it is possible to download old BvD IDs or to receive a translation table from BvD with old and new BvD IDs. Kalamli-Ozcan et al. (2022b:5) report that it should be possible to retrieve such a table form a BvD representative. The second request I had was to ask whether there was any way of acquiring a table that assigns Shareholder BvD ID numbers to unique contact identifiers (UCIs) of managers. Imagine a use case of looking for an individual majority shareholder's management board membership. Say, which corporate boards have super-rich individual shareholders such as Jeff Bezos as members. I asked our sales representatives, who claimed to have asked the data team, numerous times how to acquire such a table. I was sure that some sort of database or translation table must exist because when clicking through shareholders manually in the portal, the respective UCIs are shown together with an individual's shareholder BvD ID. In both cases, I was not successful in getting help, and most likely not even successful to make the other side understand what exactly I need. In both cases, I also did get the wrong information as a response. In the first case of changing BvD IDs, the correct or helpful answer would have been to point us to the tool integrated into the ORBIS portal. Instead, I only got the information that the old BvD ID cannot be seen in the portal. In the second case, I received the answer after weeks of continuous inquiry that it is possible to show the BvD ID number of shareholders of a company together with the UCIs of its managers. This did of course not answer our question and is not helpful. I gave up after this answer to my seventh email.

The point of these anecdotes is that I did not receive any helpful information from our contact with BvD support post-sale. I heard about other research teams working more closely together with different departments of BvD and making better experiences. Researchers thinking about working with the data should at least be aware of the possibility that no further support is offered with the sold data. Considering the numerous data problems, inconsistencies, and unknown sources and definitions of the data, this is a big problem. Researchers have to rely on the few manuals and guidelines published by other researchers cited here, and on asking each other for help and clarification. I can only speculate from our experiences, and what I heard from other experiences, that likely our financial weight by procuring the cheaper access option was not large enough to secure support with the data. The purpose of the web portal' documentation from my impression seems to be to enable customers in the business to use the basic functions of the portal. A use case might be to do some sort of small-scale analysis of businesses or competitors in their industry or area of interest. It is mostly about how to use the different functions of the ORBIS

portal. Only a small share of the documentation (certainly not sufficient for research purposes) is about the data, its sources, and its definitions, which is of course relevant and necessary information when using the data for public research. Questions such as the one about changing BvD IDs or how exactly variables are defined can be found in a few cases but not systematically. For this reason, most of the background information about the variables, how they are defined and generated, and any additional information that is not included in the data needs to be requested from the BvD ORBIS support. The next section presents the problems of the data that I discovered while working with the data.

4. Data problems and some questions to ask

As mentioned before, criteria of data quality of large firm datasets as ORBIS, e.g. representativity (Bajgar et al. 2020; Kalemli-Ozcan et al. 2022a) and the "five Vs" (Heemskerk et al. 2018), were already suggested in the literature. Figure 3 shows the best practice procedure developed by Heemskerk et al. (2018). Following this procedure, I discuss known and new problems of application in ORBIS for the four mentioned problem classes: unit of analysis, entity ambiguity, completeness, and accuracy.

Unit of analysis

The scientific applications to which this article speaks use ORBIS because they are interested in firms as the unit of analysis. However, ORBIS is the best case I know to illustrate that the "important ontological question: what is a firm?" (Heemskerk et al. 2018:11) is a very complex one. On the one hand, there are corporate groups including branches in different countries and possibly multiple other subsidiaries. These subsidiaries can either be understood as active parts of a company that differentiates into different activities possibly in different countries to exploit beneficial regulations and international competition for foreign direct investment (Reurink and Garcia-Bernardo 2020). They could also be solely administrative shell companies that only exist in a specific jurisdiction for tax purposes or other reasons, but do not produce anything (Garcia-Bernardo et al. 2017; Heemskerk and Takes 2016b:98). Which of these companies is *the* company?

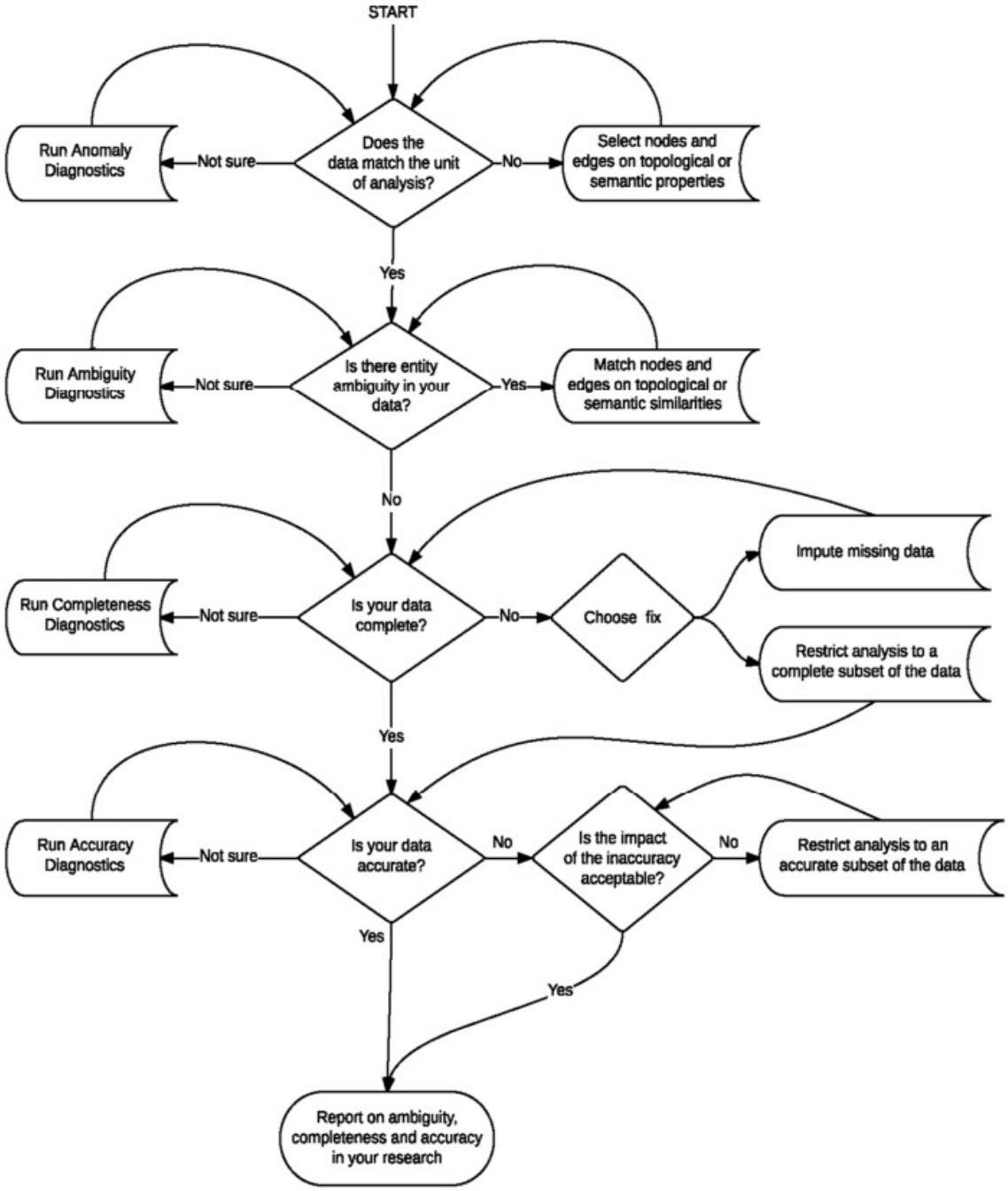


Figure 3: How should big corporate network data be prepared for research? Decision tree from Heemskerk et al. (2018:10).

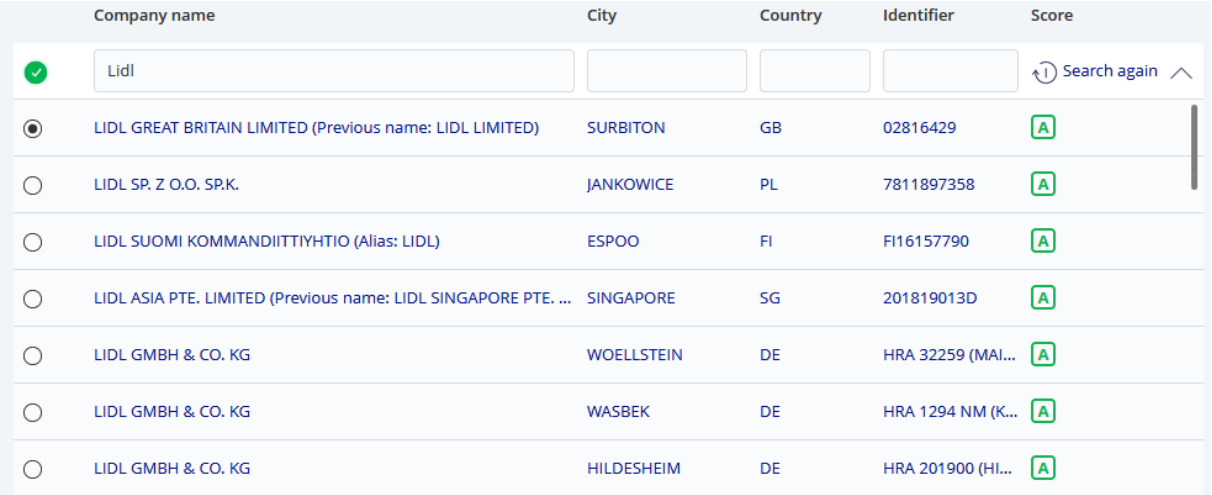
The larger the total assets of a firm, the more subsidiaries it has on average (Heemskerk et al. 2018:17). It might be intuitive trying to solve this problem by keeping only the head of a corporate group. Taking this fix may– depending on the research question - exclude firms of interest or that should be included to produce generalizable results. For example, Heemskerk and Takes (2016a) discuss that while only keeping the head of a corporate group will reduce administrative ties and shell companies which may not be relevant for several questions, they also

exclude more relevant parts of the holding structures. The authors give a few examples. One is that excluding all subsidiaries in an ORBIS sample would exclude Volkswagen AG since it is a direct subsidiary of Porsche Holding. Depending on the research question this makes a relevant difference. In my research analyzing control over firms, I argue that excluding subsidiaries and with-it Volkswagen is the better choice since members of the management of Porsche Holding should have control over Volkswagen as a subsidiary. I, therefore, assume the power to rest on the Porsche Holding board. The benefit is that tight clusters of firms belonging to the same corporate group are dissolved. The consequence of this choice is that important ties between boards that go out from a subsidiary board are excluded. When interested in all ties of an elite community of managers in contrast, it is justifiable to include all subsidiaries and their management boards, because these boards individually may have important ties to other companies – and not only the board of the top holding company. In this case, an unknown error driven by administrative ties is accepted in return.

One of the things one needs to be aware of in this regard when using ORBIS is that there is no easy way in between these two extremes of excluding all subsidiaries or keeping them. While there is a variable provided in ORBIS for some corporate groups containing the "level" within a corporate hierarchy, this is not necessarily meaningful information to solve this problem and even if it was it is only available for a small share of corporate groups. Taking an example from my research on super-rich capitalist families, the question of what is *the* firm that makes up most of the fortune of a family is not trivial either. Asking someone on the street what company made Dieter Schwarz the richest German, might lead to the quick answer "No doubt, it's Lidl." But the supermarket chain Lidl is owned by a foundation in Germany with a revenue of more than 79bn € and 120,000 employees. This foundation is owned by five other firms and foundations, one of which is Lidl SB KG which is again held by D. Schwarz Beteiligungs KG which is again held by six firms and foundations, and this continues for a couple of levels until ending up at a level where the ultimately owning individual is located among some other companies that belong to him. The point of this is that corporate hierarchies are complex, and this full complexity is mercilessly reflected in ORBIS there is no easy way to identify a firm if you do not have an exact idea of what you mean by that. To make such an idea more concrete, one example could be "the entity with the largest revenue in their unconsolidated accounts within a corporate hierarchy". But even if you have an exact definition such as this one, the question of how to operationalize it considering other data problems such as mixed-up total and direct shares within corporate groups and incorrect GUO information is quite another question.

There may be cases where researchers have an external list of firm names that they want to find in the ORBIS database. At first sight, the good news is that there is a batch firm search

implemented in the ORBIS portal that is supposed to make it easy to match such lists. The bad news is that depending on what exactly you are looking for, the result might not be the entity in ORBIS you are interested in but some other subsidiary of the corporate group.[12] Sticking with the example above, there are many companies containing the name Lidl in ORBIS. Using the batch search to find the firm gives as a first suggestion Lidl Great Britain Ltd. (Figure 4). Adding Germany as a country gives another branch, the Lidl GmbH & Co. KG which is not the head of the corporate group. Depending on your research question, this might be a problem. In some cases, it may help to simply include the whole corporate group, but in some, this will not help, not the least because the corporate group data is inaccurate as well, as will be discussed below in the accuracy section. This example also illustrates a general ironic point for the first time that is one of my main points in this text: while the database promises to enable large-scale analyses of millions of firms, often the devil is in the detail of individual cases. The definition of the correct unit of analysis might only be possible for every individual corporate group by hand – or by accepting a certain unknown error that is impossible to oversee at a large scale.



	Company name	City	Country	Identifier	Score
✓	Lidl				Search again
◉	LIDL GREAT BRITAIN LIMITED (Previous name: LIDL LIMITED)	SURBITON	GB	02816429	A
○	LIDL SP. Z O.O. SP.K.	JANKOWICE	PL	7811897358	A
○	LIDL SUOMI KOMMANDIITTIYHTIO (Alias: LIDL)	ESPOO	FI	FI16157790	A
○	LIDL ASIA PTE. LIMITED (Previous name: LIDL SINGAPORE PTE. ...	SINGAPORE	SG	201819013D	A
○	LIDL GMBH & CO. KG	WOELLSTEIN	DE	HRA 32259 (MAI...	A
○	LIDL GMBH & CO. KG	WASBEK	DE	HRA 1294 NM (K...	A
○	LIDL GMBH & CO. KG	HILDESHEIM	DE	HRA 201900 (HI...	A

Figure 4: Suggestions from the ORBIS batch search implemented in the web portal when searching for the company name "Lidl" (screenshot).

A final problem I would like to point out concerning the unit of analysis is the problem of financial account types and consolidation of financial data in ORBIS. This is something that works interested in macroeconomic aggregates are very concerned about (e.g. Bajgar et al. 2020; Kalemli-Ozcan et al. 2022b). Generally, when interested in financial accounts it makes a relevant difference

[12] Or a completely different firm since this is only based on string matching and it therefore depends on how common a name is.

which accounts types to include. Recalling Table 2, only half of the firms contained in ORBIS have recent financial data at all, roughly one-fourth have limited financials, and about 25m (5%) have detailed financial data. Of the available account data 2.5m are consolidated, i.e., including subsidiary accounts, and 42.7m are unconsolidated, i.e., only counting the respective entity.[13] For example, simply adding up unconsolidated accounts does not lead to the value of the holding company. This does not have to be problematic for all research questions, e.g. when double counting subsidiaries and holding companies is not a concern. However, for example, when interested in combining financial accounts with ownership data, this *is* a problem. How do you value shares held by a shareholder? One way is to use firm financials as a proxy for firm value and multiply it by the share held. Researchers for example have used a similar method based on the ORBIS data to identify wealthy individual shareholders for survey research (Schröder et al. 2020). However, it becomes a very difficult task in cases such as the one by Dieter Schwarz illustrated above. There are two related sets of problems: One concerns the accuracy (see below) and the complexity of the ownership data. The other is more related to the unit of analysis and concerns which accounts to use as a proxy to estimate the value of the shares. The problem is that to my knowledge there is no reliable default way to determine which accounts to use if one does not want to count double accounts of subsidiaries.[14] A feasible option might be to consolidate all accounts of firms owned by firms owned by the shareholder of interest, but the problem just mentioned that consolidating all accounts of subsidiaries does not lead to the correct consolidated account of the head of the corporate group. This way you may, but not necessarily, get a correct ranking of wealthy shareholders, but the absolute numbers e.g. of share or revenue "owned" by an individual shareholder are not going to be correct. In individual cases, it might be possible to solve this by identifying the correct largest consolidated account, but in many cases, there are only unconsolidated accounts. Even more pressingly, this would require again to look at every individual case or at least the problematic ones which contradict the goal of a large-scale analysis of millions of firms or shareholders.

[13] If not specified otherwise the cited numbers from the ORBIS portal stem were valid in January 2023.
[14] Some suggestions in this regard are made by Kalemli-Ozcan et al. (e.g. 2022b:78–81).

For their longitudinal analysis of market concentration in Europe, Kalemli-Ozcan et al. (2022a:4–5) discuss that may lead to biased results by only using a certain type of accounting statements:

> "A priori there is no reason for focusing on a certain set of accounting statements as opposed to combining all statements, as long as one is careful about not double counting the same firm reporting both statements. Focusing on a selected set of statements will lead to focusing on a selected set of firms such as listed firms, business groups, foreign firms, and will give misleading trends in concentration. This practice of selecting certain groups will also deliver biased results due to changing regulation. For example, we show a sharp increase in concentration around 2007, which coincides with a change in the European accounting legislation."

Only taking certain account types to avoid double counting subsidiaries and heads of corporate groups therefore also does not seem to be a solution.

To sum up, the more researchers think about what they mean by a "firm", the less clear this may become. Excluding all subsidiaries of a corporate group may exclude important ties in the firm hierarchy, although the "firms" along the path will mainly exist on paper and their existence may have legal, tax, or other purposes. There are at least two opposing conclusions to draw from this. On the one hand, the complexity of the firm data in ORBIS could lead to the conclusion that it is not even meaningful to think in terms of "firms" or "corporate groups" when using the data but to treat the different entities as nodes in a network that contains a relevant extent of noise. The widespread use of the term firm does not resemble the empirical reality any longer since it is undercomplex. I would rather tend to the alternative conclusion which is that ORBIS is lacking a reliable indicator to identify ownership conglomerates that allow capturing what is widely understood as new forms of modern corporations (e.g. Reurink and Garcia-Bernardo 2020). The takeaways for researchers and reviewers concerning the unit of analysis, therefore, are the following questions:

1. **How do you/ does the reviewed work define the unit of analysis or answer the question "what is a firm"? How are subsidiaries treated?**
2. **Is it possible to operationalize this definition based on the ORBIS data?**
3. **What consequences does the definition have for the analysis, results, and possibly accepted errors, uncertainties, and shortcomings?**

Entity ambiguity

The problem of entity ambiguity is closely related to the problem of the unit of analysis. In a way, it is the question of how to get from the definition of what a firm "is" to the operationalization of including all entities that should be covered (ambiguity) exactly once (duplicates). One could add

that it is also about how to structure the entities, whether subsidiaries should be separate, or all consolidated. In my view, this might be the single most problematic aspect of the ORBIS database. The data is simply messy in this regard as is also documented elsewhere (Garcia-Bernardo and Takes 2018; Heemskerk et al. 2018). Frankly, in my view, the shortcomings in ORBIS regarding entity ambiguity cannot be stressed enough. To begin with, ORBIS simply does not only include data for firms but also other entities. The ownership information we downloaded for example also includes shares held in different kinds of museums, municipal utilities, non-profit associations or clubs, but also various central banks such as the ECB. It is not trivial to filter these out systematically, but it may to some extent be possible by excluding specific sectors and non-profit entities.

Entity ambiguity can further easily be illustrated by giving an example. Figure 5 illustrates the network of interlocking directorates between the Global Ultimate Owners of all firms containing the word "blackrock". Ties mean that at least one manager from one corporate group also sits on the board of another corporate group. In an ideal data world, one might expect *one* Global Ultimate Owner (GUO) or head of the corporate group, namely Blackrock Inc. If this was the case, the GUO variable could easily be used to aggregate all entities that are economically owned by the Blackrock corporate group in this case. However, as the network illustrates this is not the case. At first sight, the availability of the GUO variable may be an apparent advantage over other databases such as Compustat which does not include a comparable variable about the GUO. It must be noted however that while the GUO variable promises a lot, namely, to identify the common global ultimate owner of a corporate group, the variable is often erroneous. This was confirmed to me also in conversations with other researchers working with the data who have made similar experiences. There are numerous examples one could give in addition to the Blackrock example. Although in this case, only part of this is due to missing links in the ORBIS database and therefore a matter of accuracy or data quality. In part, these "firms" are funds or other subsidiaries which may factually be shared among different shareholders not only from the BlackRock corporate group. A few of them are also completely different firms that are called "Blackrock" but do not have ownership ties to the asset management corporation. Again, depending on the research question this might be problematic. However, it is relatively clear that in many cases the GUO does not provide the actual global ultimate owner which is also why some researchers chose to gather ownership information from other sources (Aminadav and Papaioannou 2020:1196–97). This is yet another illustration of the ORBIS problem that while the database includes a large sample of firms, in many research applications individual cases have to be checked and additional information has to be researched (and not the least matched) by hand.

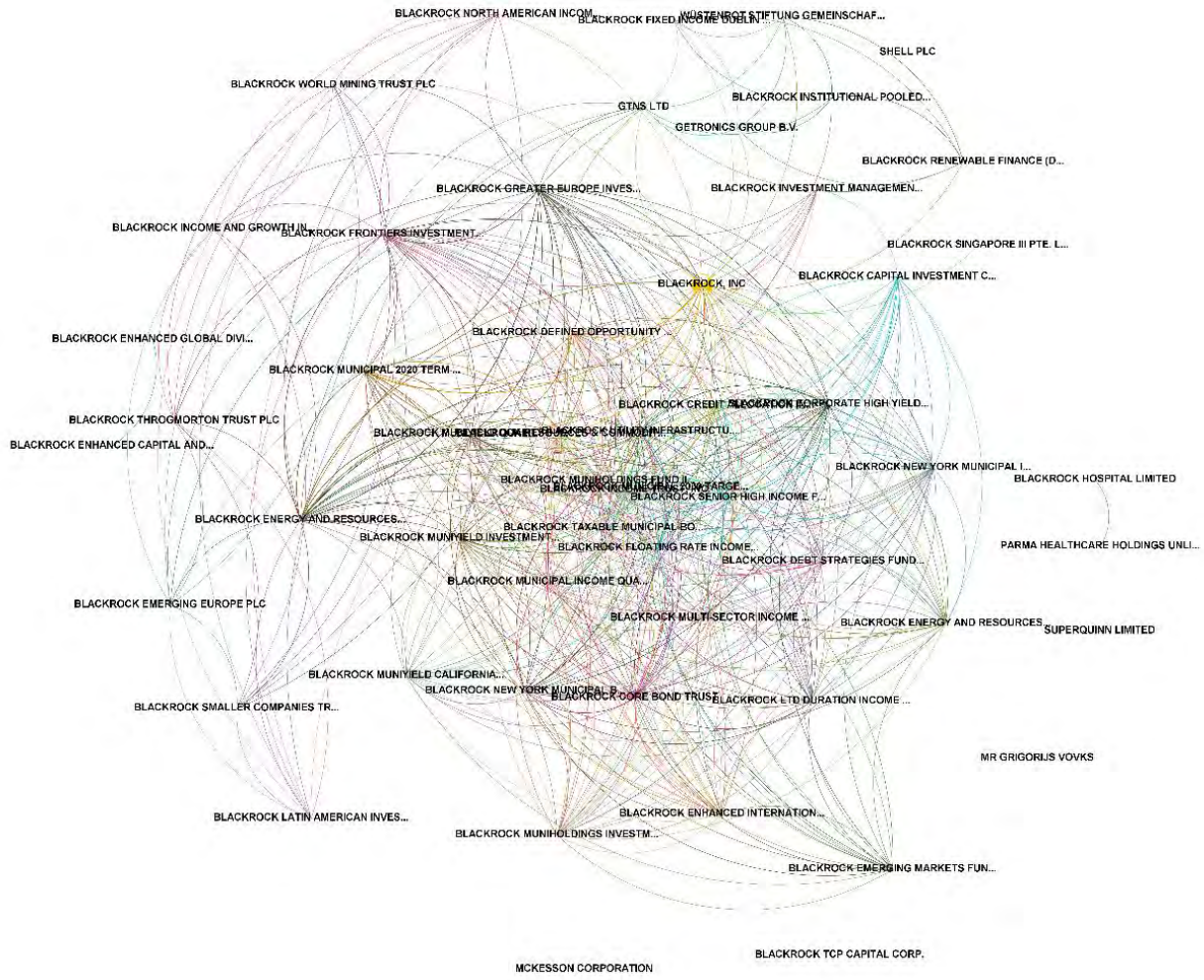


Figure 5: Who owns companies that are called "Blackrock"? The network of interlocking directorates between all Global Ultimate Owners (GUOs) owning any company with a name containing the word "blackrock".

But how to solve this issue of entity ambiguity of ORBIS? How to work around the fact that there are uncountable entities which somehow – depending on the definition – should belong to the same corporate group but according to the ORBIS data they do not? Once more, I cannot offer much hope for researchers looking for solutions completely accounting for this. One way to deal with the problem is to change the definition of the GUO to a lower threshold of shares of a company held. But this solves only part of the problem. Some literature suggests different forms of entity resolution methods based on string methods, exploiting additional data such as firm addresses, or topological methods based on networks such as the one illustrated in Figure 5 (Garcia-Bernardo and Takes 2018; Heemskerk et al. 2018). However, I was not successful in effectively improving the GUO data in ORBIS by applying any of these classes of solutions. In the following, I briefly summarize why.

In my case, I was working with a very large sample from ORBIS of around 3 million global firms and in addition around 7 million shareholding entities. The problem with data of such size is that almost every imaginable name or word will occur multiple times in the data. Blackrock is even probably one of the easier cases because you do have a relatively unique identifying word "Blackrock" that you could use. In fact, by hand, the problem may be solved with an hour of work maximum for this case. But this does not help much when you need to prepare hundreds of thousands of corporate groups and you cannot rely on the corporate group membership variables. I tried different ways of identifying the most "significant" word among a firm name, e.g., in terms of how frequently the word occurs in the whole corpus. In this example, "Blackrock" is a meaningful string to find companies of the corporate group while "Inc." is not. To give another easy example, many firms are named after their founding family. A lot of founding families are called Smith, Jackson, or Becker. So you might have strings like the fictional "Smith Bakery Shops New York", and "Smith Baker Law Firm", but also "New York Bakery Smithtown". These firms would of course be distinct corporate groups. Also, there are many firms with indistinct names such as "International Service Corporation" or "Credit Limited Holding" where there is no chance of identifying a meaningful part of the string that will help you to structure firms into correct corporate groups. [15]

It should also be mentioned here that comparing distances between strings for a large set of strings requires huge adjacency matrices and can get very costly in terms of working memory very quickly. Trying to compute the Jaccard distance between all 1.3 million Chinese companies in my sample would have required more than two terabytes of working memory. Splitting up matrices into subgroups based on some other criteria suggesting that firms may be part of the same corporate group or not may then become costly in terms of computation time.[16] There are suggestions to preselect the number of firms for which string comparison is made to reduce these matrices. But first, this is maybe a good time to constate that things get very complicated in terms of implementation, and you need good knowledge of a programming language and linear algebra to handle a large number of strings, matrices, distance calculations, etc. Developing and comparing good methods to reduce entity ambiguity can quickly become a research project by itself with a significant chance of failure. Most researchers will not even have the time to implement something that sufficiently accounts for this. But assuming such skills and time resources are available, this

[15] I am aware that there are many different string distance measures and matching methods, e.g., based on n-grams. But I do insist that none of them will solve the fundamental issue that with a sample of this size, most firm name strings become uninformative because they are simply not sufficient to decide whether two firms are part of the same corporate group or not.

[16] All of this depends on sample size of course. It may be feasible for a couple of hundred or thousand firms, but it is not for the millions of firms ORBIS claims to cover.

brings us away from the suggestion to simply use string comparisons to combine these with other measures, e.g., based on network topology.

In theory, and this is also suggested in the literature, one could use string matching algorithms as sketched above in combination with other methods such as topological network methods, although this requires network data and thereby does not apply to many applications only working with the ORBIS financial data. A very helpful of such method to identify duplicates are presented in Garcia-Bernardo and Takes (2018). It seems intriguing to use networks of overlaps in management or ownership relations to identify corporate groups. But several problems remain even using these methods. First, in both cases, information is not available for all firms. So, it depends on your sample for what share of companies and corporate groups this is even applicable. Second, the method relies on the GUO variable to identify corporate group membership which is often inaccurate, as it may not cover all potential duplicates, as discussed above. Third, one must find suitable thresholds of closeness in the network within which firms can safely be assumed to be in the same corporate group. While this works well to identify duplicates (see below), I was not successful in reliably identifying other members, and only other members, of a corporate group with these methods. Especially based on ownership data, many companies may have very similar network positions but may not be part of the same corporate group. When adding string matching to this I ran tono problems on both sides: many firms do not share any part of the string but are part of the same corporate group, and some firms share a string but are not part of the same corporate group. Merging those entities that do have similar network positions and share a string should be fine but this way you do only solve a (small) part of the problem.

A second aspect of entity ambiguity illustrating this is simply duplicate firms, i.e., entities that occur multiple times but with differing unique IDs in ORBIS. This may be the case with the same firm names or slightly different firm names and also applies to shareholding entities and managers. As Kalemli-Oczan et al. (2022b:5) report, such duplicates may also be introduced when combining snapshots from different time points and firms' IDs changed in between. Duplicates are of course a problem that may distort results. BvD claims in the ORBIS online support documents that an external data science firm already removed duplicates as far as possible. Nevertheless, a relevant number of duplicates remains. Garcia-Bernardo and Takes (2018) developed an effective method to identify duplicate firms. Based on the network of interlocking directorates, they identify those firms whose overlap in management, that is the same people employed as managers, suggesting that they are referring to the same firm. Applying this method to the Swedish network of interlocking directorates, the authors reduce the original number of entities including subsidiaries by 47% (!) (Garcia-Bernardo and Takes 2018:169). In my work where I excluded all subsidiaries, I applied the procedure suggested by Garcia-Bernardo and Takes, as

well as simple name duplicates. With these methods, I find 21,548 duplicate entities (7%) only among the 290,169 heads of corporate groups in my sample. One shortcoming of this method, however, is that it is only feasible at the national level (at least without adjustment), and for a manageable size of firms. Finding duplicates among all Chinese firms, for example, will exceed working memory capacities almost certainly with this method.

The problem of duplicates might even be more precarious among shareholders. It is easy to spot duplicate shareholding entities when simply checking individual cases by hand. For one of my research projects, I ran a supervised learning procedure to deduplicate the ORBIS shareholder data for my sample. I trained an algorithm for deduplication and matching based on multiple variables and their interactions (Dedupe.io 2022). The algorithm was trained to learn based on 1000 manually labeled cases whether two records are referring to the same shareholding entity based on their name, country, city, shareholder type, postcode, founding year (if a company), first name, and last name (if individuals). Among the 17,207,181 shareholders in my sample, I identify 186,737 (1%) duplicate entities with enough certainty.[17] However, this only captures a small share, and I spotted numerous additional duplicates among the remaining shareholders which could not be captured with enough certainty in this way. Again, the degree to which duplicates among firms, shareholders, and managers distorts results depends on the research question, sample, and research design. Yet, it is a problem that needs to be accounted for. In summary, the takeaways for researchers and reviewers concerning entity ambiguity are the following questions:

1. **How do you/ does the reviewed work operationalize the unit of analysis or classify firms as belonging to a firm or corporate group?**
2. **Does the work rely on the Global Ultimate Owner variable from ORBIS? Is it in line with the definition of "what is a firm"?**
3. **Which inaccuracies in corporate group membership does this introduce and tolerate? Is it relevant to the results?**
4. **How do you deal with duplicate firms, shareholders, and managers? Do they distort results?**

Accuracy

The problem of accuracy is one of incorrect information. The ORBIS database includes numerous errors.[18] As was described along with the background information, the data comes from different

[17] The Shareholder – BvD ID numbers of these duplicates, a unique identifier, and the matching score can be downloaded from my GitHub page.

[18] I can only mention as a side note here that this and other problems may also be very relevant for business users of ORBIS and the decisions which are based on it. More specific examination of the ORBIS data for use cases from the business world are necessary to assess this more systematically.

data providers. This may be one reason for the inaccuracies. I have spotted several such examples. Again, I could not think of a method or illustration to oversee and describe the full extent. Heemskerk et al. (2018:25–27) suggest diagnosing and fixing inconsistencies by checking for outliers which is an intuitive and helpful first step. However, this does not account for all types of inaccuracies. Figure 6 shows an excerpt from the ORBIS portal showing the number of employees of an American Branch of Munich RE over ten years. It ranges between 0 and 43,000 but is on average around 900 employees with a jump from 0 to 40,000 and back to 0 within three years. Something is wrong with this record, but it remains unclear without further research into the database whether this error is due to some definitional problem or rather correct information for some unintuitive reason. Assuming it is an error: It is true, that problems such as these may be captured and fixed by looking for outliers, but here outlier detection needs to be looking at time series and from my assessment, it will be very difficult to accurately separate actual erroneous outliers from correct large fluctuations for a large number of cases in this way.

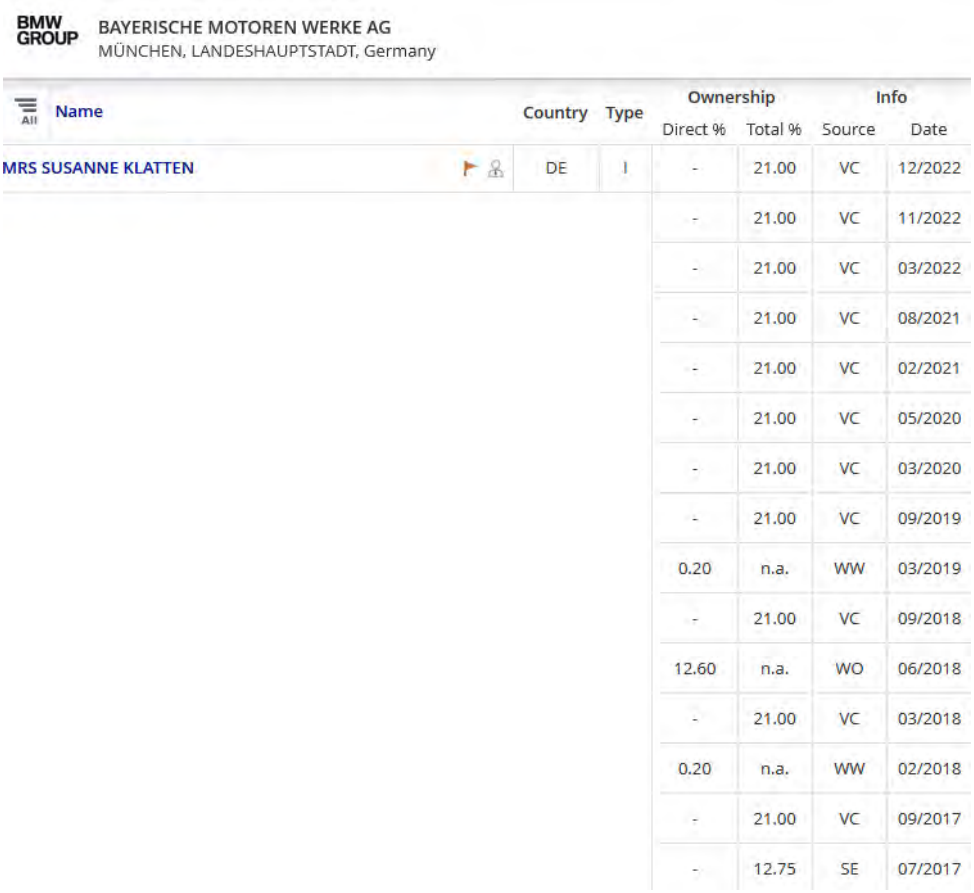
	Company name Latin alphabet		Country ISO code	Number of employees Last avail. yr	Number of employees Year - 1	Number of employees Year - 2	Number of employees Year - 3	Number of employees Year - 4	Number of employees Year - 5	Number of employees Year - 6	Number of employees Year - 7	Number of employees Year - 8	Number of employees Year - 9
1.	MUNICH RE AMERICA CORP.		US	0	539	844	964	993	0	43,000	0	1,002	996

Figure 6: Inaccuracies in the number of employees of Munich RE America Corp (screenshot from the ORBIS portal).

I cannot elaborate here on all forms of inaccuracies in the data that I have spotted. A very illustrative example is that the two largest asset managers globally, BlackRock and Vanguard, are wrongly classified in the shareholder data as a bank, and as a non-financial corporation. Some other examples from the ORBIS financial information are also implied in Kalemli-Oczan et al.'s (2022b:22) suggestions to exclude outliers in terms of total assets and sales. I also do believe that there is not one systematic method to deal with them and whether they are relevant for research results depends on the exact sample used and the research question, as well as the kind of error. The only contribution I can make here in this regard is to draw attention to the fact that such errors exist, and one needs to be careful in research applications. However, I would like to mention one more example that might be relevant for researchers analyzing ownership data. It comes from my research on super-rich ownership and management of firms. Namely, the problem is the interchangeability of direct and total shares depending on the data provider. Figure 7 illustrates this for the case of Susanne Klatten, one of the richest Germans, and her history of shareholdings in car maker BMW. The column source denotes the data provider. One can see how entries from different data providers lead to inaccuracies, or at least potential for error in data handling, in the time series but also the meaning of the variables. The entry of total shares from September 2017

stemming from SE is lower than all other values reported by VC. It is much closer to Ms. Klatten's direct share reported by WO in 2018. This, however, stands in stark contrast to the other direct shares reported by WW for two-time points. The missing values may be distorting interpretation and interpolation attempts as well which needs to be taken into account carefully in data handling. Which of these values is the correct direct and indirect ownership information at the respective time points? It seems impossible to make this call without consulting external sources. And this is just one example of one of the largest and most famous shareholders, owning shares in one of the largest companies in one of the largest economies in the world which counters hopes articulated in the literature, that data quality in ORBIS is good enough the larger the companies and the larger the countries' economies – especially in Europe.

The most constructive comment I can make on this specific point is that I have made the best experiences for the research task of identifying large owners, i.e. super-rich individuals and families, by using mostly total shares, interpolating gaps between total shares, and if entries are still missing, complementing missing total shares with direct shares. This is because, in the matching procedure, it was a necessary precondition that not only the name of the super-rich individual match the shareholder's name but that it also actually holds a share in a matching firm name. Through this data preparation procedure, I maximized the number of identified super-rich individuals among the largest firms. However, when truly interested in the distinct meanings of direct and total shares and using the data for this, I can only advise against it based on the inconsistencies in the data in this regard.



BMW GROUP BAYERISCHE MOTOREN WERKE AG
MÜNCHEN, LANDESHAUPTSTADT, Germany

Name	Country	Type	Ownership		Info	
			Direct %	Total %	Source	Date
MRS SUSANNE KLATTEN	DE	I	-	21.00	VC	12/2022
			-	21.00	VC	11/2022
			-	21.00	VC	03/2022
			-	21.00	VC	08/2021
			-	21.00	VC	02/2021
			-	21.00	VC	05/2020
			-	21.00	VC	03/2020
			-	21.00	VC	09/2019
			0.20	n.a.	WW	03/2019
			-	21.00	VC	09/2018
			12.60	n.a.	WO	06/2018
			-	21.00	VC	03/2018
			0.20	n.a.	WW	02/2018
			-	21.00	VC	09/2017
			-	12.75	SE	07/2017

Figure 7: Shareholder history of Susanne Klatten's ownership in BMW in ORBIS stemming from different sources (screenshot from the ORBIS portal).

To close the section on accuracy, researchers should consider questions along the following:

1. **How do you/ does the reviewed work check for erroneous data?**
2. **How are values over time interpolated? Are these checked for inconsistencies?**
3. **Has a serious diagnosis of accuracy taken place? Can it be visualized or otherwise demonstrated?**

Completeness

Two types of completeness must be separated and discussed distinctly. First, this is the completeness of the population, or at least the representativity of samples, from the ORBIS database. Second, this is missing information among variables provided in ORBIS. I will begin with only a few summarizing sentences on completeness and representativity since this is discussed in detail elsewhere (Bajgar et al. 2020; Garcia-Bernardo and Takes 2018; Kalemli-Ozcan et al. 2022a). I will then concentrate more on missing information after that.

Among other things, existing contributions have demonstrated that coverage in ORBIS is better for countries with higher GDP and varies by region (e.g. Garcia-Bernardo and Takes

2018:166–68). This might for example be due to varying legal requirements of limited liability companies to register, e.g. depending on their size (Kalemli-Ozcan et al. 2022a:6). Bajgar et al. (2020) report that representativity of the ORBIS data varies by year, by country, and even by industry. Coverage of large and top-performing firms seems to be good, however weighting of the sample to make it more representative does not seem to be possible since inclusion in ORBIS is non-random even when taking size into account (Bajgar et al. 2020:9). The authors conclude that "Overall, Orbis seems more suitable for studies that: i) take a global perspective rather than making comparisons across countries; ii) analyse top performers and multinationals rather than underperforming firms; iii) and focus on mean performance or within-firm changes rather than on the entire firm distribution or entry and exit" (Bajgar et al. 2020:3). All these contributions compare samples from ORBIS to external statistics and should be consulted for further details.

A distinct question is that of missing information. Financial data is generally only available for a small share of all companies in ORBIS. This was already illustrated in section two. Figure 8 visualizes the extent of missingness for some financial variables separately for six samples. First listed (a) vs. unlisted firms (b), second US S&P 500 listed firms (c) vs. Chinese Shanghai Stock Exchange listed firms (d), and third, the largest listed and unlisted firms in terms of revenue (e) vs. the largest global listed and unlisted firms in terms of their number of employees (f). All data comes from our second download in 2022. Since completeness in ORBIS depends on accounting regulations and there are more strict regulations for listed firms, it seems intuitive that listed firm data is more complete than non-listed firm data. For the latter, even the most widely available variables have a missing share of 75% among the ca. 3m largest non-listed firms globally. Also, the completeness of variables differs by listed and unlisted firms. Profit and loss for example seem to be more available for listed firms and the number of employees for unlisted firms.[19] Among listed firms, completeness of information seems to be similarly high at least for large economies as shown here for the US and China, and very similar numbers are also found for listed German firms. When going beyond listed firms, the completeness of information decreases rapidly as can be seen when comparing the completeness of data for the largest firms (Figures 8e and 8f) with data of listed firms only (Figure 8a).

[19] As discussed in section 3, I do not believe that the high share of missingness is due to the method of access and download. However, to fully clarify this it would be good to see a comparison with the missingness structure from the historical disks.

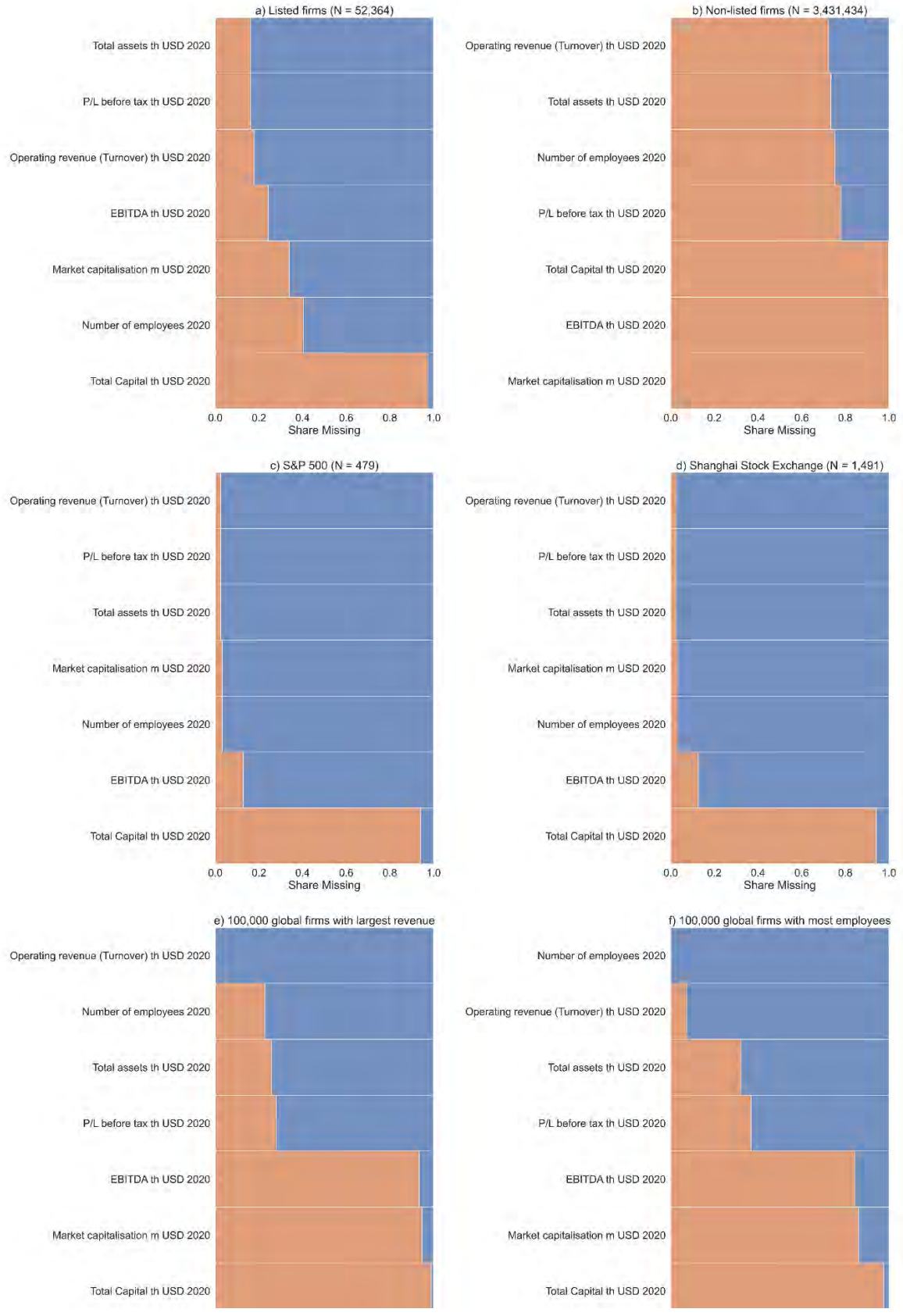


Figure 8: Degree of missingness of some key financial variables in 2020. Orange: Share of missing records. Blue: Share of non-missing records. a) All listed firms, b) all non-listed firms, c) S&P 500 listed firms, d) SSE listed firms, e) 100,000 largest firms in terms of revenue, and f) largest firms in terms of the number of employees. The download took place between January and June of 2022.

Bajgar et al. (2020), following Gal (2013), suggest imputing missing values and showing promising results for the imputation of value-added. However, as the authors acknowledge this requires the availability of other information such as the cost of employees which is in many cases also not available. Yet, they present evidence showing that "Internal imputation substantially increases coverage for about half the countries in the sample, and it reduces it for Norway. […] The internal imputation makes the firm distribution in Orbis more representative in terms of mean firm characteristics, and it moves the level of heterogeneity in the bottom half of the productivity distribution closer to that observed in the population […]. Using internal imputation moves average firm size and labor productivity closer to that observed in official microdata. It also increases the productivity dispersion in the lower half of the distribution" (Bajgar et al. 2020:32). Unfortunately, this conclusion conflicts with the fact that international comparison is often affected by differing reporting standards as gathered by Kalemli-Oczan et al. (2022b:73–76) or the meaning of e.g. management positions as constated by Heemskerk et al. (2018:8): "different countries have different governance structures, rules and regulations. A non-executive director in China is not the same as a non-executive in the UK. A big data approach easily allows for the study of, for instance, board interlocks across the globe, but decontextualizing boards and firms may lead to invalid conclusions."

From my experience imputation is a task that should not be underestimated and some of the imputations of market capitalization I performed based on random forests led to clearly wrong estimates. Most likely these were related to my sample choice for imputation, i.e. which firms and subsidiaries to include and which ones not, as well as account types and double counting of heads of corporate groups and subsidiaries. I then had to cut on further efforts due to the time restrictions of my project. For many tasks, imputation of small firms, branches, and subsidiaries may make decrease data quality because missingness itself says something about the firm in ORBIS (country, industry, firm type, reporting, and accounting requirements, transparency of companies such as potential shell companies, etc.). Not all of this information is available as variables that can be used for imputation which means the data is not missing at random and imputation will distort results to an unknown extent. Imputing the revenue of all companies in a corporate hierarchy for example including administrative companies will inflate the revenue of a corporate group. These kinds of problems possibly can be circumvented by choosing the right sample such as only heads of corporate groups and being aware of consolidated and unconsolidated account types, but this solution is dependent on the correct identification of subsidiaries as discussed above as entity ambiguity.

As a final problem of completeness to be mentioned here, I would like to refer to Bajgar et al. (2020) once more who underline that some reported financial data in ORBIS is only imputed by

estimates. This leads to the problem that for firms with limited financial accounts, the number of employees and operating revenues are only estimated as the middle of a range especially "for the Czech Republic, Poland, Slovakia and the United States [and] a smaller, but non-negligible, role for Finland, Japan, Sweden and Italy" (Bajgar et al. 2020:49). To close the section on completeness, researchers should consider questions such as the following:

1. **Does representativity/ coverage matter for results?**
2. **If the answer to 1 is yes, which countries/industries/firm sizes does your work/ does the reviewed work include? What do the existing external validity checks of ORBIS coverage say about coverage in these countries (Bajgar et al. 2020; Garcia-Bernardo and Takes 2018; Kalemli-Ozcan et al. 2022a)?**
3. **Which account types are chosen and how are missing values dealt with?**

5. Reflections

In the following, I offer some reflections on, and defense of, the points made in this article. These thoughts deal with the questions of when to believe ORBIS data after all, the origins of the problems in the political economy of corporate corporate data, the question of when a dataset is just not good enough, and finally, the idea of public goods as an answer.

When should we believe research using ORBIS firm data?

In my opinion, three things can increase trust to an acceptable level. First, authors need to show that they have engaged with the existing literature and the manifold known data problems. That is, at this point and in my view, especially Kalemli-Oczan (2022b, 2022a, n.d.)[20], Bajgar et al. (2020), and Heemskerk et al. (2018). Second, authors must deliver evidence that they have engaged in-depth with their specific sample, problems that are relevant to their research questions either already known from the literature or newly discovered by themselves. This includes being explicit about their access method of the ORBIS data, being transparent about their choices of data preparation, and at least how they tackled known existing problems. I hope that the questions gathered above may help to structure the necessary documentation of the use of ORBIS data for research. Finally, in this regard, authors should be explicit about potential errors known from the literature that they could not solve and that they, therefore, tolerate in their interpretation of results. Third, and possibly most demanding, research should ideally be reviewed by other researchers who have reasonable experience with or knowledge of ORBIS data. I do try to give a guideline here for

[20] But consider the diverging experiences we made concerning downloads from the web portal reported in section 3.

reviewers less familiar with the data of where to point the finger, but to fully evaluate which of the problems apply likely more knowledge of the data is necessary. This might be a difficult claim since the amount of researchers knowing the data as far as I can tell is very limited. It might be worthwhile to collect a public list of researchers, or set up a community some other way, of researchers familiar with the data. This could be a source from which journals could draw anonymous reviewers for the data side, or researchers working with the data could help each other out. While one could argue that this is also not done for most other datasets, I do think that the number and complexity of problems with the ORBIS data is larger and more pressing than for most other frequently used datasets for research and it is therefore necessary. I do believe (and have in fact witnessed) that in cases where researchers are not familiar to the downsides of the dataset and that there is a whole literature about it, the large promises of ORBIS – like Homer's sirens - may make blind for its perils and in the worst-case lead to wrong scientific knowledge that could have been prevented.[21]

The political economy of corporate corporate data

One could object to my comments elaborated here that I am overly pessimistic about the data, as well as the research community's use of it, and not making a constructive contribution.[22] Also, similar problems as the ones of ORBIS are known from other big data sources, as is the potential messiness of big data in general. Is this just a situation of being skeptical towards technological and methodological progress as were people of railroads in the 19th century? After all, the alternative would be not to use the data at all. Therefore, I would like to defend and justify my position before closing on the most constructive note I can. I do believe that the exploitation of large data sources is the methodological future of social sciences. However, social sciences were always known for critically reflecting on their research methods and I believe this is one of its main strengths. The sources of the problems with ORBIS from my assessment are the following. A commercial enterprise gathers prepares and sells the data mainly to a target group that has very different needs and demands than the research community. Since the enterprise is commercial, it does matter that customers from the business world probably make up the majority share of BvD's revenue with the data. Selling the data to scientist rather seems to be a side business for BvD that came later in its company history. This origin leads to the fact that the higher demands of researchers in terms of e.g., representativity, harmonization and comparability across countries, historical data and its

[21] Which in the case of ORBIS is the more problematic because there is no other dataset available or in sight for many questions to disprove findings based on other more trustworthy sources.

[22] I do also have my problems with the imperative of constructiveness and this case made here exemplifies some of the reasons why, but debating this would be beside the point here.

consistency, entity ambiguity, etc. were not considered in building the database from the start. Existing efforts e.g., to reduce duplicate individuals in the database by external data engineering companies cannot fully account for this. At the same time, ORBIS holds a monopoly position, especially for the international perspective as well as unique combinations of information e.g., of shareholders, management, and financials. Since BvD is a commercial enterprise, they have an interest in making a profit and not in being proactively transparent about shortcomings of the data for the generation of reliable scientific knowledge. After all, which company is proactive in advertising the shortcomings of its products? This is of course a fundamental difference between publicly financed datasets such as many national panel surveys. In sum, therefore, I do believe that ORBIS' history and legacy, as well as the fact that it is owned and maintained by a commercial enterprise with a different target group, are the source of its specific problems. It is therefore more problematic in this sense than other data sources. But is it too problematic to use for scientific research? This is the final point I would like to offer some thoughts on.

Just not good enough?

Heemskerk et al.'s (2018) ambition is "to begin a conversation about research process standards now if we are to advance the quality of the research community in the future." I very much agree that this is necessary, and my point here is that we need standards when a (e.g. new or particularly large and complex) data source is not good enough for research. Good enough is itself a matter of definition of course, and I do mean it here in the sense that on average over all applications it could lead to too many invalid inferences at least for very large and international samples – which is exactly the kind of applications the dataset invites itself most for in comparison to other data. In the case of ORBIS, I disagree with the authors that the necessary means to "assess the extent to which data quality issues exist and what it means for the meaning that we derive from the analysis of concern" (Heemskerk et al. 2018:6) are presented by them, are yet available, nor will ever be available as long as the data does not fundamentally improve. This, however, is an empirical question. What would be necessary is a large reproducibility initiative of trying to reproduce existing (if only the most influential) research results generated with ORBIS data. I would however not say much about their validity of course.

Public goods

To close on a more constructive note, the wide use of the ORBIS data despite its problems shows that there is a high demand for researchers for large-scale international firm data. Questions that can be answered with such data are highly relevant to the lives of many. Possibly an initiative of

well-funded international research institutions and their libraries or some other public institutions could collaborate to set up a non-commercial alternative to ORBIS. One that is constructed with the needs of public research and reliable scientific knowledge in the first place. Comparable examples although smaller in scale already exist, such as the open global firm database OpenCorporates (2023) or the related leaked German company register (OffeneRegister 2023). Until these or other alternatives can compete with ORBIS, the open question remains how many problems it needs for a data source not to invite itself for the generation of scientific knowledge – or at least only under very specific and highly costly circumstances which are especially years of experience and the necessary skills to use the data. With the comments and questions raised here, I hope to help reviewers and authors to conclude for themselves at least as long as there is no comparable alternative data source, no significant improvement of the data quality on the side of BvD, and no more systematic external checks of the data's validity and potential for reproducibility.

References

Aminadav, Gur, and Elias Papaioannou. 2020. "Corporate Control around the World." *The Journal of Finance* LXXV(3):1191–1246.

Bajgar, Matej, Giuseppe Berlingieri, Sara Calligaris, Chiara Criscuolo, and Jonathan Timmis. 2020. *Coverage and Representativeness of Orbis Data*. Paris: OECD. doi: 10.1787/c7bdaa03-en.

Bureau van Dijk. 2020. "Orbis. Powering the Business of Certainty." Retrieved March 4, 2020 (https://www.bvdinfo.com/en-gb/our-products/data/international/orbis).

Dedupe.io. 2022. "Dedupe." Retrieved March 20, 2023 (https://github.com/dedupeio/dedupe).

Gal, Peter N. 2013. "Measuring Total Factor Productivity at the Firm Level Using OECD-ORBIS." *OECD Economics Department Working Papers*.

Garcia-Bernardo, Javier, Jan Fichtner, Frank W. Takes, and Eelke M. Heemskerk. 2017. "Uncovering Offshore Financial Centers: Conduits and Sinks in the Global Corporate Ownership Network." *Scientific Reports* 7(1):1–10.

Garcia-Bernardo, Javier, and Frank W. Takes. 2018. "The Effects of Data Quality on the Analysis of Corporate Board Interlock Networks." *Information Systems* 78:164–72.

Group van Dijk. 2013. "Presentation of Van Dijk Management Consultancy Companies." *Group van Dijk*. Retrieved March 20, 2023 (https://web.archive.org/web/20150708012732/http://www.vandijkmc.com/en/group-van-dijk_35.aspx).

Heemskerk, Eelke M., and Frank W. Takes. 2016a. "Supplementary Material for 'The Corporate Elite Community Structure of Global Capitalism' by Eelke Heemskerk & Frank Takes." Retrieved January 17, 2023 (https://pure.uva.nl/ws/files/67005080/cnpe_a_1041483_sm2199.pdf).

Heemskerk, Eelke M., and Frank W. Takes. 2016b. "The Corporate Elite Community Structure of Global Capitalism." *New Political Economy* 21(1):90–118. doi: 10.1080/13563467.2015.1041483.

Heemskerk, Eelke M., Kevin Young, Frank W. Takes, Bruce Cronin, Javier Garcia-Bernardo, Lasse F. Henriksen, William Kindred Winecoff, Vladimir Popov, and Audrey Laurin-Lamothe. 2018. "The Promise and Perils of Using Big Data in the Study of Corporate Networks: Problems, Diagnostics and Fixes." *Global Networks* 18(1):3–32. doi: 10.1111/glob.12183.

Kalemli-Ozcan, Sebnem, Jingting Fan, and Veronika Penciakova. n.d. "Processing ORBIS Historical Disk." Retrieved January 12, 2023 (http://econweb.umd.edu/~kalemli/assets/policy&blog/Processing%20Orbis%20Historical%20Disk%20FINAL.pdf).

Kalemli-Ozcan, Sebnem, Bent Sorensen, Carolina Villegas-Sanchez, Vadym Volosovych, and Sevcan Yesiltas. 2022a. ""How to Construct Nationally Representative Firm Level Data from the ORBIS Global Database: New Facts and Aggregate Implications."" National Bureau of Economic Research, Cambridge, MA.

Kalemli-Ozcan, Sebnem, Bent Sorensen, Carolina Villegas-Sanchez, Vadym Volosovych, and Sevcan Yesiltas. 2022b. ""Online Appendix to How to Construct Nationally Representative Firm Level Data from the ORBIS Global Database: New Facts and Aggregate Implications."" National Bureau of Economic Research, Cambridge, MA.

Liu, Grace. 2020. "Data Quality Problems Troubling Business and Financial Researchers: A Literature Review and Synthetic Analysis." *Journal of Business & Finance Librarianship* 25(3–4):315–71. doi: 10.1080/08963568.2020.1847555.

OffeneRegister. 2023. "Licht Ins Dunkel Der Firmengeflechte." Retrieved March 20, 2023 (https://offeneregister.de/).

OpenCorporates. 2023. "The Largest Open Database of Companies in the World." Retrieved March 20, 2023 (https://opencorporates.com/).

Reurink, Arjan, and Javier Garcia-Bernardo. 2020. "Competing for Capitals: The Great Fragmentation of the Firm and Varieties of FDI Attraction Profiles in the European Union." *Review of International Political Economy* 1–34.

Selenium. 2023. "The Selenium Browser Automation Project." Retrieved February 27, 2023 (https://www.selenium.dev/).

2nd Article:

The Peak of Capital?[1]

Super-Rich Capitalist Families in Global Interlocking Directorate Networks

Abstract

Do capitalist families concentrate economic power through coordinating capital in networks of interlocking directorates? This study analyzes how super-rich capitalist families are embedded in the 2020 network of interlocking directorates common to the world's largest 254,207 firms and their 1,473,149 top managers and directors. I identify 4,250 family members from 2,313 super-rich families, each with an estimated family net worth of US$250 m or more. I map these family members in the networks between firms and between individuals. Factors facilitating super-rich participation in global and national networks of interlocking directorates are tested using logistic regression models. Results show that, first, directors belonging to super-rich families are more likely to be embedded in corporate elite networks than are other directors. Family-level explanatory variables such as family wealth and size of a dynasty further increase the odds of forming part of the largest connected component. The share of firms with any type of super-rich involvement increases towards the core of the global network – but only up to a share of about 20% of all firms. Second, two individuals' being members of two different super-rich families does not increase the odds of their sitting on the same board. This speaks against the idea, suggested by the class cohesion model, that super-rich capitalists coordinate on corporate boards because they are driven by shared interests based on the commonality of owning or controlling capital. Instead, large, listed firms seem to hire members of super-rich families to supervisory positions, possibly to utilize either their prestige, as a signal of credibility for stakeholders, or their contacts, as potential clients, and investors.

Submission status:

Accepted for publication in Revue Française de Sociologie.

[1] I would like to thank Olivier Godechot for fruitful continuous discussions while I was working on previous versions of this work. Further thanks go to Nils Neumann, Daria Tisch, Emily Anne Wolff, and Kevin L. Young for their helpful comments on earlier versions. This work was presented at the European Graduate Network Conference, April 20–21, 2022 in Florence, and the summer school on "Elites, Power and Inequalities" in Lausanne, June 20, 2022. Special thanks for the very helpful comments and critiques by these events' participants, and especially the discussants Magdalena Breyer, Stéphanie Ginalski, Andreas Jozwiak, Rebecca Lévy-Guillan, and Nora Waitkus. Further thanks to the participants of the presentations at the meeting of the research group on wealth and social inequality at MPIfG Cologne on May 17, 2022, the MaxPo colloquium at SciencesPo Paris on May 23, 2022, and in the virtual corporate network session at Sunbelt Conference on July 13, 2022. Sincere thanks also to the administration and IT department of the MPIfG Cologne, without whose efforts and support in helping me to acquire the necessary hardware and software I could not have carried out the analyses presented here.

1. Introduction

The wealth held by the world's richest individuals and families has increased tremendously over the past decades. The estimated wealth held by global billionaires has grown from about US$2 tn in 2008 to more than US$10 tn in 2020 (UBS and PwC, 2020, p. 35). In 2014, Thomas Piketty suggested as one of his central hypotheses that return on capital was larger than the average growth rate (r>g) and that this trend, and therefore wealth concentration was projected to continue (Piketty, 2014). Wealthy individuals were already central to social thought more than a hundred years ago (Marx, 1887; Sombart, 1913). Karl Marx argued that concentration of control over capital can be the source of economic and therefore also political power in modern societies. Following his distinction between a class in-itself and a class for-itself, one could ask whether this group's members share anything other than similar positions within the relations of production or wealth distribution. This question has, however, been left aside in recent decades because separation of ownership and control of capital was assumed, which would make ownership less relevant in terms of economic power (Berle and Means, 1991 ; Mizruchi, 2004). After the famous claim that capital is back (Piketty, 2014), it is time to re-examine this assumption.

The present study aims to shed light on the potential of the globally largest capitalist families to be considered as a community, and therefore as concentrating and coordinating massive economic power as a group. I analyze one classic channel through which this may take place. This is embeddedness in networks of interlocking directorates, i.e. membership of the same company boards. Connecting different firms and individuals in this way is a classic form of economic network to which different functions have been ascribed, such as information exchange, facilitating friendships, assimilation and cohesion of its members, and concentrating capital. The guiding question of this study is: *Do capitalist families concentrate economic power by coordinating capital in networks of interlocking directorates?* In other words, can we find a "peak of capital" at the center of interlock networks where the richest capitalists coordinate with each other? The study uses the following definitions. First, the population of interest is defined as all families whose estimated net worth exceeds US$250 m – here understood as "the super-rich" – *and* who own or control business assets.[1] According to wealth data provider Wealth-X, the global population of individuals possessing US$250 or more was 19,640 in 2020 (Wealth-X, 2021, p. 5).[2] It is therefore a bigger

[1] The exact threshold of 250 m does not have a substantive meaning, but simply reflects the limit of our budget covering labor costs to collaborate the global wealth lists. It may be useful for the purposes of interpretation to note that this wealth level is lower than the famous billionaire category, but above the threshold of US$50–100 m which "most of the better wealth advisory firms have [as] minimum account sizes" (Curtis, 2020, p. 1).

[2] Other data sources report similar numbers with Credit Suisse stating the group of those possessing US$ 100 m and more roughly at 30,000 (Credit Suisse, 2021, p. 33 ; Wealth-X, 2021, p. 5).

group than the oft-referenced number of roughly 2,755 billionaires (Forbes, 2021) and it hence allows us an understanding of the top level of wealth globally that is less visible. This study analyzes a list with information on 11,406 individuals from 9,518 families, which is equivalent to a sample containing about 58% of this group's estimated global population.

The second defining criterion, that of being involved in a company, is likely only a marginal selection criterion to narrow this group down. Here, "capitalist" families means those that generate income from business assets (rentiers) or control the business assets they possess via management positions (owner-managers). The higher the position in the wealth and income distributions, the larger the share of business assets making up an individual's wealth portfolio (Keister, Li, and Lee, 2021 ; Schröder et al., 2020, pp. 41–42 ; Smith et al., 2019). The vast majority of the global fortunes we know of were once generated by founding a company, and even if shares in that original company were sold, asset management is, as far as we are aware, in most cases still organized through family offices or other holding companies.

The next section presents the recent theoretical and empirical literature on the interlocking directorates and their functions that this study speaks to. It is followed by, data, samples, and methods, the results section presenting explorative evidence, and evidence for and against the developed hypotheses, and is followed by the conclusion.

2. Concentration of capital and interlocking directorates

While interest in the world's wealthiest individuals and families has only been reawakened in the past decade, theoretical concepts with which to understand this group were being proposed in sociology as early as its beginnings in the 19th century. Most notable here is the tradition of social class analysis, which always entailed viewing the owners of the means of production as one central category (Marx, 1887) but also the concepts of (economic) elites holding powerful positions in modern societies (Mills, 2000 ; Mosca, 1939 ; Pareto, 1961 ; Sombart, 1913). From a Marxist perspective, it was suggested that the owners and managers of capital coordinate large amounts of capital, as a capitalist class, through shared boardroom membership (Hilferding, 1955; Jeidels, 1905). Zeitlin (1989; Zeitlin, Ewen, and Ratcliff, 1974) claimed that even after the separation of ownership and control (Berle and Means, 1991), it was capitalist families who controlled substantial shares of the largest economies through complex mechanisms of ownership, control, and family ties.

In recent decades, only a small number of sociological studies have analyzed interlocking directorates systematically as a mechanism of cohesion of a capitalist class (e.g. Allen and Broyles, 1989 ; Banerjee and Murray, 2020 ; Burris, 2001, 2005 ; Mizruchi, 1992 ; Murray, 2017). Since Michael Useem (1984) presented the idea of highly interlocked communities of top managers, the

existence and political influence of such "inner circles" has been shown frequently for different countries (Bond, Glouharova, and Harrigan, 2010; Broyles, 1993; Comet, 2019; Heerwig and Murray, 2019; Larsen and Ellersgaard, 2018). With the availability of off-the-shelf large-scale datasets including ownership relations and information on management (Bureau van Dijk, 2020; Heemskerk et al., 2018) it has become easier to analyze the national and transnational networks of interlocking directorates. This is probably also why the number of large-scale studies on interlocking directorate networks comprising up to one million firms has increased in the first two decades of the 21st century (Heemskerk, 2007, 2013; Heemskerk and Takes, 2016; Valeeva, 2021).

Perhaps the most striking finding of interlock research in the last twenty years is that national interlock networks are fracturing in almost all the countries studied (e.g. Beyer and Höpner, 2003; Chu and Davis, 2016; David and Westerhuis, 2014; Mizruchi, 2013). The most prominent explanation suggests a fracturing of the corporate elite because capital had won the battle against labor that had begun in the 1970s, making coordinated action less necessary (Mizruchi, 2013). Another important factor is that in the wake of fraud scandals in the USA in the early 2000s there were changes in law and hiring preferences, which meant that the demand for highly interlocked directors from corporations was reduced (Chu and Davis, 2016). It could be, however, that "ebbs and flows" in the density of the overall corporate interlock network are a feature of the historical development of capitalism and that they may strengthen again when coordination becomes more necessary (Francois and Lemercier, 2014). Despite this evidence of a fracturing corporate elite and declining national networks of interlocking directorates, data continues to be found that shows interlocks' prevailing relevance; for example, for firm political action such as party contributions. Where they exist, therefore, interlocking directorates still seem to fulfill a coordinating function for those connected (Banerjee and Murray, 2020; Goerres and Höpner, 2014; Heerwig and Murray, 2019; Murray, 2017; Schoeneman, 2022).

To sum up what is known about the relationship between interlocks and super-rich capitalists from the literature so far, the theoretical concern, stemming from Marxism, is whether it is plausible to think of capitalist families as concentrating and coordinating economic power through networks of interlocking directorates, and thereby as potentially exerting exceptional structural power, especially in capitalist democracies (Hilferding, 1955 ; Zeitlin, 1989). In sociology, interlocking directorates have in recent years mostly been studied as evidence of an elite community of managers rather than owners, and in order to explain similarities in political action taken by firms. This literature shows that the density of national networks has been declining since at least the 1990s (Chu and Davis, 2016; Mizruchi, 2013). Consequently, interlocking directorates still show effects in some domains but not in others (e.g. Banerjee and Murray, 2020; Benton, 2018). The role of large capitalist families within these networks, however, has not been studied systematically

recently enough to reflect the concentration of wealth that has taken place in recent decades (Mizruchi, 2004; Piketty, 2014). The next section develops some hypotheses regarding what we would expect to find based on these previous works.

3. Hypotheses

The research question asked here is whether capitalist families concentrate economic power by coordinating as a class in networks of interlocking directorates. For this to be true, members of the capitalist class should be more likely to sit on the same company boards, compared to managers or directors who do not own relevant shares of their companies, following the Marxist class model (esp. Zeitlin, 1989). There are different ways of defining and operationalizing the capitalist class. A rigorous development is given by Wright (1976, 1998), who differentiates between ownership and control over capital. On the one hand, there are classic capitalist owners who legally possess business assets. On the other, there are top managers who inhabit intermediate class positions by controlling capital but who own comparably little themselves. By this definition, the CEOs of the largest corporations and capitalist investors and owner-managers should be members of the same class. Following this, one would therefore expect the possession of high net wealth – in this stratum, and especially in the sample here, largely made up of business assets and therefore capital ownership – to make it more likely that an individual is connected to other members of this class. I therefore assume ownership *and* control over capital to be the determinants of position in the class structure. The exact mechanisms linking this structure to interaction cannot be identified, as a result of the data used; there could be several such mechanisms. To mention two, one could argue that ties represent social capital – on top of economic capital – within a class, such as membership of social clubs (Cousin and Chauvin, 2014; Domhoff, 1975; Mace, 1971) and other informal networks such as shared elite education (Bourdieu, 1996; Lillie, 2020), to which only other members of that class have access. A second potential mechanism might be (intentional) coordination and political unification to realize class interests based on shared economic interests (e.g. Banerjee and Murray, 2020 ; Murray, 2017 ; Young, Banerjee and Schwartz, 2018).

To confirm the research question, it is a necessary precondition that a relevant share of the population of super-rich family members are at minimum connected to the largest share of the network. Since possessing significant amounts of capital should qualify one as a member of this class, according to this "class cohesion model" we would expect that:

H1: Estimated net wealth increases the odds of being connected to networks of interlocking directorates.

This could take place either in person, when super-rich family members sit on multiple boards themselves, or when other members of super-rich firms' management do. However, Zeitlin's (1989; 1974) argument allows us to expect another finding and to answer the research question from a slightly different angle. Large families or family dynasties should be in control of a relevant share of the largest corporations. Zeitlin et al. (1974, p. 102) write:

> It is generally recognized that many legally distinct personal holdings, together with those held through personal and family holding companies, trusts, and estates (and/or such intermediaries as nominees and brokers), may form a single-family bloc for purposes of control. Aggregating such holdings (and penetrating their anonymity) is a primary task in any study of corporate control.

According to this family-control model, we would therefore assume that looking at the network of the largest global firms:

H2: Super-rich families hold controlling shares (through complex ownership relations) in large firms at the center of the interlocking directorate network.

For the first two hypotheses, managers and owners were assumed to belong to a unified capitalist class, following Wright (1976, 1998). Managerialism gives an alternative definition (cf. e.g. Mizruchi, 2004, pp. 581–592). From this perspective, managers are a distinct social group that in some situations have interests that compete with those of owners. For Hypothesis 3, I take the latter perspective. My purpose is not to take sides in this debate upfront. Instead, seeing managers as distinct, a corporate elite rather than the capitalist class, allows for a stricter test of the class cohesion model and therefore of the question studied here. No matter whether we expect managers to be in the same class as the owners of capital, it is a necessary precondition that the owners of capital are connected to each other. If this were not the case, it would not only call into question the cohesion of capitalists and managers, but – more significantly – even cohesion among capitalists, and therefore the class model more broadly.[3] The strongest confirmation of the class cohesion model would therefore be to show that members of different super-rich families are more likely to sit on the same corporate boards. If, controlling for other factors, the odds of both sitting on the same boards are greater for two members of the largest capitalist families than for two non

[3] This is true at least for coordination across different fractions of capital, and in 2020 as a single point in time.

rich managers, this would be powerful evidence of the class cohesion model.[4] Therefore, the final hypothesis is:

H3: Super-rich individuals are more likely than non rich managers and directors to have a common tie with members of other super-rich families who own separate sets of capital.

These hypotheses are tested using heterogenous evidence. The data, sample, and methods used are presented next.

4. Data and sample

The study is based on two data sources. The first is a large sample from the firm database ORBIS (Bureau van Dijk, 2020). In 2022, the full database consists of data – depending on the definition of what is a "firm" - for almost 500 m firms globally, with information on financials, directors, and management, including their names, as well as shareholders' shares and names. ORBIS gathers data from company reports, private data providers, and official registers. Since it is the only data source that combines information on management and ownership and it is the largest global firm database, it seems to be the ideal source for this study. This is because the involvement of individuals in corporations can be analyzed in both of the relevant channels: being in an executive position in a firm or having a relevant share in it. Unfortunately, the database is also very complex and some severe data-quality problems must be overcome to make it usable. This includes duplicates, entity ambiguity, and varying coverage which was shown to be acceptable mostly for the largest firms in the largest global economies (Bajgar et al., 2020, Compston, 2013; Garcia-Bernardo and Takes, 2018; Heemskerk et al., 2018; Kalemli-Ozcan et al., 2015). We accessed the data through automatized downloads from the web portal according to our contract with BvD. Many but not all known problems of ORBIS were tackled during data preparation. The most relevant step might be to only include the heads or holding companies of corporate groups and exclude all subsidiaries. This was operationalized by excluding all firms that are owned by another *firm* with a share of 50% or more.[5] Other steps included removing 21,207 duplicate firms – by means of the procedure based on clustering of similar network positions suggested by Garcia-Bernardo and Takes (2018) – as well as removing nonprofit corporations and other entities which are not firms. In terms of ownership data, share percentages were complemented (direct shares for missing total shares) and

[4] For simplicity reasons, "nonrich" here refers to individuals who could not be identified as belonging to one of the known super-rich families with an estimated family net worth of US$250 m or more. Many of them will, however, be in the top of the global income distribution or could be members of other relatively wealthy families, and therefore qualify for a different variety of affluence (Arndt, 2019).

[5] For a detailed discussion see Appendix 3c.

interpolated over the period 2010–2018. One problem that could not be solved is the known inaccuracy of the global ultimate owner variable which was used as provided by BvD. For a more detailed discussion of ORBIS and its shortcomings see Arndt (2023).

The sample was downloaded in early 2020. The final sample used for this study consists of 254,207 firms that are heads of corporate groups, defined via the global ultimate owner variable from ORBIS. These were the largest global firms in the ORBIS database at the beginning of 2020 in terms of revenue or number of employees. More precisely, about half of the sample consists of the largest firms in terms of number of employees, with at least 200 personnel each, and the other half the largest firms in terms of income, with at least US$ 50 m revenue. To identify the super-rich in the ORBIS data, a much larger set of 7,761,467 entities overall was downloaded. This is equal to the 254,207 firms including their shareholding entities, shareholders' shareholders, and subsidiaries. All data on firms, including those concerning their financials, ownership structure, and management, stems from this data source. In particular, data on individual directors stems from ORBIS, including information on age, gender, exact job position, and country.

The second data source is a list of names of super-rich families, individuals, and the firms that are the origin of their fortunes. The list also includes each family's likely country of residence, and an estimation of net wealth where available. It was gathered by and acquired from a private researcher, who aggregated different sources; in particular, the global rich lists that were available (Bornefeld, 2019; the top 10 sources are presented in Appendix 1a). Estimates of family net wealth stem from these sources. They are therefore estimates made by groups of journalists. While the professionalism of these estimates seems high, they still need to be taken with a grain of salt since they are based only on publicly available information such as that found in ORBIS (Fisher, 2018). It is difficult to assess the accuracy of these estimates. On the one hand, there is evidence that these approximations are higher than tax records show (Raub, Johnson, and Newcomb, 2010). On the other, there is evidence that a relevant share of dynastic wealth might be hidden (Cummins, 2022).

The list used for this study consists of overall 9,532 families globally. This list was automatically matched with the ORBIS data. Inclusion criteria in terms of year of the wealth estimate for the original list were kept relatively loose; to be included, the family had to have been listed as possessing at least $US250 m in about the last 10 years. Some individuals on the list might therefore have since lost their fortune or passed away. This should not be problematic, however, because the final sample only contains those individuals and families who are also in the 2020 ORBIS sample. Yet, it might be that a current wealth estimate would look different and might fall below the $US250 m threshold in some cases.

An eight-step rule-based string-matching procedure was implemented in Python to find as many super-rich families and individuals as possible in the complex ownership and company

structures (cf. a detailed description in Appendix 2). Table 1 presents the results of this procedure, and therefore the sample of super-rich families and individuals identified that forms the basis of this study, on the three levels of the data structure: individuals, families, and firms. It gives a rough idea of potential selection, i.e. how many super-rich families could not be identified.

One shortcoming of this matching procedure is that family names were used to snowball family members in the data. This means that if, and only if, a person with the same last name was also a member of the same board or senior management or was a relevant shareholder in the same company, this individual was added as a family member. If family members had different last names, they could not be identified as such. It could be that this selects against women, because last names are – at least in Western countries – passed on to male descendants in the lineage, as noted also by Cummins for the case of England (Cummins, 2022, p. 13), while women often change their last names when they marry. Nevertheless, customs vary by culture. In Japan, for example, it is common for sons-in-law to join family businesses and take up the founding families' names, so the nature of this bias will also vary (O'Hara, 2004). However, it is not quite clear whether this works in the same way in the top strata of societies. Ginalski, for example, presents many examples from the wealthiest Swiss families in which men had taken the name of their wives' more prestigious families (Ginalski, 2015, pp. 222–223). In sum, the matching procedure based on last names introduces bias. While further research on last-name customs among super-rich dynasties would be helpful, in the meantime at least two consequences of this bias are likely. First, later generations of wealth dynasties might have a lower chance of ending up in the sample due to having changed their last names. Second, in the former case but also in general, daughters- and sons-in-law might be systematically underrepresented in those cases where they have different last names.

It should be briefly mentioned that the main level of analysis is the family level, operationalized here as individuals with the same last name as a confirmed super-rich family, and who share ownership and/or control in the same firms. There are two reasons for this method. The first is that wealth research has shown that the management and reproduction of family fortunes largely takes place at the family level (cf. e.g. Beckert, 2022; Bessière and Gollac, 2019; James, 2006; Stamm, 2016). As Zeitlin (1974) argues, families may also be the relevant level of concentration of economic power. The second reason is that data restrictions prevent access to other types of information. Not only are rich lists often collected at the family level, but also much of the shareholder data in ORBIS only contains family shares, not differentiated by the individual.

Table 1: Findings and sample composition of super-rich families, individuals, and firms in ORBIS.

Selection	Cases	
	Absolute	**Percent**
Families		
Total number of families known from the original rich list	9,532	100
Identified in ORBIS with at least one shareholder and/or director/senior manager[6]	4,250	44
Identified in ORBIS with at least one director or senior manager	2,808	29
Identified in ORBIS with at least one director holding a position in senior management or on a supervisory board in 2020	2,313	24
	Absolute	
Individuals		
Total number of individuals known in the original rich list	11,418 individual members of the 9,532 known families	
Identified in ORBIS as a shareholder and/or director/senior manager	13,597 individual members of the 4,250 identified families	
Identified in ORBIS with at least one director or senior management position	5,975 (44%) of the 13,597 identified individuals	
Individuals that hold a position in senior management or on a supervisory board in 2020	4,250 (71%) of the identified managers	
	Absolute	
Firms		
Total number known in the original rich list	7,441	
Identified in ORBIS as having a super-rich shareholder or ultimate owner and/or director (including subsidiaries)	201,788[7]	
Identified in ORBIS as having at least one super-rich individual in management (including subsidiaries)	79,450	
Largest global firms where super-rich individuals hold positions in senior management or on supervisory boards in 2020 (excluding subsidiaries)	3,963	

[6] For Germany, we identified the missing super-rich names in ORBIS manually, using individual Google searches rather than firm names. It turns out that about 35% of the nonmatched names do exist in the ORBIS database, but not in our drawn sample. This means that they are not related to the world's largest companies, as defined by the thresholds for employees or revenue, their shareholders, shareholders' shareholders, or subsidiaries. Reasons for this could be manifold: e.g. that they do not hold the majority of their wealth in business assets but in real estate or other assets which we have no data on. It could also be that they have lost their fortunes, hold some of it in a form that is untraceable to them, or similar. Obviously, the firm selection thresholds are also only imperfect predictors of company or share value, which are, of course, the relevant measure for business wealth. We will investigate this further in future studies.

[7] In contrast to the sample of firms used for the main analysis, this number not only includes the largest 254,207 global corporate groups but also all their subsidiaries and shareholders, which adds up to about 3 m firms. The fact that subsidiaries are included here means that many of these firms and positions might be artificial; for example, they may include shell corporations or administrative positions in subsidiaries (Heemskerk and Takes, 2016, p. 98). The billionaire founder of the British optician chain Specsavers, Dame Mary Lesley Perkins, technically holds management positions in 861 distinct entities and branches within the Specsavers corporate group. These positions are not relevant for the questions under study here. Ties of interest are those between different corporate groups. Therefore, all subsidiaries are excluded and only directorial or senior management positions in global ultimate-owner companies are considered. For detailed discussion of this entity-ambiguity problem, see Heemskerk et al. (2018, pp. 14–19).

Figure 1 presents the geographic distribution of the findings.[8] Another relevant choice is that I include not only the members of the board of directors (BoD) but also of senior management, in line with the original work by Berle and Means (1991) and more recent applications of their approach (Heemskerk and Takes, 2016). This method prioritizes finding as many super-rich individuals as possible and defines ties more loosely, while placing less emphasis on accurately differentiating the distinct roles of the board and senior management. The method is also less vulnerable to possible variation in the definitions of supervisory and executive positions by country (cf. Scott, 1985, pp. 17–18). "Senior management" refers to all available executive positions in a firm such as Chief Executive Officer, Chief Operations Officer, and so on. "Supervisory management" includes supervisory positions such as those on the BoD or other supervisory boards. Besides the exact name of a manager's position, ORBIS also includes the classifications "Executive" and "Supervisory," and these identifiers are used to distinguish between these two types in the regression analysis.

The question of whether a firm is owned by a family was determined in the following way. Identified shareholding entities that belong to a family (individuals, family shares, or foundations/trusts) are aggregated at the family level. As well as direct shareholdings, those held by firms controlled by the family (with a share of 50% or more for these aggregation purposes) were also considered, up to the third network degree; i.e. up to the level of firms owned by firms owned by firms owned by the family. This was done imitating the procedure suggested in Rungi, Morrison, and Pammolli (2018). All firms controlled directly or indirectly by a family, as ultimate owners or intermediate shareholdings, were coded as such. For the purposes of the regression analysis below, a controlling share is defined as owning 20% or more of a company.

[8] Appendix 1b shows the geographic distribution of the original list.

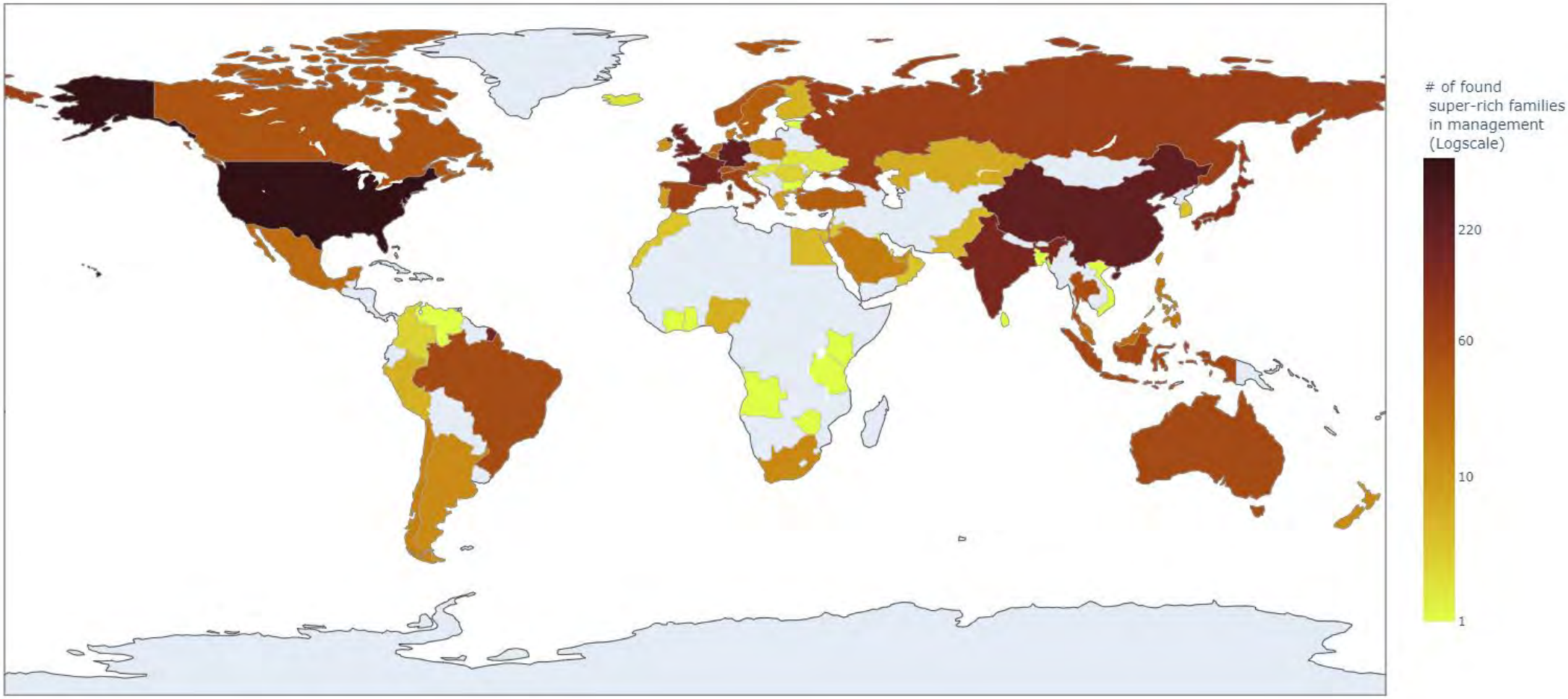


Figure 1: Geographic distribution of the analyzed sample of identified super-rich families in senior management or on the supervisory board of a company.

Further data-preparation steps included imputation of missing values for the firm founding year, board size, revenue, number of employees, and individual's age. Family size was also imputed for those cases where there were only family shares in ORBIS that were not differentiated at the individual level. These missing values were imputed using all other variables in the analysis with random forests in Python (Wilson, 2021; cf. Appendix 3b). Five datasets were imputed and were used to adjust standard errors for multiple imputation in the analysis. If a director's country was missing, it was replaced by the country where the largest number of companies they were involved in were located. Data on the individual, firm, and family levels was used for the analyses. The methods of analysis are presented in the next section. Descriptive statistics of the variables used, including missingness, are reported in Appendix 3a.

5. Methods

I use two different methods to analyze the international network of interlocking directorates, and the role of the super-rich within it, as well as test the developed hypotheses. These are, first, network visualization, and second, logistic regression analysis. Graph visualization is conducted with the ForceAtlas2 algorithm in Gephi, which "simulates a physical system in order to spatialize a network. Nodes repulse each other like charged particles, while edges attract their nodes, like springs. These forces create a movement that converges to a balanced state" (Jacomy et al., 2014, p. 2). This method is used to visualize membership of super-rich family members in the largest component of the networks, as well as the centrality of their positions, and whether they are embedded in national or transnational clusters.

It is relevant to briefly explain the two different forms of networks used. One way to analyze interlocking directorates is to look at two firms and how the individuals sitting in the management of both entities work as ties between them. Especially when looking at the centers or clusters of such a graph, one can identify firms that are very well connected to many other firms. No matter whether interlocking directorates are understood as a mechanism of class cohesion, or simply as a channel for information flows that allow businesses to be run more efficiently, firms at the center of this network and the individuals who sit on their boards should have the easiest access to all other firms in the network. The position of a firm or individual in this network therefore tells us something about their structural position in relation to those of the other large global firms that are connected to the network. One relevant feature of this network is that it does not require a member of a super-rich family itself to be central; it is enough when other members of the board or senior managers at the same company are. This type of network is used to test the first two hypotheses, which make claims that can be answered by looking at the firm-level network.

A second network that can be constructed is the network of individuals where firms are ties. Individuals at the center of these networks are exposed to information from many different companies. Understood from the perspective of class cohesion, for example, they know about the concerns, obstacles, and interests of many other members of the network. One could therefore expect these individuals to have a central position in relation to all other individuals involved in these firms and the network. This type of network is used to test the third hypothesis, which claims that, controlling for other factors, two distinct super-rich capitalist families should be members of the same board.

In addition to visualizing these networks, I apply logistic regression to board positions in order to estimate the correlates of membership of the largest connected component of the global interlocking directorate network in the firm network. The purpose of this analysis is to understand what determines being tied to the network of firms, and therefore to test hypotheses 1 and 2. More precisely, being connected to the largest component of the network is used as a dependent variable. Board appointments are the unit of observation; i.e. if a director sits on three company boards, there are three observations for them. To account for this, standard errors are clustered at the firm level in this analysis. This first step of the regression analysis is applied to three different samples: all directors (rich and nonrich), only nonrich directors, and only members of super-rich families.

In addition to this, I estimate odds ratios using logistic regression based on dyadic data from the network of individuals to test Hypothesis 3. All individuals from the largest connected component of the network of individuals (N = 75,639) are taken as a basis. All possible dyads between these individuals are used as observations (N = 2,860,591,341 dyads) in a first sample, and all dyads only between the 3,404 super-rich family members in the largest component are used in a second (N = 5,791,906 dyads). The dependent variable is whether there is a tie between each pair of individuals; i.e. whether they sit on the same board.

Table 2 The independent variables and operationalization methods used and the hypotheses tested.

Independent variable	Operationalization	Hypothesis tested
Is the individual member of super-rich family?	Binary variable for each director indicating whether they are a member of one of the super-rich families or not.	Hypothesis 1
(Log) Estimated family net wealth	Logged family wealth estimate from the collaborated rich list.	Hypothesis 1
(Log) Family size	Logged count of distinct family members who are involved in any firm as shareholders or managers/directors.	Hypothesis 1
Ownership of own family in firm	Categorical variable indicating whether the director's family's members, directly or indirectly, collectively own 20% or more of the firm (controlling share), less than 20% (minority share), or no share at all.	Hypothesis 2
Type of involvement and position	Categorical variable indicating whether a director's family's members hold a controlling share of a firm and/or whether their management position is executive (e.g. CEO) or supervisory (e.g. BoD).	Hypothesis 2
Both super-rich (but from different dynasties)?	Indicating whether both directors in this dyad are members of super-rich families, but from separate families related to different fortunes (i.e. their wealth does not stem from the same original firm).	Hypothesis 3

Table 2 presents the independent variables used to test the three different hypotheses, and descriptive statistics for all the variables used are presented in appendices 3a and 4b. Due to the nature of the dyadic data and its collinearity, standard errors should ideally be clustered with specific procedures of multi-way clustering (Cameron and Miller, 2015). However, due to the large size of the data this is not feasible in terms of working memory and computation. Therefore, single-way clustering at the dyad level is used. This should not be problematic for interpretation of the results presented here, as shown in the results section. The analysis is implemented using weighted logistic regression. All regression results are adjusted for multiple imputation. Finally, it is necessary to mention that neither the full population of the super-rich nor that of global firms could be used, which might lead to missing links in the presented networks. This should be kept in mind when evaluating the results, most notably for the k-core and dyadic analyses. Results for the two networks and different methods are presented in the following.

6. Results

The first hypothesis claimed that estimated net wealth should increase the odds of being connected to networks of interlocking directorates. As a first test of whether the super-rich are in fact connected to the largest component of the firm network, I determine and visualize the network of interlocking directorates that link the 254,207 largest global firms.[9] In total, 1,473,149 directors are involved in these companies, of which 94,519 (6%) sit on the board of more than one company. There are more than 10,000 smaller components, consisting of only a few firms, and which are unconnected to the larger network. The most relevant part of the network is its largest component, which includes 46,045 (or 18% of all) firms, and 762,979 directors (51%).

Table 3: Measures of the super-rich in the largest connected component (CC) of firms

Measure	
Number of super-rich families in largest CC	1,772
Percentage of all known super-rich families in largest CC (N = 9,518)	18
Percentage of super-rich families found in ORBIS with at least one director in largest CC (N = 2,313)	76
Percentage of the estimated net wealth of all known super-rich in largest component (N = 9,518)	26
Percentage of estimated net wealth in largest CC relative to that of all super-rich found in ORBIS (N = 2,313)	86

The largest component of the network between firms in 2020 is illustrated in Figure 2. Several interesting observations can be made. Turquoise edges denote interlocks across national borders (18%), whereas brown edges denote national ties (82%). Some families and their countries of origin are highlighted as examples. Table 3 shows the participation of super-rich families in the largest component of the network. Of all the super-rich families from the original rich list used, 18% were identified in this study and are in some way involved in a firm that is part of the largest component. At this point, this includes any shareholdings or management positions. Of those families that were identified, 76% can be found in the largest component. This is already an interesting finding per se in that those firms with super-rich involvement form part of the largest component; this is the first evidence we have found in favor of Hypothesis 1. The counterfactual

[9] Appendix 4a presents a comparison of the results with those of Heemskerk and Takes (2016), who used the same data source and a comparable sample in 2013.

extreme case would have been that the super-rich remained isolated in their private firms. Even more interesting is that it seems especially the richest super-rich families are more likely to be represented in this network, with 86% of net wealth estimates from rich lists of the identified families being represented, while 26% of the estimated wealth (vs. 18% of families) of all families we know of are part of the largest component.

To further test the first two hypotheses, Table 4 presents odds ratios estimated from logistic regression for two samples: first, for all directors, super-rich or not, and second, only for super-rich family members. The binary dependent variable is membership in the largest component, i.e. being connected to a large number of the other largest firms globally (1) or not (0). Hypothesis H1 suggests that super-rich individuals should be more likely to be members of the largest component compared to nonrich directors and managers. In fact, the chances of being tied to this component are increased x 3.6 for those who are members of a super-rich family compared to a nonrich average director and controlling for other factors. In model 3, different positions of rich and nonrich directors are compared. The results provide further evidence of the relevance of being super-rich. All those in rich categories are more likely to be part of the largest component than nonrich BoD members. To a striking extent, super-rich members of the BoDs of companies that they do not control have by far the highest odds of being connected to the largest component.

Further evidence in this regard is provided by the results from the sample of only super-rich family members. Higher logged estimated family wealth increases the odds of being connected to the largest component of the network of firms. This can be interpreted as evidence that the wealthier a family is, the more likely it is to be connected to the interlock network between the largest firms globally. Logged family size shows a similar effect. This factor can be seen as a proxy of how much of a dynasty a family is. While there are some single entrepreneurs, the largest families are often those from old money, whose families have been involved with a company and significant financial means for generations.[10] The greater the number of members of such a dynasty being involved in any firm, the more likely it is that the family will be connected to the largest component. Being male is related to significantly higher odds of being part of the largest component. While this effect seems to be less pronounced for the super-rich, it is also significant when controlling for family characteristics. This speaks for the relevance of male networks for interlocked board positions, a relevance which also seems to apply for networks of super-rich men (cf. Young et al., 2021).

[10] However, it is a somewhat top-coded measure because the largest dynasties often only have one family share for the majority of its members, besides several individuals mentioned in the shareholder data.

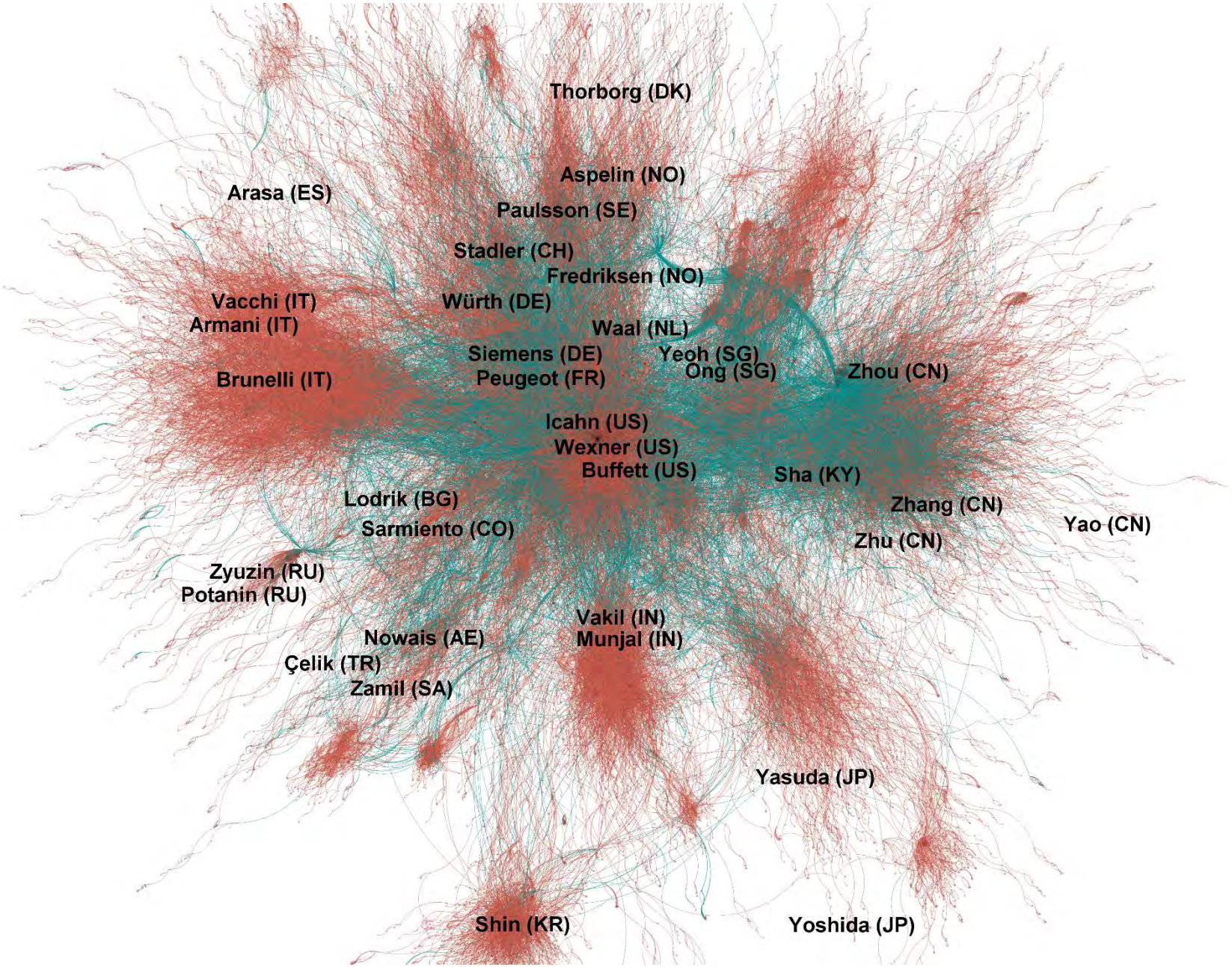


Figure 2: Largest component of the global network of interlocking directorates between firms. Nodes: firms; ties: individuals. Transnational ties = turquoise, national ties = brown. N = 46,045 firms connected by 176,556 ties. Some of the super-rich families involved are labeled as examples, with their home countries indicated in brackets.

	Firm is part of largest component?		
	All Directors	All Directors	Super-Rich Family Members
	(1)	(2)	(3)
Family level			
Director is a member of a super-rich family?	3.683***		
Ownership of own family in firm (ref=controlling share)			
Minority share			1.480***
No ownership			2.109***
Log (Estimated net wealth)			1.494***
Log (Family size)			1.231***
Individual level			
Male (ref = Female)	1.093***	1.091***	1.302*
Age	1.007***	1.007***	.996
Supervisory position (ref = Executive)	1.461***		1.702***
Family level x Individual level			
Type of involvement and position (ref= Nonrich on BoD of managerial firm)			
Nonrich exec. of managerial firm		.683***	
Nonrich BoD of super-rich firm		2.001***	
Nonrich exec. of super-rich firm		1.280***	
Rich owner-manager		2.527***	
Rich member of BoD without controlling share		8.702***	
Rich exec. without controlling share		1.733**	
Rich owner on own BoD		2.199***	
Firm level			
Log (Revenue)	1.108***	1.106***	1.051*
Log (Employees)	1.052***	1.049***	.880***
Log (Board size)	3.104***	3.085***	2.663***
Founding year	1.002***	1.002***	.998
Firm list status (ref=Listed)			
Unlisted	.221***	.222***	.399***
Constant	0.000***	0.000***	3.427
Observations	1,608,244	1,608,244	6,927

Table 4: Odds ratios from logit regression models. Standard errors are clustered at the firm level and adjusted for multiple imputation. All models also include control variables for individuals' home continent, continent of firm headquarters, and NACE sector division. For full regression table see Appendix 4b. *P < 0.05, **P < 0.01, ***P < 0.00.

To sum up, results from these first models give support for hypothesis H1: that the super-rich are more likely to be part of the network of interlocking directorates that enable them to exchange information with other families and other members of the capitalist class. Not only does being a member of a super-rich family increase the odds of forming part of the largest component, but being wealthier also does so, as does being a member of a larger family. When looking at the core of the network, there is additional evidence that this is not because these individuals or families own relevant shares in the firms that are at the core of the network. To further test hypothesis 2, Figure 3 presents the share of super-rich involvement in terms of ownership in relation to closeness to the core of the largest component..

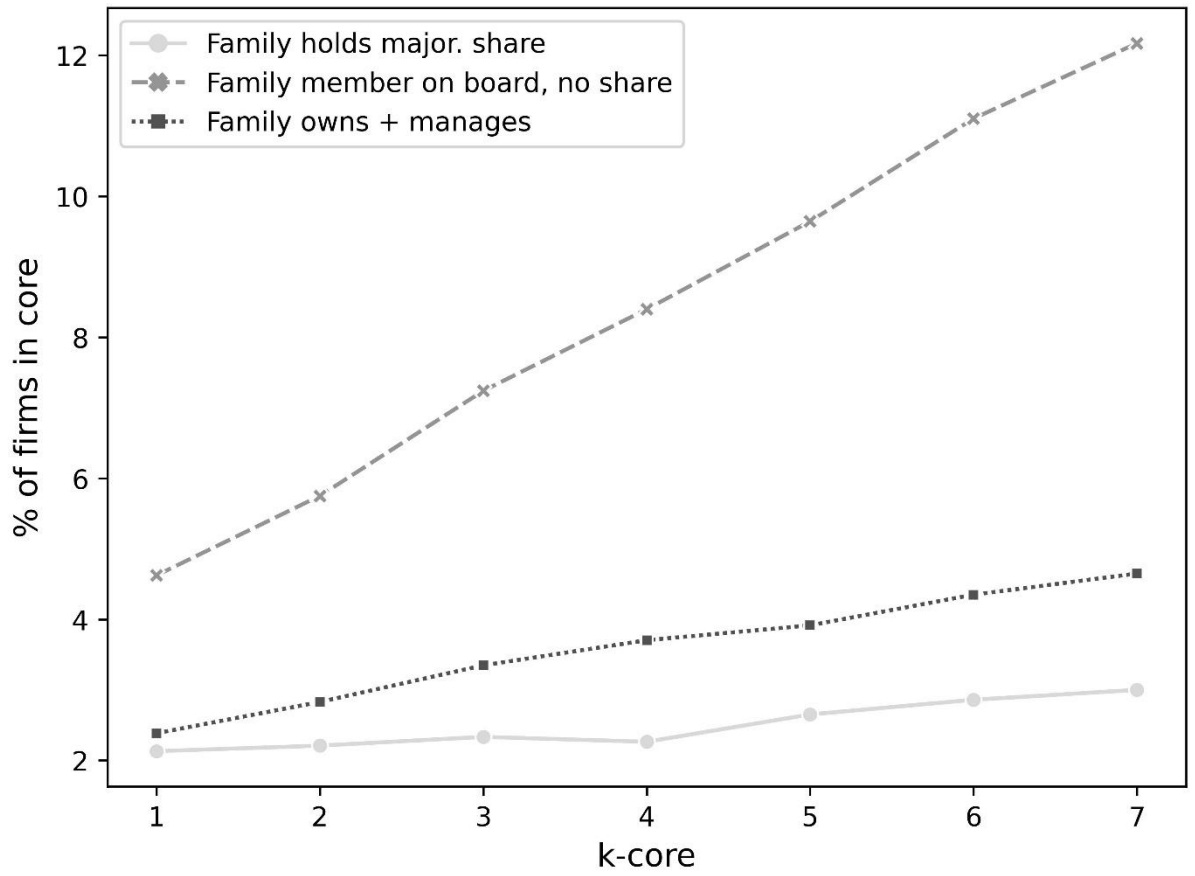


Figure 3: Types of super-rich involvement in firms in different k-cores of the largest global component of the network of firms.

The k-core means to keep the largest subgraphs of firms in the network who have at least k ties. A k-core of 2 means that we only consider the largest subgraph of firms with connections to at least two other firms. Therefore the higher the core number, the lower the number of firms in the core. As Figure 2 shows, the deeper one investigates the core of the largest component, the higher the share of firms for which a super-rich family member only sits on the board (and does

not own a share). At the same time, the share of firms for which super-rich family members have a controlling share remains relatively low and increases less – although it does, in fact, increase.[11] What this suggests is in line with the regression results from all the models presented above: Being in the largest component of the network is strongly correlated with sitting on the boards of large, listed firms, probably in a supervisory position as an external director. It is not the firms controlled by super-rich families that are at the core of the network. This is strong evidence against Hypothesis 2, and therefore also against the family-control model proposed by Zeitlin (1989). That said, another striking takeaway from Figure 2 is that, in the data used, most firms in the largest component do not have any form of super-rich involvement; these represent a majority of 80% and greater, depending on the core number.

Summing up, super-rich family members and especially the richest and largest families are the most likely to be connected to many other of the largest firms. However, they do not seem to be so through their own firms, but rather through membership of the supervisory boards of other large firms, particularly listed firms. This does not rule out some sort of coordination that is attained by sitting on these boards (cf. Murray, 2017, pp. 1622–1628) or a tendency for the super-rich to be connected to other super-rich families. However, it speaks against the idea that the firms owned by the super-rich are the most highly connected.

[11] One could argue that this is due to the international network and a lack of transnational clustering, so the core breaks into the national clusters, or possibly regional clusters as claimed by Heemskerk and Takes (2016). However, the finding is the same at the national level for Germany, the USA, India, and China, and also for Chile, the case Zeitlin, Ewen, and Ratcliff (1974) analyzed in the 1970s. While levels and differences between shares vary, super-rich BoD members on the boards of firms they do not own are the most prevalent in all analyzed national cores (cf. Appendix 4c).

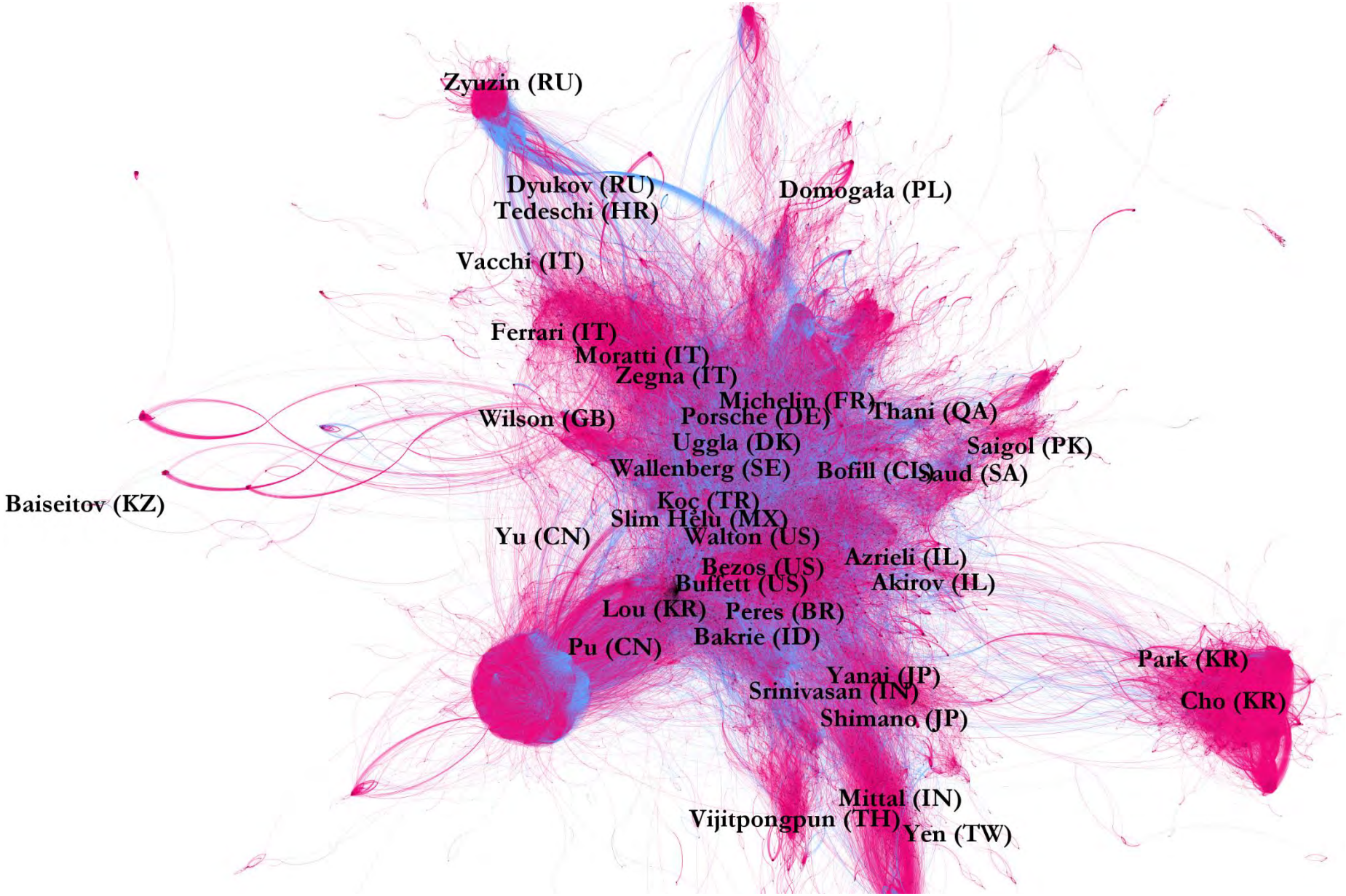


Figure 4: Largest component of the global interlocking directorates networks between individuals only. Nodes: individuals; ties: firms. Transnational ties = blue; national ties = pink. N = 75,639 individuals with 1,367,168 ties (sitting on the same board) between them. Some super-rich families are labeled as examples, with their home countries indicated in parentheses.

Next, we consider the network of individuals in order to better understand which individual appointments and characteristics foster the existence of a link between two individuals, that is, their being involved in the senior management or the BoD of the same company. Figure 4 presents a visualization of the network's largest component. Like firms, individuals seem highly clustered at the national level. National ties are shown as pink here (76%), and transnational ties as blue (24%). Hypothesis 3 suggests super-rich family members should be more likely to sit on the same board as members of other super-rich families than with nonrich managers. This is a relatively strict test of the idea of coordination between large capitalist families. The main independent variable is also operationalized strictly here: It is set to 1 only if two members of dynasties stemming from two different fortunes are involved in management at the same firm. This means the firm their wealth came from must be a different one. To make it more concrete: If Charles Koch and David Koch (or Susanne Klatten and Stefan Quandt) sat on the same board this would not be counted, but if Warren Buffett and Bill Gates did (or Delphine Arnault and Piero Ferrari), it would. This is necessary, because otherwise there would be too many cases in which the descendants of founding families of the same firm sit on the boards of family holding companies or similar. Such cooccurrences are not relevant for the class explanation, which suggests instead that two families are more likely to have a tie as a result of their both owning and controlling some kind of capital, rather than because they manage the same fortune.

Table 5: Odds ratios estimated from weighted logistic regression on dyadic data.

	Dependent variable: Both on the same board? (=1)	
	All directors in largest component	Super-rich family members in largest component
	(1)	(2)
Family level		
Both super-rich (but from different dynasties)?	.358	
Difference in log estimated family wealth		.315**
Difference in log family size		.288**
Individual level		
Both male?	1.179	.720
Both female?	1.050	1.489
Firm level		
Same sector?	64.253***	69.055***
Same country?	39.130**	170.701***
Same firm size?	4.944	1.593
Distinct observations[12]	95	1,673
Directors	75,635	4,324
Dyads/ Sum of weights	2,860,288,795	5,700,376
Intercept	0.000***	0.000***

Note: Standard errors are clustered at the dyad level and adjusted for multiple imputation. Results remain without control variables (Appendix 3e) and are therefore not driven by potential overcontrol bias.

Table 5 presents the results from logistic regression on all possible dyads: between all directors, first, and between only super-rich family directors, second. Surprisingly, dyads of members of two different families defined in this way have lower odds of sitting on the same board. More importantly, the coefficient is not significantly different from zero. There is therefore no evidence that members of two different super-rich families are more likely to sit on the same board

[12] The regressions are implemented as weighted logistic regressions. In the first model, there are six binary independent variables and one binary dependent variable. Both being male and both being female are mutually exclusive. This comes down to only 95 observed different combinations of these variables which are frequency weighted summing up to the 2.8 billion possible (undirected) dyads between the 75,635 directors. The unique observations made for the second sample are 1,673 because they include differences between continuous variables and therefore more possible values than combinations of only binary variables.

simply based on this fact. Among the super-rich, differences in wealth levels and family size also do show significant effects. This means that *if* the individuals are connected to members of other families, then they are more likely to be connected to other families with similar wealth levels and similar family sizes – the proxy used for age of wealth here. This speaks for some degree of homophily of an upper class, but this has to be seen in light of the finding that two different families are not more likely to sit on the same board. By far the most relevant drivers in these models are at the firm level, namely that the two firms are in the same country or sector. This is strong evidence against H3 and the idea that the super-rich coordinate through interlocking directorates simply based on their wealth level.[13] As mentioned in the methods section, standard errors are likely underestimated because only single-way clustering is accounted for at the dyad level. Since the relevant independent variable is already insignificant and likely points, if is different from zero at all, in the wrong direction to support H3, this should not affect the results presented here.

If interlocks of super-rich family members do not seem to be a case of class coordination, how else can we possibly understand their position in interlock networks? Some relevant observations arise throughout all the models presented above. Holding a supervisory position increases the odds of an individual's being part of the largest component of firms. In particular, this implies that being a supervisory director rather than an executive makes it more likely that an individual will be in the largest component, as well as being affiliated to listed firms. This emphasizes the role of external board members as participants in the network as more significant than that of the owner-managers of unlisted, fully owned private firms, such as classic capitalists. This finding suggests that large, listed companies invite the super-rich to act as external directors.

[13] One could argue that the super-rich do not sit on the same boards themselves but are connected through other members of their boards. However, this would imply that their firms are more likely to be connected. I present evidence against this alternative mechanism in Appendix 4d.

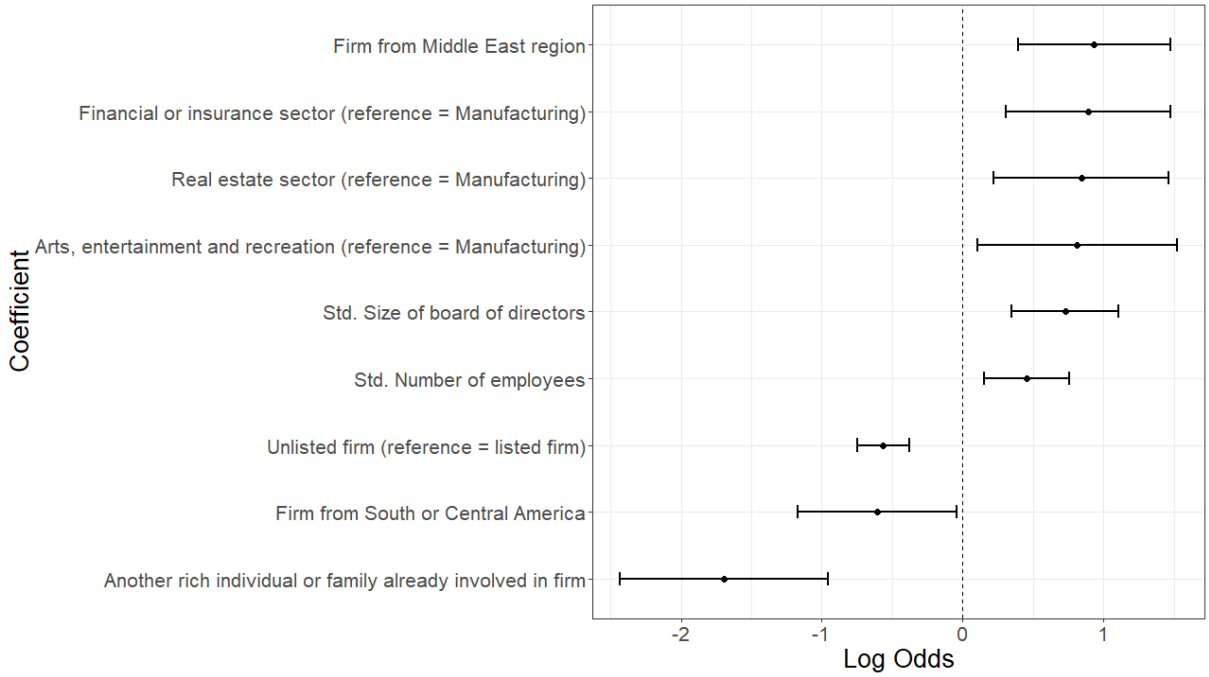


Figure 5: Log odds from a logit model estimating presence of a super-rich family member on a firm's supervisory board. Data on all 254,207 firms is used for observations. Only coefficients significant at the 95% confidence level or higher are presented. Board size and employees are standardized with mean 0 and SD 1. Appendix 3f presents results without the last variable, as a test of endogeneity.

To provide directions for future studies on this topic, Figure 5 presents a final piece of evidence which helps us to understand these relationships better. The figure illustrates log odds for all significant coefficients from a logit model, performed on the full sample of 254,207 firms. The outcome variable is whether a firm has a member of a super-rich family on their BoD. The predictor variables are all variables on the firm level used in previous models. In particular, three impressions can be gained from this. First, in line with results from previous models, it seems to be relatively large (in terms of employees), listed firms that recruit members of super-rich families to their supervisory boards. This is interesting because it indicates that members of super-rich families are actively hired by management and/or shareholders of public firms, rather than having any direct ownership interests in these firms. A possible reason for this could be the firms' desire for prestige among their board members and for the effects of such prestige (D'Aveni, 1990; D'Aveni and Kesner, 1993). Second, the strongest predictor for the absence of a member of a super-rich family on a firm's supervisory board seems to be that there is already another form of involvement on the part of such a family in that firm – be it in executive management, through ownership, or both. To hire one member of such a family seems to be enough. Third, region and sector both seem to play a role and further research is necessary to explain variation within these

categories. Overrepresentation of members of super-rich families on the boards of Middle Eastern firms might be due to a high level of involvement on the part of royal families in state-owned firms which they control de facto, but this needs to be investigated further. The super-rich seem to be overrepresented on the boards of financial, real estate, and art and entertainment firms such as private museums. In all these cases, motivating factors for both the firms and the super-rich individuals themselves could be hypothesized. From the firms' perspective, financial and real estate firms in particular could benefit from ties to super-rich families and their contacts as potential clients. From the perspective of wealthy individuals and families, arts institutions and museums, especially in the USA, are a popular and prestigious meeting point for the richest families and it might therefore be seen as attractive and desirable for super-rich families to join these clubs. However, none of these potential alternative explanations of super-rich involvement in interlock networks have been tested here in detail, and they can only be viewed as starting points for future studies on the topic.

7. Conclusion

This study mapped 4,250 members of super-rich families, with an estimated net worth of US$250 m, in networks of interlocking directorates. It is the first study to analyze interlocking directorates of the super-rich on such a large scale. Classic theories that engaged with families controlling capital – the capitalist class – saw interlocking directorates as potential measures for class cohesion. Shared board membership was seen as either an indication of family control of some of the largest and most central corporations, or as interaction between the wealthiest families and between those families and other members of the capitalist class. The results presented here show two things. First, super-rich families are more likely than other individuals to be part of the largest component of global interlocking directorate networks. The share of firms with all forms of super-rich involvement also increases at the core of the network. Second, however, this does not take the form of the involvement of the super-rich in their own companies, but through their sitting on the boards of large, listed companies. Also, they are not more likely than other individuals to sit on the same boards with members of other super-rich families.

The widespread idea of the fracturing of interlock networks and how it affects the super-rich within them could not be tested here due to the nature of the cross-sectional sample. However, this was not necessary to support the argument being made here. The presented evidence speaks against the relevance of ownership of capital for class coordination in general in these networks – regardless of whether fracturing makes these networks less relevant or not. These two facts could be interacting in several ways and capital ownership might not matter because of the ongoing fracturing. However, the results presented still stand, independent of ongoing fracturing for 2020.

Several shortcomings of the study need to be disclosed. The identification of super-rich families is imperfect, with only about 45% of the global list being identified. There are several possible reasons for this, including the data quality of the ORBIS data or the data in the original rich list, or the way in which the firm sample was defined. One problem is that firm names change regularly, and the super-rich tend to found new companies or whole company hierarchies to manage their wealth and presumably tax burdens and legal requirements. This makes it difficult to verify an individual only by their name in an automated manner, when it is only related to firms that have completely unknown names. Furthermore, the search for super-rich individuals heavily relies on using their last names, which excludes those individuals and family strains that have since taken different surnames. Finally, the ORBIS database's varying coverage by country and sometimes federal state needs to be kept in mind when evaluating results (Liu, 2020, pp. 344–346). In general, higher-income countries seem to be better covered, with the exception of the USA, for which fewer firms are included (cf. Garcia-Bernardo and Takes, 2018).

The results presented here reject the idea of coordination of a capitalist class simply based on the commonality of possessing or having control over large amounts of capital. There is therefore no identifiable "peak of capital," or level at which the richest families connect with each other. It must be stressed that this was only tested here for the channel of interlocking directorates. There are also suggestions that coordination might take place through other channels such as other networks and boards, for example those of nonprofits, cultural organizations, thinktanks, foundations, business associations, or policy-planning networks (Barnes, 2015 ; Comet, 2019 ; Domhoff, 1967 ; Kinderman, 2017 ; Skocpol and Hertel-Fernandez, 2016 ; Useem, 2015). However, the results presented here seem to suggest that super-rich family members tend to be hired by large, listed firms. While possible reasons for this should be investigated in future studies, it does seem at least that these appointments are unrelated to the ownership interests of the super-rich – and therefore that they serve some purpose for firms and their management.

References

ALLEN M.P., BROYLES P., 1989, "Class Hegemony and Political Finance: Presidential Campaign Contributions of Wealthy Capitalist Families," *American Sociological Review*, *54*, 2, pp. 275–287.

ARNDT L., 2019, "Varieties of Affluence. How political attitudes of the rich are shaped by income or wealth.," *European Sociological Review*, *36*, 1, pp. 136–158.

ARNDT L., 2023, "When should we believe research using ORBIS firm data?," *MPIfG and Sciences Po, Cologne and Paris*.

BAJGAR M., BERLINGIERI G., CALLIGARIS S., CRISCUOLO C., TIMMIS J., 2020, "Coverage and representativeness of Orbis data", OECD Science, Technology and Industry Working Papers, n° 2020/06, OECD, Paris

BANERJEE T., MURRAY J., 2020, "Class Dominance or Fracturing? Sources of Broad Interest in Lobbying by Fortune 500 Corporations," *Sociological Perspectives, 62,* 2, pp. 157–175.

BARNES R.C., 2015, "Structural Redundancy and Multiplicity within Networks of US Corporate Directors:," *Critical Sociology*, *43*, 1, pp. 37–57.

BECKERT J., 2022, "Durable Wealth: Institutions, Mechanisms, and Practices of Wealth Perpetuation," *Annual Review of Sociology*, *48*, 1, pp. 233-255.

BENTON R.A., 2018, "Brokerage and Closure in Corporate Control: Shifting Sources of Power for a Fractured Corporate Board Network:," *Organization Studies*, *40*, 11, pp. 1631–1656.

BERLE A.A., MEANS G., 1991, *The Modern Corporation and Private Property*, New York, Routledge.

BESSIERE C., GOLLAC S., 2019, *Le genre du capital. Comment la famille reproduit les inégalités*, Paris, La Découverte.

BEYER J., HÖPNER M., 2003, "The disintegration of organised capitalism: German corporate governance in the 1990s," *http://dx.doi.org/10.1080/0140238031233128073*8, *26*, 4, pp. 179–198.

BOND M., GLOUHAROVA S., HARRIGAN N., 2010, "The political mobilization of corporate directors: Socio-economic correlates of affiliation to European pressure groups," *British Journal of Sociology*, *61*, 2, pp. 306–335.

BORNEFELD A., 2019, "Vermögensforschung," *https://www.reichsten.de/en/index*.

BOURDIEU P., 1996, *State Nobility. Elite Schools in the Field of Power*, Stanford, Stanford University Press.

BROYLES P.A., 1993, "The Corporate Elite Goes To Washington: Presidential Campaign Contributions of Corporate Officers," *Social Science Research*, *22*, pp. 72–91.

BUREAU VAN DIJK, 2020, "Orbis. Powering the business of certainty." https://www.bvdinfo.com/-/media/brochure-library/orbis.pdf

BURRIS V., 2001, "The Two Faces of Capital: Corporations and Individual Capitalists as Political Actors," *American Sociological Review*, *66*, 3, p. 361-381.

BURRIS V., 2005, "Interlocking Directorates and Political Cohesion among Corporate Elites," *American Journal of Sociology*, *111*, 1, pp. 249–283.

CAMERON A.C., MILLER D.L., 2015, "Robust Inference for Dyadic Data," *Winter North American Meetings of the Econometric Society*. http://cameron.econ.ucdavis.edu/research/dyadic_cameron_miller_december2014_with_tables.pdf

CHU J.S.G., DAVIS G.F., 2016, "Who Killed the Inner Circle? The End of the Era of the Corporate Interlock Network," *American Journal of Sociology*, *122*, 3, pp. 714–754.

COMET C., 2019, "How does the inner circle shape the policy-planning network in France?," *Socio-Economic Review*, *17*, 4, pp. 1021–1041.

COMPSTON H., 2013, "The Network of Global Corporate Control: Implications for Public Policy," *Business and Politics*, *15*, 3, pp. 357–379.

COUSIN B., CHAUVIN S., 2014, "Globalizing forms of elite sociability: varieties of cosmopolitanism in Paris social clubs," *Ethnic and Racial Studies*, *37*, 12, pp. 2209–2225.

CREDIT SUISSE, 2021, "Global wealth report 2021". https://www.credit-suisse.com/media/assets/corporate/docs/about-us/research/publications/global-wealth-report-2021-en.pdf

CUMMINS N., 2022, "The hidden wealth of English dynasties, 1892–2016," *The Economic History Review*.

CURTIS G., 2020, *Advising the Ultra-Wealthy. A Guide for Practitioners*, Chur, Springer Nature Switzerland.

D'AVENI R.A., 1990, "Top Managerial Prestige and Organizational Bankruptcy on JSTOR," *Organization Science*, *1*, 2, pp. 121–142.

D'AVENI R.A., KESNER I.F., 1993, "Top Managerial Prestige, Power and Tender Offer Response: A Study of Elite Social Networks and Target Firm Cooperation during Takeovers on JSTOR," *Organization Science*, *4*, 2, pp. 123–151.

DAVID T., WESTERHUIS G., 2014, *The power of corporate networks : a comparative and historical perspective*, London, Routledge.

DOMHOFF G.W., 1967, *Who Rules America?*, Englewood Cliffs N.J., Prentice Hall.

DOMHOFF G.W., 1975, *The Bohemian Grove and Other Retreats. A Study in Ruling-Class Cohesiveness*, New York, Hagerstown, San Francisco, London, Harper & Row.

FISHER F., 2018, "Forbes 400 ranks America's wealthiest people using data from Orbis," *Bureau van Dijk*. https://www.bvdinfo.com/en-us/blog/corporate-finance-and-ma/orbis-data-in-forbes-400-2018

FORBES, 2021, "World's Billionaires List. The richest in 2021." https://www.forbes.com/worlds-billionaires/

FRANCOIS P., LEMERCIER C., 2014, "Ebbs and Flows of French Capitalism," dans DAVID T., WESTERHUIS G. (eds.), *The Power of Corporate Networks. A Comparative and Historical Perspective*, New York and London, Routledge, pp. 149–168.

GARCIA-BERNARDO J., TAKES F.W., 2018, "The effects of data quality on the analysis of corporate board interlock networks," *Information Systems*, *78*, pp. 164–172.

GINALSKI S., 2015, *Du capitalisme familial au capitalisme financier? Le cas de l'industrie suisse des machines, de l'électrotechnique et de la métallurgie au XXe siècle*, Neuchâtel, Éd. Alphil.

GOERRES A., HÖPNER M., 2014, "Polarizers or landscape groomers? An empirical analysis of party donations by the 100 largest German companies in 1984–2005," *Socio-Economic Review*, *12*, pp. 517–544.

HEEMSKERK E.M., 2013, "The rise of the European corporate elite: Evidence from the network of interlocking directorates in 2005 and 2010," *Economy and Society*, *42*, 1, pp. 74–101.

HEEMSKERK E.M., TAKES F.W., 2016, "The Corporate Elite Community Structure of Global Capitalism," *New Political Economy*, *21*, 1, pp. 90–118.

HEEMSKERK E.M., YOUNG K., TAKES F.W., CRONIN B., GARCIA-BERNARDO J., HENRIKSEN L.F., WINECOFF W.K., POPOV V., LAURIN-LAMOTHE A., 2018, "The promise and perils of using big data in the study of corporate networks: problems, diagnostics and fixes," *Global Networks*, *18*, 1, pp. 3–32.

HEERWIG J.A., MURRAY J., 2019, "The political strategies and unity of the American corporate inner circle: Evidence from political donations, 1982–2000," *Social Problems*, *66*, 4, pp. 580–608.

HILFERDING R., 1955, *Das Finanzkapital. Eine Studie über die jüngste Entwicklung des Kapitalismus*, Berlin, Dietz Verlag.

JACOMY M., VENTURINI T., HEYMANN S., BASTIAN M., 2014, "ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software," *PLoS ONE*, *9*, 6.

JAMES H., 2006, *Family capitalism : Wendels, Haniels, Falcks, and the continental European model*, Cambridge, MA, Belknap Press of Harvard University Press.

JEIDELS O., 1905, *Das Verhältnis der deutschen Grossbanken zur Industrie mit besonderer Berücksichtigung der Eisenindustrie*, Leipzig, Duncker & Humblot.

KALEMLI-OZCAN S., SORENSEN B., VILLEGAS-SANCHEZ C., VOLOSOVYCH V., YESILTAS S., 2015, "How to construct nationally representative firm level data from the ORBIS global database?", NBER Working Paper Series, 21558, Cambridge, MA.

KEISTER L.A., LI M., LEE H.Y., 2021, "Do You Need Business Assets to Be Rich?," *Socius: Sociological Research for a Dynamic World*, *7*, pp. 1–3.

KINDERMAN D., 2017, "Challenging varieties of capitalism's account of business interests: Neoliberal think-tanks, discourse as a power resource and employers' quest for liberalization in Germany and Sweden," *Socio-Economic Review*, *15*, 3, pp. 587–613.

LARSEN A.G., ELLERSGAARD C.H., 2018, "The inner circle revisited: the case of an egalitarian society," *Socio-Economic Review*, *16*, 2, pp. 251–275.

LILLIE K.J., 2020, *Transnational Class Formation in a School for the Global Elite*, Thèse de doctorat, University College London.

LIU G., 2020, "Data quality problems troubling business and financial researchers: A literature review and synthetic analysis," *Journal of Business & Finance Librarianship*, *25*, 3–4, pp. 315–371.

MACE M.L., 1971, *Directors: Myth And Reality*, Boston, Harvard Business School.

MARX K., 1887, *Capital. A Critique of Political Economy. Volume I. Book One: The Process of Production of Capital.*, Moscow, USSR, Progress Publishers.

MILLS C.W., 2000, *The Power Elite*, Oxford, Oxford University Press.

MIZRUCHI M.S., 1992, *The structure of corporate political action : interfirm relations and their consequences*, Cambridge, Harvard University Press.

MIZRUCHI M.S., 2004, "Berle and Means revisited: The governance and power of large U.S. corporations," *Theory and Society*, *33*, pp. 579–617.

MIZRUCHI M.S., 2013, *The fracturing of the American corporate elite*, Cambridge, Harvard University Press.

MOSCA G., 1939, *The Ruling Class*, McGraw-Hill Book Company.

MURRAY J., 2017, "Interlock Globally, Act Domestically: Corporate Political Unity in the 21st Century," *American Journal of Sociology*, *122*, 6, pp. 1617–1663.

O'HARA W.T., 2004, *Centuries of success. Lessons from the World's Most Enduring Family Businesses*, Avon, Adams Media.

PARETO V., 1961, "The Circulation of Elites," dans PARSONS T., SHILS E., NAEGELI K.D., PITTS J.R. (eds.), *Theories of Society. Foundations of Modern Sociological Theory*, New York, The Free Press of Glencoe, pp. 551–558.

PIKETTY T., 2014, *Capital in the twenty-first century*, Cambridge, Belknap Press.

RAUB B., JOHNSON B., NEWCOMB J., 2010, "A Comparison of Wealth Estimates for America's Wealthiest Decedents Using Tax Data and Data from the Forbes 400," *103rd Annual Conference on Taxation*.

RUNGI A., MORRISON G., PAMMOLLI F., 2018, "Global Ownership and Hierarchies of Firms. 'What is essential is invisible to the eye.'" https://ssrn.com/abstract=3031955

SCHEIDEL W., 2018, *The Great Leveler . Violence and the History of Inequality from the Stone Age to the Twenty-First Century*, Princeton, Princeton University Press.

SCHOENEMAN J., 2022, "Corporate connections: how centrality, structural dependence, and transnationalization relate to lobbying activity," *Interest Groups and Advocacy*, pp. 1–23.

SCHRÖDER C., BARTELS C., GÖBLER K., GRABKA M.M., KÖNIG J., SIEGERS R., ZINN S., 2020, "Verbesserung der Forschungsdateninfrastruktur im Bereich Hochvermögender mit dem Sozio-oekonomischen Panel (SOEP)," Berlin.

SCOTT J., 1985, "Theoretical Framework and Research Design," dans STOCKMAN F.N., ZIEGLER R., SCOTT J. (eds.), *Networks of Corporate Power. A Comparative Analysis of Ten Countries*, Cambridge, Polity Press, pp. 1–19.

SKOCPOL T., HERTEL-FERNANDEZ A., 2016, "The Koch Network and Republican Party Extremism," *Perspectives on Politics*, *14*, 3, pp. 681–699.

SMITH M., YAGAN D., ZIDAR O., ZWICK E., 2019, "Capitalists in the Twenty-First Century," *The Quarterly Journal of Economics*, *134*, 4, pp. 1675–1745.

SOMBART W., 1913, *Der Bourgois*, Berlin, Duncker & Humblot.

STAMM I.K., 2016, "Coordination Tasks and Negotiation Modes of Linked Lives in Entrepreneurial Families," *Journal of Marriage and Family*, *78*, 4, pp. 939–956.

UBS, PWC, 2020, "Riding the storm. Market turbulence accelerates diverging fortunes. Billionaires insights 2020." https://www.ubs.com/content/dam/static/noindex/wealth-management/ubs-billionaires-report-2020-spread.pdf

USEEM MICHAEL., 1984, *The inner circle: large corporations and the rise of business political activity in the U.S. and U.K*, Oxford; New York, Oxford University Press.

USEEM M., 2015, "From classwide coherence to company-focused management and director engagement," *Research in the Sociology of Organizations*, *43*, pp. 399–421.

VALEEVA D., 2021, "Where is the backbone of the transnational corporate elite?," *Global Networks*, 22, 3, pp. 547-563.

WEALTH-X, 2021, "World Ultra Wealth Report." https://go.wealthx.com/world-ultra-wealth-report-2021

WILSON S., 2021, "miceforest." https://github.com/AnotherSamWilson/miceforest

WRIGHT E.O., 1976, "Class Boundaries in Advanced Capitalist Societies," *New Left Review*, *98*, 1, pp. 3–41.

WRIGHT E.OLIN., 1998, *Classes*, London and New York, Verso.

YOUNG K.A., BANERJEE T., SCHWARTZ M., 2018, "Capital strikes as a corporate political strategy: The structural power of business in the Obama era," *Politics and Society*, *46*, 1, pp. 3–28.

YOUNG K.L., GOLDMAN S.K., O'CONNOR B., CHULUUN T., 2021, "How white is the global elite? An analysis of race, gender and network structure," *Global Networks*, *21*, 2, pp. 365–392.

ZEITLIN M., 1989, *The Large Corporation and Contemporary Classes*, New Brunswick, New Jersey, Rutgers University Press.

ZEITLIN M., EWEN L.A., RATCLIFF R.E., 1974, "'New Princes' for Old? The Large Corporation and the Capitalist Class in Chile," *American Journal of Sociology*, *80*, 1, pp. 87–123.

3rd Article:

Linking Wealth and Power[1]

Direct Political Action of Corporate Elites and the Wealthiest Capitalist Families in the United States and Germany

Abstract

This study inquires whether two groups of individuals with power in the economy directly translate it into political power in capitalist democracies: The corporate elite and super-rich capitalist families. It does so by analyzing three potential "avenues of influence": Lobbying or party donations through controlled firms, and individual party donations. Shareholders and managers of the largest 1,091,151 German and US firms (from the ORBIS database) are analyzed. First, 6,227 members of 1,854 US and German families with an estimated family net worth of at least 250 million USD or EUR are identified. Second, the national corporate elites are identified. Individual and firm data is then used to predict lobbying and party contribution in 2019-2021 with logistic regressions. Results suggest that direct political action on the part of the super-rich and the corporate elite is much more prevalent and more ideological in the United States than in Germany. If they engage at all, the super-rich tend to be a very conservative group who use all three avenues of influence complementarily. However, the magnitude of super-rich and elite money does not favor the idea of an "oligarchy" in either of the two countries, at least not through these direct and visible channels.

Submission status:
Submitted to American Sociological Review.

[1] I am thankful to Olivier Godechot and Cornelia Woll for helpful discussions of this project during my stay at Sciences Po Paris. Further thanks to Jens Beckert, Mark Mizruchi, and Joshua Murray for comments on previous versions of this manuscript. This project benefitted greatly from comments and feedback received at the MaxPo closing conference in Paris on 21 October 2022, and the workshop on top wealth at Leibniz Centre for Contemporary History on 10 and 11 November 2022 in Potsdam.

1. Introduction

Do individuals with power in the economy translate it into political power in capitalist democracies? And if yes, what do these individuals want to achieve? Are they just pragmatically using access to influence economic policy, maximizing their own profits? Or do they use their economic power to advocate for an ideological agenda that might be different from the rest of voters? The question what power "business" has is a classic one in the social sciences and already had a prominent place in Marx's (1887) chapter on the working day from 1867. Some recent developments give it ennewed special relevance. As Piketty (2014) argued, wealth concentration at the top of the distribution in many economically developed countries might lead to continuous concentration of large fortunes among few super-rich families. At the same time, what he calls superstar managers receive astronomical salaries. Critical observers warn that such concentration could lead to the concentration of political power and potential for oligarchy (Winters and Page 2009).

Money at the bank does not necessarily enable to buy relevant influence in functioning democracies. The mechanisms spelling out exactly which forms of owned assets can be used how to influence politics are not easily specified. The richest families for example own to a significant extent business wealth (Keister, Li, and Lee 2021). This does not only provide them with a continuous share of profits, but also with potential power over future investments and jobs. The latter is only the case if they can also control the behavior of firms to which their business assets belong. Firms with dispersed ownership may also be controlled by hired executives who may control the power related to a firm and its assets, investments, and divestments. Specifying such mechanisms today therefore raises old questions of ownership and control over capital which have been discussed since the genesis of the joint stock company (cf. Dahrendorf 1959:21–23; Scott 1990). Another way to frame the problem is to ask: who are the individuals controlling capital, and therefore having the potential to leverage it for political power? Are owners back or does the separation of ownership and control hold and it is corporate executives who control capital through the firms which hired them (Berle and Means 1991; Burnham 1975; Mizruchi 2004)? Is there (still) an inner circle corporate elite recruited from the two groups that feels responsible for taking instrumental political action in favor of capital in general (Chu and Davis 2016; Heerwig and Murray 2019; Mizruchi 2013; Useem 1984)?

In this article I approach this question by empirically examining three avenues of influence that individuals and "their" corporations can pursue in two of the largest democratic capitalist economies, namely the USA and Germany. Following the work of Bonica (2016), these avenues of influence are individual party contributions, firm party contributions, and firm lobbying. The focus of the study is to differentiate between the countries' corporate elites, and national super-rich

owning capitalist families as instrumental political actors. In addition to the question how these groups might take direct political action, I also study what they are trying to achieve in terms of support for conservative or liberal parties. Three research questions are asked for the two countries: *How relevant is the magnitude of super-rich and corporate elite money in the political process? What explains which of the three avenues of influence are used by the super-rich and the corporate elite? How do the super-rich and the corporate elite lean ideologically?* I exploit three types of data sources to examine these questions. First, a collaborated rich list stemming from journalistic and other public sources. It includes names and firm names of 3,250 families with an estimated family net worth of 250m USD or EUR in recent years (Bornefeld 2019). The study focusses on capitalist individuals and families only, defined as having at least one family member somehow involved in a company as either manager or shareholder. Second, a large sample of overall 4 million firms including financial, ownership, and management data from the Bureau van Dijk ORBIS database (Bureau van Dijk 2020). Third, multiple sources on individual and firm party contributions, as well as lobbying expenses of firms in Germany and the US provided by official sources and NGOs (Deutscher Bundestag 2022; Lobbycontrol 2022a; OpenSecrets 2022). The article is structured as follows. Section two briefly presents previous literature and derives hypotheses from it. Section three discusses relevant country differences that need to be kept in mind to evaluate results. Next, section four presents' data sources and data preparation, variables, used samples, and methods. Section five presents results followed by discussion and conclusion in section six.

2. Previous literature and hypotheses

To locate the three research questions outlined above, it is necessary to briefly recapitulate existing bodies of literature. I present three different strands of literature for the present study. The first surrounds the question who holds economic power or who rules the corporations, the second is the question whether there is any unity or coordination within and between these groups of powerful individuals, and the third is what we can expect members of these groups to want ideologically from the political process. The three strands and related hypotheses are presented subsequently in the following.

Especially in Marxist thought, the question of who controls the corporation is a long debated one (cf. Mizruchi 2004; Scott 1990). During Marx's lifetime, individual capitalists controlling single companies were the rule. However, this configuration of ownership changed in the historical development of capitalism. In the 1930s, Berle and Means (1991) diagnosed a widespread separation of firm ownership from firm control. They found that the largest share of US companies was not run by its owners anymore, but rather by hired managers. This challenged Marxist conceptions of a clearly defined ruling class which required a powerful group that owned

and controlled the means of production, and to leverage this power to "rule". Theoretically this implied that ownership of, and control over, capital had to be seen as two different things that *can* but do not have to be present simultaneously. It was widely regarded as evident that independent managers with interests separate to those of the owners seemed to have taken over (Allen 1981; Berle and Means 1991; Burnham 1975; Chandler 1977), and were from then on analyzed as separate groups (e.g. Gourevitch and Shinn 2005; James and Soref 1981; Jensen and Meckling 1976). An underlying question when analyzing political action of individuals with economic power is how to understand this group conceptually: Should they be seen as a capitalist class advocating for some general interest of capital? Or is the empirical reality better described as more loose networks of elite managers whose motivation is to maximize profits by realizing interests of the specific fractions of capital which hired them, e.g., sectors or single corporations? A central question therefore is whether the corporate elite and (super-rich) firm owners are connected among each other, have the same interests, and therefore act in the same way. This is the starting point for a first set of hypotheses.

In terms of political action, I expect those individuals who own businesses to a significant extent (most often have founded them), to be deeper and in more ways devoted to their firms. This is because they are often not only economically related to them, but they are firms that they or their ancestors built up and that therefore also have value beyond the economic value. Their relationship to their firms is also often deeply intertwined with their family relations (Stamm 2016). They can therefore be regarded as having more at stake in emotional, status, and personal wealth terms, in comparison to hired managers. For the latter group, economic success is more about their individual occupational performance but less tied to ownership interests or loyalty towards specific capital. Economic success also translates less directly to individual gains for managers – only mediated by their employment contracts through negotiated bonusses, etc. Their incentives to influence politics to secure firm interests should therefore be lower than for owners. In comparison to families owning smaller or less valuable firms, they can be expected to receive more attention and possibly also more responsiveness from politics. This is due to their relative importance and structural power by representing a relatively large block of capital personified in one family or even one individual. I therefore expect that due to this cooccurrence of incentives (controlling family having much at stake) and opportunities (being more likely heard):

H1: Firms in which super rich are involved are more likely to be politically active.

In line with this, I would also expect the extent of this effect to be related to the degree of control super-rich families have. If they only hold a controlling share but are not otherwise involved

in management, they only have some occasions on which to raise issues that are important to them to be implemented by management. If they are at least on a supervisory board of a firm, they should have more frequent and institutionalized opportunities to influence firm behavior and strategy. If they are themselves members of the executive management, they have the largest degree of control and opportunities to establish a political action strategy or more short-term political action such as party contributions. I therefore expect:

H2: The more control a super-rich family has over a firm (controlling share < supervisory position < executive position), the more likely the firm to is to take political action.

The second relevant preposition is based on a strand of literature which speaks to the question whether there is any unity between individuals that have control over different firms. This is important because business has long shown to coordinate and act in a unified way as a critical number of businesspeople as Lindblom called it. The larger the share of capital acting together, the more critical I expect its representatives to be. Ideas about cohesion of a capitalist class were long formulated in Marxist thought (Hilferding 1955; Mosca 1939; Sweezy 1951), and later for example by Zeitlin (1989) or Domhoff (1967). It was argued in one way or another for the idea that a central elite group of owners and managers of the large corporations represent a capitalist class, i.e., a group which pursues the interests of capital by influencing politics. In the 1980s, Useem (1984) suggested that a highly interlocked circle of managers controlling firms at the center of national economies take action in the name of business in general. It can only be noted here briefly that in distinction to previous Marxist ideas, it was now about an elite of hired managers and not about a class of rich owning families anymore (cf. Mizruchi and Bey 2003; Scott 2008). Cohesion among individuals from both groups was long analyzed by looking at interlocking directorates, i.e. shared board memberships that indicated coordination. While it has been argued that the inner circle fractured in the US since about the 2000s, and was no longer existent today (Chu and Davis 2016; Mizruchi 2013), others maintain that an inner circle persists at least in terms of direct political action in the US and in other countries (e.g. Banerjee and Murray 2020; Comet 2019; Heerwig and Murray 2019; Larsen and Ellersgaard 2018). From my reading, the common suggested mechanism in most of these versions still is what Mizruchi (2013:129) summarizes the former functions of networks of interlocking directorates to be:

> "First, they served as centers through which information of relevance to the entire corporate community was exchanged and transmitted. Second, they served as a source of normative consensus and stability among the leaders of the largest corporations, in part by helping to forge

> similar worldviews and behavior and in part by sanctioning deviant and/or irresponsible actors. And third, they served as a source of cognitive range— or breadth in outlook— in terms of concern for the larger business community as well as the larger society. These roles— information, normative consensus, and cognitive range— were accomplished largely, although not exclusively, through bank boards of directors."

Murray (2017:1622–26) remarks that coordination keeps happening but not necessarily only through bank boards but also boards of other firms. I follow Murray's (2017) suggestion and results, and assume as a working hypothesis that the inner circle still exists. In line with previous research (Broyles 1993; Burris 2005; Goerres and Höpner 2014; Heerwig and Murray 2019), the inner circle is expected to be more likely to take political action.[1] I expect this group to act on behalf of business in general because they have the most central access to information, and this still has the effects as described by Mizruchi above. If a member of a super-rich family is also a member of this inner circle, I would expect them to have even higher chances to take political action – not only for their own specific economic interests but also for the broader economic interest of business in general. This is also in line with previous research on party contributions of wealthy families in the US of the 1970s (Allen and Broyles 1989). I therefore expect that:

H3a: Firms with inner circle members in management are more likely than other firms to take political action.

H3b: Firms with super-rich inner circle members in management are more likely than firms with other inner circle involvement, and more likely than other firms to take political action.

In terms of individual party contributions, previous evidence for the US shows that wealthy donors are an important and generous group to finance US PACs. Heerwig (2018:1005) for example highlights that "approximately 70% of super PAC funds have come from megawealthy individual donors". I expect some variation within the super-rich. First, and along with the current set of hypotheses, I expect super-rich inner circle members to be more likely than other super-rich family members to make individual party contributions for the reasons just mentioned:

H3c: Super-rich inner circle members are more likely than other super-rich individuals to make party contributions.

A famous differentiation among large fortunes furthermore is the idea of "old money", i.e. durable wealth that was generated in the past and transmitted to further generations (Beckert 2022a). Burris (2000) discusses a longstanding debate in US social science assuming "new money", i.e. more

[1] There was however also some contrary evidence reported in the past (Clawson and Neustadtl 1989).

recently generated fortunes, to be related to more conservative political views and action. However, his empirical findings analyzing political action of the 1995-1997 Forbes 400 rich list shows that the groups do not differ much in terms of political action but there is a tendency of old wealth families to be more conservative. The author explains this by old wealth to be more likely socialized in more conservative upper classes and more densely connected to a social network of this class, both fostering more conservative views. For the wealth distribution more broadly, Arndt (2019) also finds that if a German individual receives inheritances, prizes or gifts, it places itself more left on a left-right scale. In addition, Allen and Broyles (1989) find that old wealth donated more than newer wealth in the US. Following these findings, I expect:

H4: Individuals from "old wealth" families are more likely to take political action.

As discussed above, I expect the size of a fortune to be related to the incentives to take political action. The larger a fortune, the more is at stake and the more individuals and families are part of the public discourse and may be the target of political claims to introduce redistributive measures such as a wealth tax, inheritance taxes, or even expropriations. These of course are rather backed by left leaning parties than conservative ones. In line with this, Allen and Broyles (1989) showed that more visible wealthy families occurring in the elite directory *Who is Who?* were more likely to make a party donation. I therefore expect more wealthy families to have higher odds to make individual party contributions at all, and more conservative contributions.

H5: Estimated family net wealth increases the odds to make any party donation.

As a final hypothesis concerning the question *What explains which of the three avenues of influence the super-rich and the corporate elite use?*, I speculate about the strategic use of the three avenues of influence complementary. One could expect "two faces of capital" (Burris 2001) and that political action by individuals differs from firm political action. Alternatively, one could expect super-rich families to engage in strategies of exploiting multiple avenues of influence complementary. I would expect a combination of both. On the one hand, I expect individual contributions to be more ideological than firm contributions in line with previous findings (Bonica 2016; Burris 2001):

H6a: Individual contributions of super-rich family members are more conservative than their firms' contributions.

On the other hand, I also expect some families to act politically through multiple channels complementary as part of more complex political strategies in line with recent findings on firm party contributions (Kim, Stuckatz, and Wolters 2020).

H6b: If a family made a party contribution, their firms are also more likely to make a contribution.

H6c: If a firm made a party contribution, their controlling family is also more likely to make a party contribution.

The third research question asks how the corporate elite and the super-rich lean ideologically. A multitude of previous empirical evidence gives the chance to derive hypotheses on this question. Several analyses of individual and corporate contributions of the super-rich as well as corporate elites in the US have shown the remarkable conservatism of this group (Allen and Broyles 1989; Bonica and Rosenthal 2015; Burris 2000). Studies on wealth more generally also show that wealth possession is related to more conservative views in the US, as well as in Germany (Arndt 2019; Page, Bartels, and Seawright 2013). It was mentioned before that several authors suggested that individuals with economic power can act either individually or through their firms, and that action might differ between the channel used (Bonica 2016; Burris 2001). Their contributions show that individual political action, such as direct individual contributions, are on average more ideological and support candidates also on non-economic issues, than political action of firms (Bonica 2016; Burris 2001). Firm political action, on the other hand, was better explained as being more pragmatic and is more strongly related to profit interests of a firm and therefore for example donating to candidates who are most likely to win (Burris 2001). Kim, Stuckatz and Wolters (2020) show how firms make party contributions to candidates for them to be more responsive to later lobbying activities which was already suggested by studies from the past decades (Bonica 2016:370).

When it comes to ideology of the two groups, some studies find that inner circle members are more conservative than other executives (Broyles 1993; Burris 1991; Clawson and Neustadtl 1989), while others find that inner circle members are actually rather pragmatic than ideologically driven in their political action (Heerwig and Murray 2019). Differentiating between individual and firm political action, I expect the inner circle to be more conservative than other corporate managers when acting as individuals, and their firms to act more pragmatic in line with (Bonica 2016; Burris 2001):

H7a: Inner circle members make more conservative individual contributions than managers who are non-members.

I assume an ordinal rank of the strengths of correlations of being super-rich and being a member of the inner circle corporate elite. Since the super-rich have more at stake and will be more targeted

and affected by potential left policies and campaigns, I expect them to be even more conservative than the inner circle:

H7b: Super-rich family members make more conservative individual contributions than the inner circle, and more conservative than other managers.

For firms controlled by the super-rich or connected to the inner circle, I expect more pragmatic political action than for individuals from these groups, i.e., less leaning towards conservative parties. This is due to the profit motive dominating in the institutional context of the firm and they need access to successful politicians of all couleur:

H8a: Firms in which inner circle members are involved take pragmatic political action not leaning towards conservative parties.

H8b: Firms controlled by the super-rich take pragmatic political action not leaning towards conservative parties.

The next section presents case selection of Germany and the US and the rational behind choosing these cases to test the developed hypotheses.

3. Case selection and differences between the US and Germany

This study and its hypotheses are framed as studying similarities in political action of the super-rich and the corporate elite as powerful groups in the economy. The US and Germany were selected for several reasons. First, both are among the largest global economies under democratic rule, in which the question of political action of rich individuals from the economy is of given normative relevance. Second, in 2022 the USA is the country with most billionaires and most ultra high net worth individuals (UNHWs)[2] in the world followed by Germany on rank 4 after China and India in terms of billionaires (Forbes 2022), and rank 3 in terms of UHNWs (Wealth-X 2022:13). Third, both selected countries are nation states (in contrast to the EU as a whole). Fourth, both are capitalist democracies (in contrast to China), and fifth, both have decent data availability and quality on individual and firm political action (to my knowledge in contrast to India). Furthermore, sixth, there were recent federal elections in both countries (US: 2020; Germany: 2021) which also makes timing of political action more comparable. However, despite these similarities there are relevant country differences that need to be considered when comparing the US and Germany. I will briefly discuss three of those differences especially relevant for the study at hand. First, these are differences of corporate governance such as the ownership structure and shareholder

[2] Wealth-X defines UHNWs as owning 30m USD or more.

concentration, as well as management positions. Second, differences in party financing systems. And third, general differences in the legitimacy of affluence in both countries.

Ownership of companies is much more dispersed in the US compared to Germany (e.g. De La Cruz, Medina, and Tang 2019). This implies that companies less frequently have a controlling shareholder which is relevant here because I partly assume groups of individuals to be able to exert control over firms. The German "Mittelstand", i.e. mid size family led companies, is famous for being world market leaders in niche markets, and an important reason for why Germany has so many super-rich families (Berghoff 2006; Lehrer and Celo 2016). The fact that a super-rich family controls a firm with a relevant size is therefore more prevalent in Germany. However, the most valuable family-controlled firms in terms of market capitalization with super-rich involvement are the US firms of billionaires such as Jeff Bezos, Elon Musk or Warren Buffett. As a final point to be mentioned here, management boards especially of listed companies are organized very differently in the US and Germany. In the US, a one board system dominates in which most of the power is concentrated in one board with a chairman and other business executives. In Germany, a two-board system dominates in which the executive board of managers is supervised by a supervisory board often called *Aufsichtsrat*. These boards are often very large and may include professional trade unionists and thereby sometimes professional politicians. This also implies that an inner circle or the core of the corporate elite may include such trade unionists – and not only managers - when including *Aufsichtsräte* in the network. I circumvent this by only taking into account those members of *Aufichtsrat* who also hold an executive position in some company.

There are relevant differences in the financing of parties between the US and Germany. Cagé (2020) describes these differences for several countries. Three differences seem especially relevant. First, party financing systems work very differently in the US. As possibly the most relevant consequence, party contributions are much more important to parties and candidates in the US, while German parties finance themselves much more through membership fees and public financing (Cagé 2020:135–61; Fink 2017:222–23). It might therefore be that German parties simply do not need the money from private donations – or at least to a lesser extent. In both countries, there is no maximum contribution an individual can contribute to a party.[3] In Germany, only donations above 10,000€, from individuals or from firms, must be reported with the identity of the donor. Smaller donations up to 3,300€ are incentivized by doubling the amount with public funding, but only up to this threshold. In the US, there is an identifying name for almost all individual contributions. For individual contributions of 200$ and more, it is legally required to state the donor's employer. The second difference is that while German firms can make direct

[3] This is de facto the case in the US as well since the introduction of super-PACs in 2010 (Cagé 2020:39).

contributions to parties, in the US so called political action committees (PACs) must be set up and are well regulated. PACs are something different than direct contributions. A firm can set up a PAC and collect contributions of up to 5,000$ by employees (or other individuals) and other stakeholders. The collected sum is then distributed to political parties and candidates at the discretion of the firm such as the government affairs department in line with the firm's interests. Nevertheless, one can see that these PACs clearly do have an ideological component. The PAC of Koch industries, the main firm of the famous Republican supporters Charles and David Koch, donated 97.14% of the collected amount to Republicans in the 2020 election cycle. At the same time, Soros Fund Management the firm of Democrats supporter George Soros overwhelmingly collected contributions for democratic candidates (OpenSecrets 2022). US PACs have been studied massively to understand firm political action (Burris 1991, 2001; Clawson and Neustadtl 1989; Mizruchi 1992; Murray 2014) but also donations by individuals to PACs (Burris 2001; Heerwig 2018; Heerwig and Murray 2019). Firm contributions in Germany are far more direct political engagements of organizations, but have been studied far less frequent (Cagé 2020; Fink 2017; Goerres and Höpner 2014).

Wealth concentration and affluence can be expected to vary in many ways by culture and institutional context. One important dimension of this is the legitimacy of wealth, implying whether the group of the rich tends to show-off their wealth or whether they are humbler and more secretive about their lifestyles and even how they look. While there will obviously be variation within these groups and within countries, the German super-rich on average tend to be the latter. Several of the richest German individuals actively suppress the publication of pictures of them, including Lidl founder Dieter Schwartz (Würzer 2019), and Aldi founders Theo and Karl Albrecht (Rühle 2010). There might also be generational differences and younger members of super-rich families are much more present in German media, be it in interviews as young entrepreneurs such as the cookie heiress Verena Bahlsen (Kapalschinski 2019) or the drug store chain heirs Christoph Werner and Raoul Rossmann (Gnirke, Kühn, and Salden 2021). Or be it as philanthropes such as Marlene Engelhorn (Bubola 2021) or Antonis Schwarz (Friese 2021). In other words, especially the older and first-generation richest Germans do not seem to want public attention which is obviously relevant for the study of publicly visible political action. Gajek (2016) shows that this "dialectic of visibility and invisibility" for the richest Germans was already historically relevant in the 1960s. But what is the difference to the US? First of all, being secretive and almost socially invisible does not seem to be the case for the richest Americans as was already suggested by Veblen's (1899) concept of "conspicuous consumption". It is easy to find contemporary images, interviews as well as political interventions by the richest among the US Forbes 400. Page, Seawright, and Lacombe (2018) find that while the majority of US billionaires do not speak about politics publicly, there are

at least a hand full who frequently do – and that these are among the most wealthy ones. What I suggest is that there might be different cultures of legitimacy of wealth in both countries which makes for example conspicuous consumption but also political attitudes and involvement by the super-rich more acceptable or even appreciated in the US. Cagé (2020:22–23) rejects the idea that cultural differences could be responsible for country differences in political action of the wealthy. But the empirical reality is hardly explained by differences in institutional settings alone in my view. However, despite this anecdotal evidence and journalistic sources, a systematic explanation of these differences is so far absent and comparative wealth research with a focus on them seems to be developing only now (cf. e.g. Beckert 2022b). It can therefore only be mentioned here that the possession of large fortunes underlies norms which should be expected to vary by culture which might be relevant also for how publicly those who possess large fortunes engage in politics. Although further research is necessary on this, the idea will also be included in the discussion of the results below.

4. Data and methods

Data sources

This section describes the used data sources, the firm samples used, and the matching procedures applied, first, to identify super-rich family members in the firm data and, second, to match ORBIS individual and firm data to political contribution data. Figure 1 visualizes all data sources to give a quick overview over the multiple sources and connections of the database used to test the developed hypotheses.

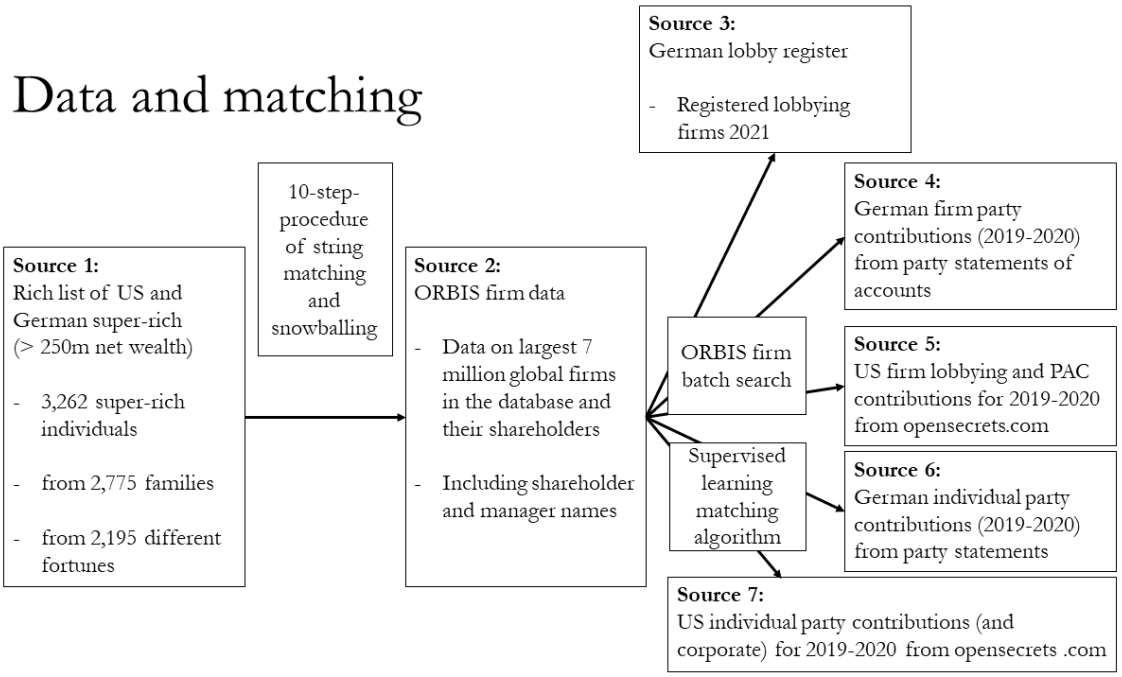


Figure 1: Data sources and an illustration of the applied matching processes.

Source 1 is a compilation of global rich lists extended with additional public data on the richest global families. It was provided by a private researcher in Cologne who is also involved in the research of the German rich list published by manager magazin (Bornefeld 2019; Neßhöver and Bornefeld 2018). The list includes family names, in some cases individual names, firm names, estimated net wealth, and in some cases founding years for 10,831 of the global wealthiest families with an estimated net worth of 250m EUR or USD. For the present study, only the 3,250 German and US families are considered.

This rich list is then matched to a second data source: a sample from the ORBIS firm database, the largest database with global coverage including data on 400m firms (Bureau van Dijk 2020). I scraped and downloaded a sample each from the database at two different time points: one in early 2020, and a second in the first half of 2022. Together these samples comprise financial data, shareholder data including names of shareholders, and data on management including individual names for the largest 3m global firms and their shareholders. The super-rich individuals and families from source 1 were then matched to shareholders and managers in the ORBIS sample with a 10-step rule-based string-matching procedure.[4] 1,854 (57%) of the US and German individuals and families known from the list could be identified with enough confidence in the ORBIS data. Table 1 presents numbers for the US and German families before and after matching and different forms of selection. This combined data source of the largest global firms and super-

[4] For details, see appendix 2.

rich involvement among them was then matched to four other data sources with information on firm and individual political action.

For firm political action, three sources were used. First, the novel German lobby register introduced in 2022 which includes data on individuals and firms. Registration is required by law for all lobbying organizations which intend to contact members of the Bundestag or the German government to lobby for their interests. It also includes information on estimated expenditures (Deutscher Bundestag 2022). Second, party contributions made by firms stem from publications of the Bundestag for donations of more than 50,000€, and from party statements of accounts for donations between 10,000 and 50,000€. They were gathered by German NGO Lobbycontrol (Lobbycontrol 2022a). The third source contains information on firm political action in the US and is provided as bulk data by the independent research group OpenSecrets (OpenSecrets 2022). Firm political action committee (PAC) party contribution data is based on Federal Election Commission data, lobbying data comes from the Senate Office of Public Records. Firm names from these three sources are matched to firm names from the ORBIS data with the batch search tool provided in the ORBIS portal. It uses firm names, aliases, and firm location to find the best match of a firm name among firms in the ORBIS database.

Individual party contributions stem from similar sources. For Germany, individual party contributions also come from Lobbypedia (Lobbycontrol 2022a), for the US they are provided by OpenSecrets including data on more than 4.8 million individual contributions (OpenSecrets 2022). However, individual names had to be matched to several million names of individual shareholders and managers from ORBIS. This was done by training a supervised learning matching algorithm with the Dedupe library using Python (Dedupe.io 2022). For Germany, matching was based on names, how common a name is, and cities of donors and firms. For the US data, employers name was also provided and used in addition which should lead to increased accuracy for matching of the US data. On the other hand, the US matching task is much more challenging because the number of individual donors is more than 4.8 million compared to only 2,071 in Germany. Based on these data consolidation and preparation steps, the next section presents the used independent variables and their distributions.

Table 1: Findings and sample composition of super-rich families, and individuals in ORBIS.

Selection	Cases					
	US		DE		Total	
	Absolute	Percent	Absolute	Percent	Absolute	Percent
Families						
Total number of families known from the original rich list[5]	1,978	100	1,272	100	3,250	100
Identified in ORBIS with at least one shareholder and/or director/senior manager[6]	1,081	54	773	60	1,854	57
Identified in ORBIS with at least one shareholder and/or director/senior manager and matched individual donation	421	21	33	3	454	14
Identified in ORBIS with at least one managed or controlled firm which made a party contribution or lobbied	381	19	111	9	390	12
	Absolute					
Individuals						
Total number of individuals known in the original rich list	2,311 individual members of the 1,978 known families		1,499 individual members of the 1,272 known families		3,810 individual members of the 3,250 known families	
Identified in ORBIS as a shareholder and/or director/senior manager	2,592 individual members of the 1,081 identified families		3,635 individual members of the 773 identified families		6,227 individual members of the 1,854 identified families	
Identified in ORBIS with at least one shareholder and/or director/senior manager and matched individual donation	563 (21%) individual members of the 2,592 identified families		37 (1%) of the 3,635 identified individuals		600 (10%) of the 6,227 identified individuals	

[5] Also individuals not identified in ORBIS were matched to individual donation data and included in the analysis of the individual sample.

[6] Only firms controlled by these families could be identified as such in the analysis here. Firms controlled by families that could not be matched might therefore be misclassified as not controlled by a super-rich family.

Variables

To test the hypotheses developed in section 2, three main independent variables are operationalized using the generated database. First, an indicator of the exact management position and whether it is a supervisory or an executive position. Second an indicator for corporate control, showing whether all members of a super-rich family together own 20%, as one of several common thresholds for control, or more of a company. Third, an indicator whether a member of the inner circle is involved in a firm's management.

Supervisory positions are defined as those positions that have a supervisory or advisory role in a firm's management (cf. Stokman, Ziegler, and Scott 1985:17; Windolf 2002:67–76). Supervisory boards can also be expected to have a say in long-term strategy making of firms, under which political action strategies such as opening a lobbying office are also likely to fall. Executive positions are defined as management positions with actual decision power over operations and virtually all processes in a firm. This includes senior management such as the "C-Suite" of CEOs, CFOs, etc. While the two roles differ, both are expected to have an influence on decisions over firm political action in the short-term and long-term and they are therefore used to test how super-rich family members may influence these decisions.

Exerting control over a corporation can be seen as success in the competition over influence on decisions over the allocation of resources within a company. Groups which compete over such influence include financiers, management, and owners (Scott 1997). Before the genesis of the joint stock company, firms were often led by owner-managers who owned a factory and thereby had full ownership and control over a firm. Since the separation of ownership and control, remaining control of owners depends on the size of their shares held versus that of other owners. In the study of corporate governance, several thresholds are commonly used to assume control of one owning individual or group over a firm (e.g. Gourevitch and Shinn 2005:17). I use the threshold of 20% as necessary to exert control over a firm. If an individual or a family own 20% or more of a firm, they are assumed to have an influence at least over long-term strategic decisions including firm political strategies. Therefore, I sum up all shares held in a firm that could be matched to one of the known super-rich families – be it direct holdings or holdings through other companies, foundations, trusts or the like. If a family owns 20% or more in a firm, the independent variable is set to 1, if not it is 0.

The idea of an inner circle of highly interlocked managers is operationalized here following the procedure suggested by Larsen and Ellersgaard (2017; see also Huijzer and Heemskerk 2021). In short, this can be summarized as follows. I map the network of interlocking directorates between all heads of corporate groups in the sample. An interlocking directorate is defined as a manager who is involved in management of more than one company and thereby forms a tie between these

companies. I include not only the board of directors in the analysis, but also senior management in line with previous work (Berle and Means 1991; Heemskerk and Takes 2016). Only heads of corporate groups are considered because otherwise too many highly clustered subsidiaries would falsely be identified as dense elites although ties actually only reflect bureaucratic ties and board memberships (Heemskerk and Takes 2016:98). The procedure according to Larsen and Ellersgaard (2017) then is to prune this network based on weighted edges and a measure of the minimum absolute number of individuals an individual and its direct neighbors are connected to. In other words, it is the number of other individuals an individual is connected to either directly or through its 1st degree neighbors. Interlocks are weighted with measures approximating tie strength. The weighting procedure suggested by the authors considers board size, as well as the absolute number of ties between pairs of individuals. In short, ties in large boards with more than 14 members are weighted less and redundant interlocks between the same individuals decrease logarithmically in their strength (for details cf. Larsen and Ellersgaard 2017). I use the calculated coreness score as a closeness to the inner circle. Inspected the networks, visually and defined inner circles for the two countries as groups which are reasonably closely connected and not implausibly to large to be connected to each other. Therefore, I define the German inner circle as all interlockers with a coreness score of 4, and the US inner circle all members with a coreness score of 10. Furthermore, for the German case I only keep those managers who are executive managers in at least one firm. This way I only include managers and not the many other kinds of members of *Aufsichtsräte* such as trade unionists or politicians. Figure 2 shows the full 2020 inner circles, as well as super-rich involvement in both countries in comparison.

In addition to these independent variables, several control variables are included in the regression analyses. These are separated between the sample of families, and the sample of firms. For the sample of families, mainly data provided in the original rich list was used. Family wealth estimates were simply used from the original sources without adjustment for inflation. They should be seen as a noisy signal of the wealth category they belong to (e.g. owning several billion vs. owning 300m) rather than exact values. The number of family members is determined in the matching process. Since family members are identified when they have the same last name, it is equal to the number of family members who are either shareholder or manager in the same companies as other members of super-rich families with the same last name. The more family members are involved in a company, the older I assume the fortune to be. This is because families grow over time, and entrepreneurs are assumed to include other family members subsequently and not from the start of a business endeavor. Finally, a variable is included indicating whether a firm in which a super-rich family is involved in either with a controlling share or in management has made a party contribution and/or lobbied.

For the firm sample, revenue, number of employees, firm founding year, global ultimate owner, and whether it is a listed firm or not is simply used as provided in the ORBIS database. In addition to that, the degree was added as a variable indicating the sum of interlocks of all managers in a firm's management. In case there was no data on management, no interlocks were assumed. The two samples and the distribution of variables is described in the next section.

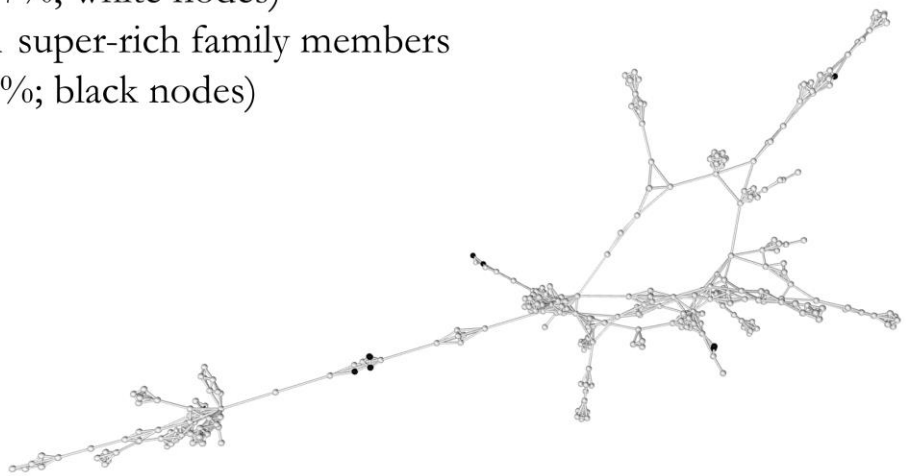
The largest global firms were selected from the ORBIS database in the following way. First, firms were selected that ORBIS classifies as large or very large (the two largest categories) in their database. This is equal to all companies with any of the following criteria: a revenue of 13m USD or more, total assets of 26m USD or more, or 150 or more employees. In early 2022, the ORBIS database included 3,081,589 firms fulfilling these criteria. In addition to that, all shareholders were added to the samples to be able to find shareholding individuals and more complex corporate structures higher up in corporate hierarchies. Therefore 1,365,322 firms that held shares in the 3m largest firms were added. From these, only the 347,242 German and 743,909 US firms were kept. This includes firms with holding companies in other countries but German or US subsidiaries or branches. In general, it is important to note that the sample includes subsidiaries. Firms are structured into corporate groups based on their global ultimate owning entity (GUO). The GUO is a variable provided by ORBIS which denotes the ultimately owning entity which can be any form of entity such as a holding company, an individual or family, or a state. In case no GUO was provided, it was assumed that a firm is independent and is its own GUO. In Germany, the 347,242 firms are subsidiaries of 72,753 distinct corporate groups. In the US, the 743,909 firms belong to 262,611 distinct corporate groups. Table 2 presents descriptive statistics for the sample of individuals and the used variables separated by country. Table 3 presents the same for the sample of firms.

Methods

Logistic regression models are estimated to test the developed hypotheses, and some additional descriptive findings are presented to set results into context. I estimated multiple imputation models, but results deviated from complete case analysis and there is not enough reason to be confident in the missing at random assumption. Therefore, single imputation by the mean was applied for continuous variables, and mode imputation for categorical variables. Complete case analysis is provided in Appendix 4b. It shows that results are consistent between these two methods. Standard errors are clustered at the corporate group level for the firm analysis, and at the level of original fortunes for the family analysis.[7]

[7] Original fortunes mean all families whose wealth origins from the same company. E.g. the Koch brothers and their families are counted as one fortune, but Warren Buffett and Bill Gates and their families as two fortunes.

a) German inner circle
329 other managers
(97%, white nodes)
11 super-rich family members
(3%; black nodes)



b) US inner circle
346 other directors
(97%, white nodes)
13 super-rich family members
(3%; black nodes)

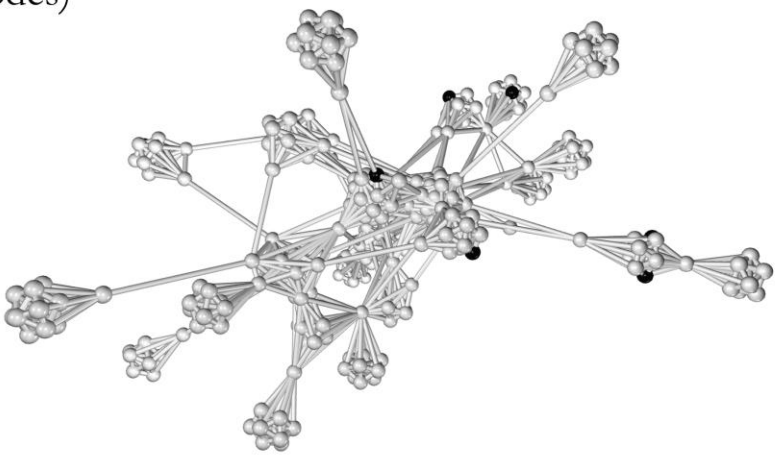


Figure 2: Inner circles of the 2020 interlocking directorate networks of individuals in the US and Germany. Inner circles were identified applying the procedure suggested by Larsen and Ellersgaard (2017) for Germany (a) and the US (b). Super-rich family members are denoted in black and other managers in grey. Networks are visualized with Graphia (Freeman et al. 2022).

Germany								USA							
Variable	**N**	**Complete**	**Missing**	**Mean**	**SD**	**95% Conf. Interval**		**Variable**	**N**	**Complete**	**Missing**	**Mean**	**SD**	**95% Conf. Interval**	
						Low	High							Low	High
Est. family net wealth	1,272	998	274	1,991	5,010	1,679	2,302	Est. family net wealth	1,978	1,368	610	4,559	13,346	3,851	5,267
Family members	1,272	739	533	4.95	4.59	4.62	5.28	Family members	1,978	1,005	973	1.71	1.56	1.62	1.81
Founding year	1,272	936	336	1,945	70	1,941	1,950	Founding year	1,978	1,223	755	1967	50	1964	1969

Variable	**Count**	**Percent**
Inner circle		
Not in inner circle	1,270	99.45
Inner circle	7	0.55
Family donated		
False	1,237	96.87
True	40	3.13
Firm political action		
No firm acted	1,165	91.23
Related firms lobbied and contributed	81	6.34
Related firm lobbied	16	1.25
Related firm made party contribution	15	1.17

Variable	**Count**	**Percent**
Inner circle		
Not in inner circle	1.958	98.89
Inner circle	22	1.11
Family donated		
False	1.513	76.41
True	467	23.59
Firm political action		
No firm acted	1.598	80.71
Related firms lobbied and contributed	191	9.65
Related firm lobbied	170	8.59
Related firm made party contribution	21	1.06

Table 2: Descriptive statistics of used variables in the sample of super-rich families.

Table 3: Descriptive statistics of used variables in the firm sample.

	Country	Complete cases	Binary variable = 1 (yes)	Firms	Missing values	Min	Max	Mean	q50	q90
Founding year	DE	204,364	-	347,242	142,878	1197	2022	1998.21	2018	2018
	US	538,963	-	743,909	204,946	1622	2022	1968.92	2010	2010
Revenue th. $	DE	72,123	-	347,242	275,119	-96,437	283,756,896	219,633	199,765	199,765
	US	185,740	-	743,909	558,169	-1,186,516	523,964,000	250,652	138,000	138,000
No. of employees	DE	172,376	-	347,242	174,866	1	662,575	311	306	306
	US	175,275	-	743,909	568,634	0	2,200,000	735	673	673
No. of interlocks	DE	347,242	-	347,242	-	0	32	0.024	0	0
	US	743,909	-	743,909	-	0	196	0.058	0	0
Firm is listed?	DE	347,242	1,155	347,242	-	0	1	0.003	0	0
	US	743,909	16,408	743,909	-	0	1	0.022	0	0
Firm is controlled by super-rich family?	DE	347,242	27,451	347,242	-	0	1	0.079	0	0
	US	743,909	28,095	743,909	-	0	1	0.038	0	0
Firm made party contribution to German party?	DE	347,242	181	347,242	-	0	1	0.001	0	0
Firm made party contribution to US party?	US	743,909	1,290	743,909	-	0	1	0.002	0	0
Controlling family donated?	DE	347,242	1,495	347,242	-	0	1	0.004	0	0
	US	743,909	16,858	743,909	-	0	1	0.023	0	0
Firm lobbied?	DE	347,242	985	347,242	-	0	1	0.003	0	0
	US	743,909	3,523	743,909	-	0	1	0.005	0	0

5. Results

In the following, results are presented relevant for the three research questions: *How relevant is the magnitude of super-rich and corporate elite money in the political process? What explains which of the three avenues of influence the super-rich and the corporate elite use? How do the super-rich and the corporate elite lean ideologically?* They are answered subsequently in conversation with the derived hypotheses.

The first research question sets the context for the subsequent two and puts into perspective the relative magnitude of money spent by the super-rich and the corporate elite in the US and Germany: *How relevant is the magnitude of super-rich and corporate elite money in the political process?* I present some descriptive explorative evidence to evaluate it. To begin, Table 4 shows the share of identified super-rich family members and inner circle members who engaged at all in any of the three avenues of influence. Quite strikingly, political action of both groups is much lower in Germany than in the US. A first finding is therefore that the number of known and matched super-rich individuals who engage in direct political action at all is very low in Germany with 6% being either donating individually or being involved with firms that act politically. It is moderate in the US with 49.4% pursuing any of the three avenues of influence.

Table 4: The share of individuals from super-rich families and the corporate elite that could be matched to the three avenues of influence.

		Donated individually		**Related to a firm which donated or lobbied**	
	Sample	Absolute	Percent	Absolute	Percent
Germany					
Known super-rich families	2,311	40	1.7	112	4.8
Inner circle[8]	138	0	0	9	2.6
US					
Known super-rich families	1,979	827	55.1	382	10.2
Inner circle	359	146	40.6	285	79.3

Note: The basis is all families included in the original list. *Donated individually* is the number of individuals to which an individual party contribution could be matched. *Related to a firm which donated or lobbied* is the number of firms that could be matched to a super-rich family as shareholder or in management, and either lobbied and/or made a party contribution.

[8] This includes only the 138 inner circle members who are also executives in any company. The used database only includes the family level for the super-rich and not for the inner circle. This is because the super-rich were matched from a separate source, while the inner circle is determined deductively from the ORBIS data.

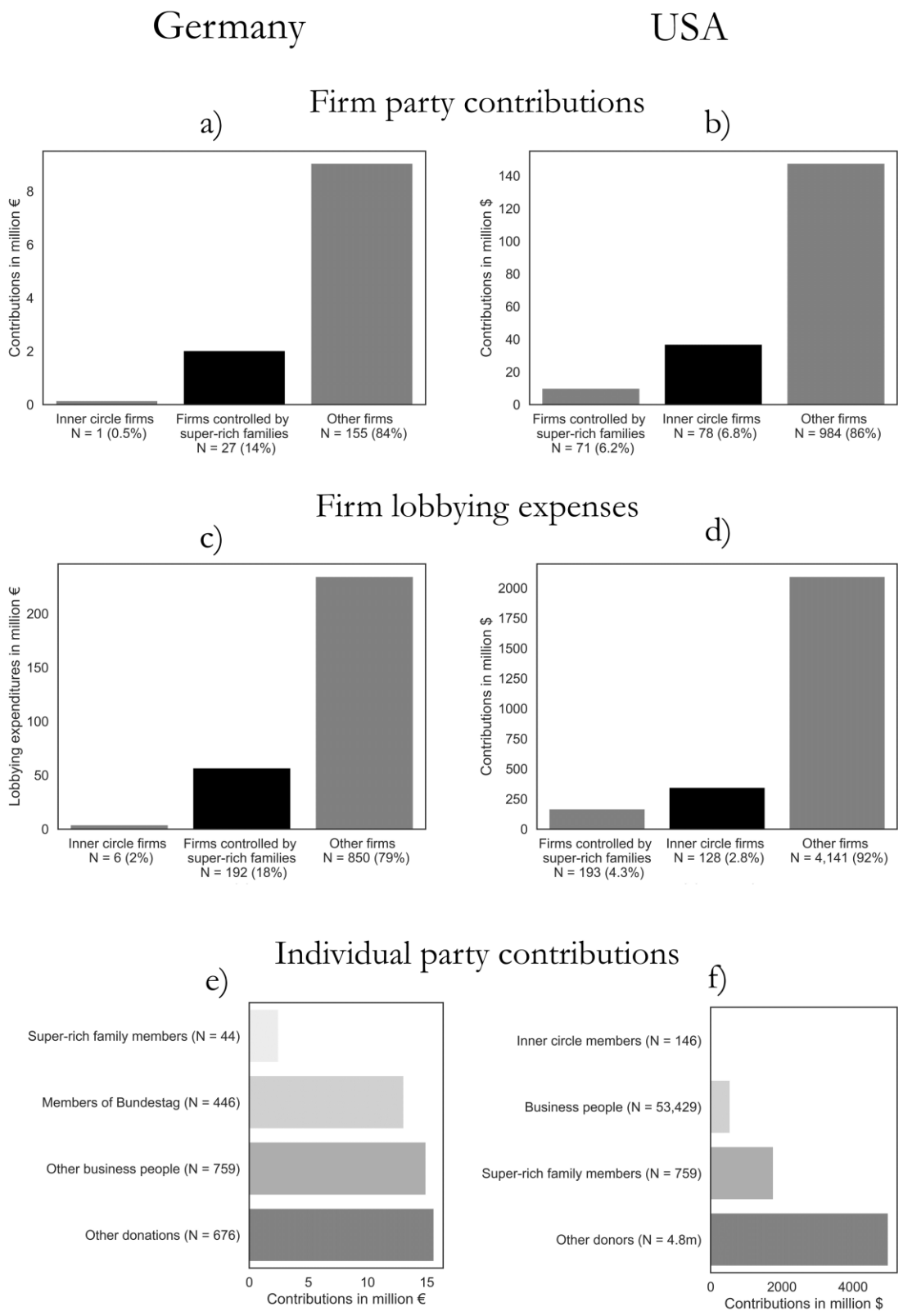


Figure 3: The relative sums of firm party contributions in Germany (a) and the US (b), firm lobbying expenditures in Germany (c) and the US (d), and individual party contributions in Germany (e) and the US (f). Inner circle firms are firms with at least one member of the inner circle in their management.

Figure 3 shows the relative sum of money that was spent by corporations on party contributions, spent by firms on lobbying, and donated by individuals in the two countries. Several interesting patterns of differences between countries and between the three avenues emerge. First, the relative amount spent either by the super-rich or by the corporate elite are much lower than money spent by other firms. It does not seem to be the case that money controlled by either of the two groups is dominant in any of the three avenues of influence. Second, at the level of corporate political action, money mobilized by the corporate elite is more relevant in the US, while money mobilized by the super-rich is more relevant in Germany. This is a first interesting and very relevant difference to understand the structure of business power exercised by individuals in the two countries. Third, in terms of individual contributions to parties, the inner circle is impressively negligible in Germany, while the super-rich seem to be relevant especially in the US. In the former case, only a tiny minority of 44 individuals contributes at all. The total value of their contributions is large given their small number, but in fact tiny compared to all money donated by other groups such as the members of the Bundestag themselves. While there is a bias through super-rich donations who each donate much more than individual members of other groups, the total amount is still low. In the US case, the 827 identified super-rich family members contribute about one third of the amount that the 4.8 million other individual contributors donate. Even when excluding the top outliers Michael Bloomberg and Tom Steyer, who financed their own candidacy with together about one billion USD, it is still about one fifth of all other donations. Therefore, a clear overrepresentation of super-rich money in comparison to all other individuals is visible in the US.

With these figures in mind, I now turn to the more explanatory second research question: *What explains which of the three avenues of influence the super-rich and the corporate elite use?* The first set of hypotheses concerning this question suggests that firms controlled by super-rich families are more likely to take political action. Table 5 presents odds ratios estimated from logistic regressions. Three different variants of control were suggested above: Having a controlling share, being a member of a supervisory board, or being in executive management of a firm. Neither in Germany, nor in the US, firms in which super-rich families have a controlling share are more likely to make a party contribution or to set up a PAC. This part of the results speaks against hypothesis H1. However, when it comes to firm lobbying, firms in which super-rich families have a controlling share are 1.732 times more likely to lobby in Germany. This coefficient is highly significant at the 99.9% confidence level. US firms controlled by the super-rich in contrast do not have higher odds to lobby. Hypothesis H1 therefore finds support for Germany, but not the US. When looking at management involvement, the story is different. If a member of a super-rich family is either in executive or in a supervisory position, the firm has significantly higher odds to make party contributions and to lobby in both countries. The only exception is lobbying in the US when a

super-rich member only sits on an advisory board. In more detail, the correlation seems to be much stronger for Germany than for the US. Party contributions in Germany are more than 7 times more likely if a super-rich family member is involved in any type of management role. The odds ratios are higher in all cases for Germany than the US. In summary, therefore, H1 finds good support both in Germany and the US. However, it depends on the avenue of influence and there are relevant country differences of which avenues are affected by super-rich involvement.

Table 5: Odds ratios estimated from logistic regressions.

	Germany		**USA**	
	Dependent variable		Dependent variable	
	Firm made party contribution?	**Firm lobbied?**	**Firm has PAC?**	**Firm lobbied?**
Log (Closeness to inner circle)	.772	.699	1.819***	1.597***
Firm made a party contribution?		21.075***		80.635***
Firm is controlled by super-rich family?	1.022	1.731***	1.079	.721
Controlling family made a party contribution?		.974		.935
Member of a super-rich family is on a supervisory board?	8.377***	3.301***	1.164	1.031
Member of a super-rich family is on executive board?	7.776***	1.841**	2.171***	1.735***
Constant	334.032**	0.000	326.631***	0.000***
Observations (Firms)	346,655	346,655	733,175	733,175

Notes: Standard errors were clustered at the corporate group level. Control variables are included as suggested by Bayesian Information Criterion (BIC). For full regression tables see appendix 4a. *p<0.05; **p<0.01; ***p<0.001.

Hypothesis H2 suggested that the more direct control super-rich family members have on firm decisions, the more likely a firm is to take political action. To evaluate this hypothesis, Figure 4 plots predicted probabilities of taking firm political action for interactions between super-rich involvement in management, and the family holding a controlling share over the firm. While the

relationships are less linear than expected, overall the results show that a firm has higher odds to take political action when a super-rich family member is in management than simply holding a controlling share. In some cases such as lobbying in Germany, it seems that being on the supervisory board increases the predicted probabilities by more than being an executive. But since the differences between the two forms of management are not significantly different of each other, while those between being involved in management or not are, I conclude that the hypothesis finds some support: the more control the super-rich have over a firm, the more likely it is to take political action.

The third set of hypotheses focuses on the corporate elite inner circle. Results show that in Germany, the inner circle as operationalized here does not play any relevant role for firm political action through the channels analyzed here. Firms in which they are involved are neither more likely to make a party contribution, nor more likely to lobby. Hypotheses H3 can therefore be rejected for the German case. This is also very much in line with the findings presented in Figure 3, that only one firm from the inner circle made a party contribution in Germany, and only six of those firms lobbied. For the US, results are very different: The closer the board of a firm is to the core of the US inner circle network, the higher the odds to have a PAC, as well as to lobby. This gives good support to hypothesis H3 in the US.

The next set of results relates to individual party contributions by super-rich families and individuals. Table 6 presents odds ratios estimated from logistic regressions for the sample of super-rich families, and for the sample of matched super-rich individuals. For Germany, the independent variable whether a family is in the inner circle does not show a statistically significant effect on whether a family member donated or whether an individual donated. H3c is therefore rejected for the German case which gives further evidence against the relevance of an inner circle in Germany. In the US, families in which at least one member is part of the inner circle show 4 times the odds to make a party contribution in comparison to members of other families. This correlation is not existent on the individual level though. There is therefore mixed evidence for hypothesis H3c in the US.

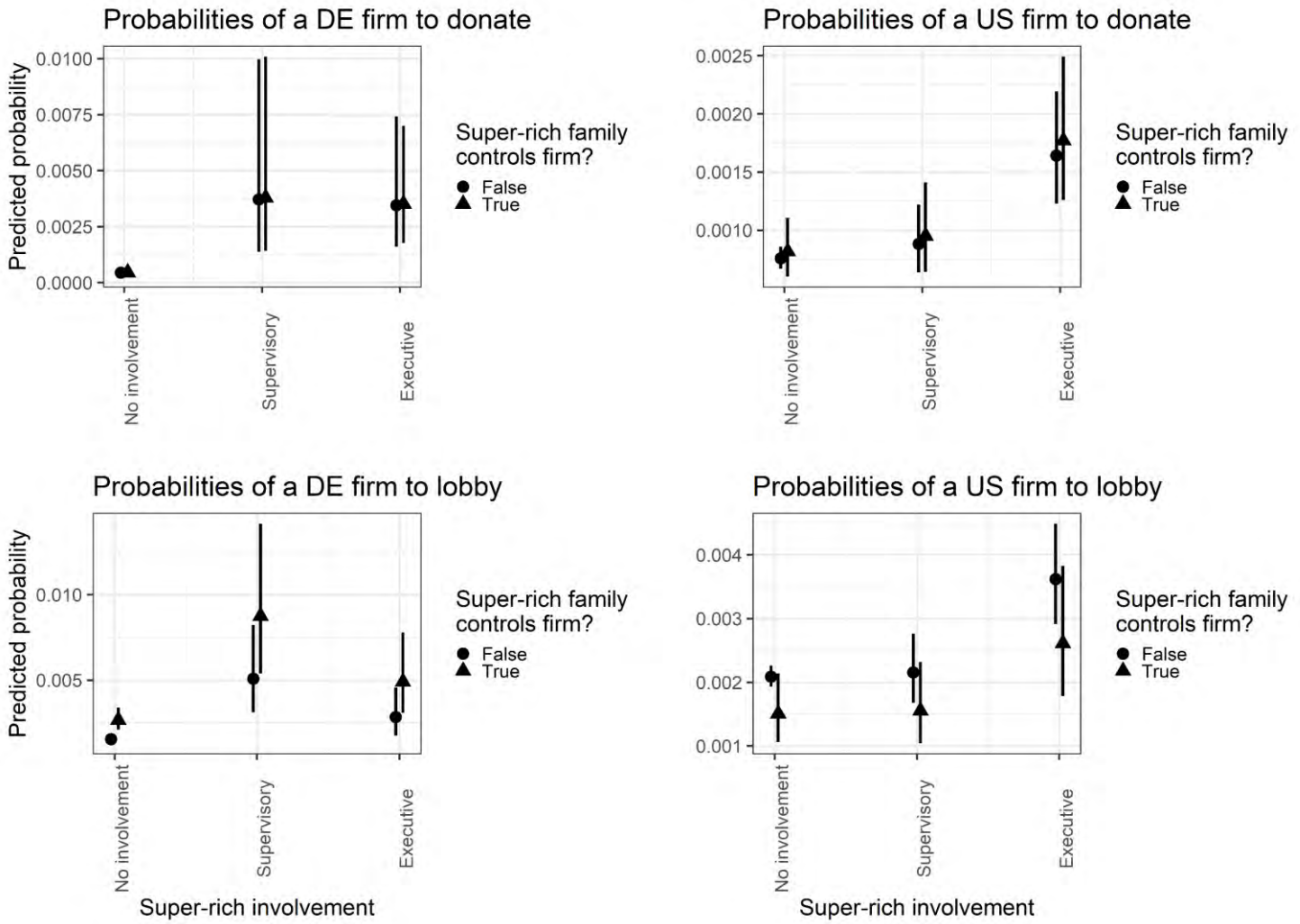


Figure 4: Predicted probabilities of different forms of super-rich involvement in a firm. The plot shows predicted probabilities of interactions between super-rich involvement in management and holding a controlling share of 20% or more.

Table 6: Odds ratios estimated from logistic regressions based on the sample of individuals.

	Dependent variable:			
	Made a party contribution?			
	Germany		USA	
	Family	Individual	Family	Individual
Family level				
Log (estimated family net wealth)	1.337*	.902	1.022	1.054
Log (number of family members)	1.121	.975	1.248***	.837*
Family member is in the inner circle?	3.524	4.500	4.130*	2.174
Individual level				
Female (reference = Male)		.084*		.560**
Age		1.035		1.007
Involvement of individual (reference=Only shareholder)				
Supervisory board of non-owned firm		.575		2.580**
Executive of non-owned firm		0.000		3.616***
Supervisory board of owned firm		1.647		1.943**
Executive of owned firm		1.085		3.633***
Firm level				
Firm political action (reference=No firm took political action)				
Related firm lobbied	2.235	.656	6.353***	1.083
Related firm made party contribution	2.204	0.000	7.322***	1.783
Related firms lobbied and contributed	3.224	5.299**	10.543***	1.598**
Constant	.003***	.015***	.173***	.195***
Observations	1,277	3,464	1,980	1,795

In the US but not in Germany, larger super-rich capitalist families are related to higher odds of a family to make a party contribution. At the family level, this speaks for the relevance of "old wealth" in the US. Yet, this correlation reverses when controlling for individual level factors and in fact, the coefficient is then negative. I conclude that H4a is therefore neither supported for the US, nor for Germany. In other words, I do not find evidence for the relevance of old verses new money for making party contributions. The size of a fortune does only increase the odds to make a party contribution in Germany at the family level. Higher estimated family net wealth increases the odds of a family to contribute. When also controlling for individual factors, family wealth does not seem to explain much variation. The coefficient is also not significant in the US. H5 must therefore be rejected here.

When it comes to political strategies of the super-rich, one can see that especially in the US, a family is clearly more likely to donate if any of the firms they are involved in as controlling shareholders or managers took any firm political action. This is strong support for hypothesis H6c in the US. For Germany, the effect is weaker since there is only a significant coefficient if any of their firms lobbied and made a party contribution at the individual level. Nevertheless, this shows that there is some relation between the different channels and H6c also finds some support in the German case. In the reverse direction, the relationship does however not seem to hold. Table 5 shows that the coefficient for a family having contributed to a party does not significantly increase the odds of a firm to make a party contribution. H6b is therefore not supported in both countries.

How do the super-rich and the corporate elite lean ideologically? To evaluate polarization and ideology of individual and firm party contributions, distributions of the share of donations going to conservative or liberal parties are commonly used in the literature (Bonica 2016; Heerwig 2018). Figure 5 presents the distribution of contributions according to the share they donated to conservative or liberal parties in the US and Germany. As first impressions, it seems that except US PAC contributions all contributions seem very polarized by donating to either liberal or conservative parties only – and not to split between them. For the German case, the inner circle and its political contributions can be neglected. Only one firm with an inner circle member on board could be matched to party contributions. The only result for the political action of the German inner circle therefore is that it overwhelmingly does not pursue any of the analyzed avenues of influence to a relevant extent. For the German case H7a and H8a are therefore rejected. In the US, the inner circle seems to be the most pragmatic group in terms of individual donations. While the overwhelming majority of more than 60% of the 146 identified inner circle members seems to be donating to Democrats only, several members of this group also split donations between the two parties. If anything, inner circle members are less conservative than other businesspeople who could be identified as donors. H7a must therefore be rejected.

The super-rich as individual donors are an overwhelmingly conservative group judged by their party donations in the US and Germany. In Germany, 40 of the 44 super-rich individuals who donated gave 100% of their donations to conservative parties.[9] In the US, the super-rich are the only among the differentiated groups within which a higher share (around 45%) gave 100% to Republicans than the share that gave 100% to Democrats (around 40%). Mean donations of these groups in both countries also clearly lean towards conservative parties. The super-rich seem more polarized between liberal and conservative in the US than in Germany where they are clearly almost exclusively giving to conservative parties. H7b is therefore supported for both countries, and more emphasized for Germany. However, it must not be forgotten that with 40 super-rich individuals only a fraction of the super-rich give at all in Germany, in contrast to 827 matched super-rich individuals in the US. In both countries, donations by super-rich individuals also seem more conservative than their firms which gives support for hypothesis 6a. But surprisingly, the difference especially in Germany is smaller than some of previous findings for the US would suggest.

Firms in which members of the corporate elite are involved seem to behave more pragmatic than other firms. More than 25% of firms with inner circle involvement, the largest share, split their donations equally between Republicans and Democrats in the US. While there is also a relevant share donating to conservative parties, firms with inner circle involvement are on average more pragmatic than individual contributions of this group. H8a therefore finds support. When it comes to firms in which the super-rich hold a controlling share, US PACs actually seem rather normally distributed but with 15%, a relevant share donated 100% to republicans while almost no super-rich controlled firm donated 100% to democrats. While it is not easy to evaluate, in sum PAC contributions of super-rich firms can be regarded as less conservative than individual contributions because a lower share donated 50% or more to democrats. The image points in the same direction but is more emphasized in Germany: Firms clearly donate mostly to conservative parties and only very few give less than 100% to them – although some do. All in all, this speaks against hypotheses H8b that super-rich controlled forms take pragmatic political action: they clearly give more support to conservative parties than to liberal parties.

[9] The four others gave 95%, 88%, 46%, and 0% to conservative parties.

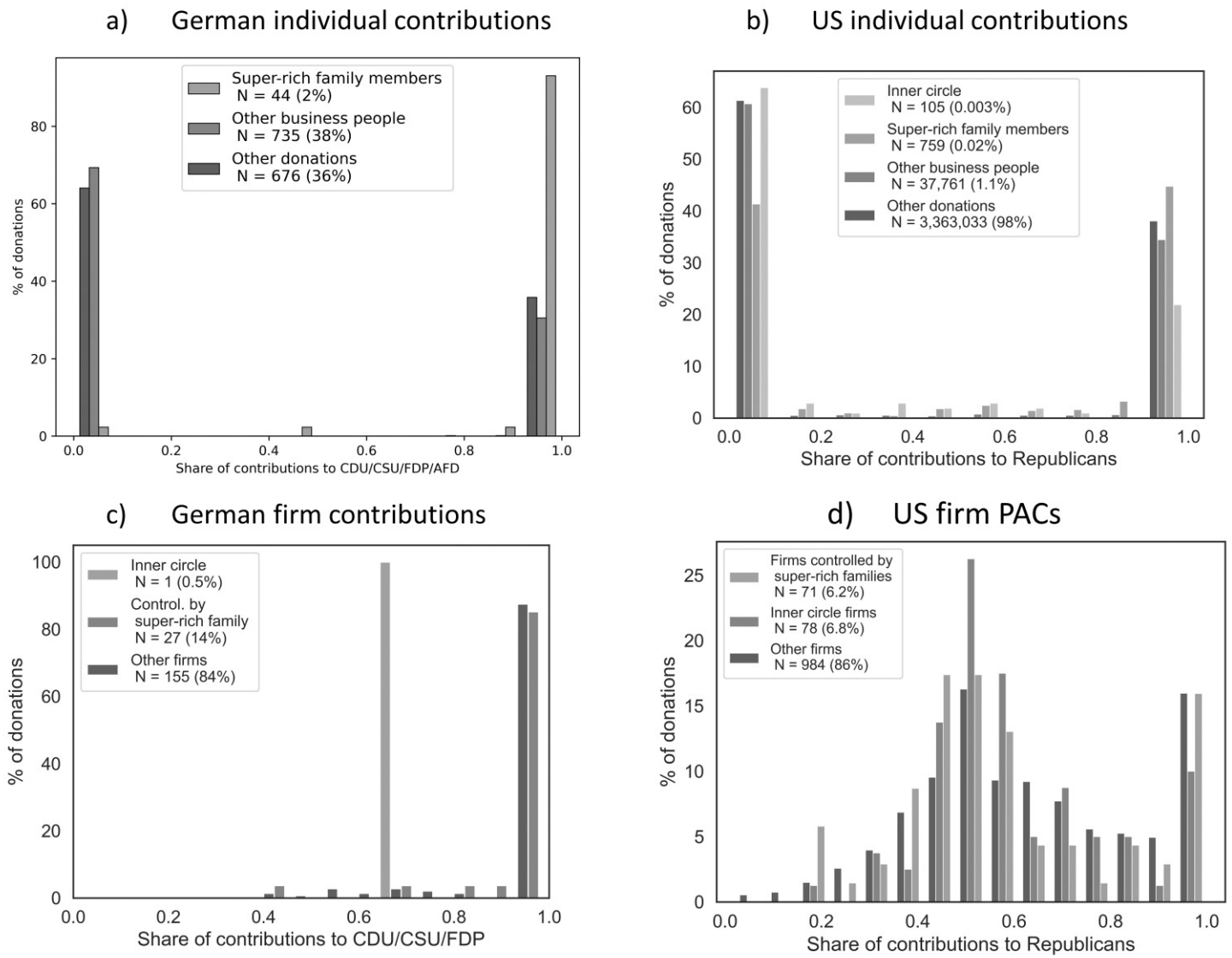


Figure 5: Partisan giving by individuals and firms in Germany (left) and the US (right).

6. Discussion

In summary: what do these results tell us about super-rich and corporate elite direct political action in the US and Germany? This section first gives brief answers to the three research questions, and then discusses alternative avenues of influence, as well as shortcomings, and paths for future research.

How relevant is the magnitude of super-rich and corporate elite money in the political process? The magnitude of super-rich and corporate elite money is moderate to low in both countries, but more relevant in the US than Germany. The super-rich in Germany hardly make any individual party donations, and their overall amount is small in relation to all other donations. The inner circle does not make any party donations in this country at all. It can therefore hardly be said that any of these groups exerts relevant influence through this channel in Germany. While the sum of firm party contributions and lobbying expenditures of firms influenced by these groups is more relevant than individual donations, the total amount contributed or spent on lobbying by firms controlled by these groups is still much smaller than that of other firms. It does seem that the German super-rich are less hesitant to take political action through their firms than in their individual name, but the overall amount is not disproportionally high. In the US, the amount donated and spent on lobbying of firms controlled by the super-rich is quite low. This may well be due to the discussed country difference that shareholdings are more dispersed and there are less controlled firms in the US. Firms in which the inner circle is involved play a much more relevant role compared to super-rich controlled firms and compared to the inner circle in Germany. The corporate elite therefore is a relevant group in the US. However, the absolute amount is by far outweighed by political action of other firms. The most relevant of all avenues of influence in the two countries is likely individual contributions of the richest Americans. With about one fifth of all contributions coming from 827 super-rich individuals their weight is extraordinary. Taken together with the fact that individual party contributions are the most important source of financing for US parties, the largest impact of all analyzed here can be found in this channel in the US.

What explains which of the three avenues of influence the super-rich and the corporate elite use? Both groups use all three avenues of influence as part of political strategies. If a super-rich family member or a member of the inner circle is in executive management of a firm, this firm is much more likely to make a party contribution and to lobby in both countries. This is possibly the most striking and clearest finding concerning this question. Weaker forms of control such as holding a controlling share show only weaker or insignificant correlations. The super-rich are therefore the more likely to use their firms to influence the political process the more control they have over it. That means especially if they are involved in a firm's management. The corporate elite is very relevant in the

US, but negligible in Germany. Political inhibition and disengagement of a corporate elite in politics does therefore seem to be even more prevalent in Germany than was diagnosed for the US in recent years (Mizruchi 2013). The fact that both groups use the three avenues as part of political strategies is underlined by the fact that individual families are more likely to make a party contribution if their firms also lobbied and made a party contribution. In the US, there is some evidence that the interaction of being super-rich and being in the inner circle increases the odds for a rich family to donate. Differences between the two groups and the two countries also underline that it is much more useful to think of the corporate elite and capitalist owners as two distinct groups who act differently – and not as one capitalist class. On top of that, there is variation within these groups in both countries – but there is much more relevant variation in the US since more individuals donate. In the US also, evidence at the family level suggests that old wealth is more likely to make party contributions.

How do the super-rich and the corporate elite lean ideologically? The super-rich are an overwhelmingly conservative group in both countries, the corporate elite in comparison seems much more pragmatic. In Germany, almost all party donations made by the super-rich exclusively go to the conservative parties of FDP and CDU. The far majority of super-rich controlled firms donates more to conservative parties than to liberal parties. It must not be forgotten that we can only say something about a small share of the German super-rich. This group will likely be selective towards being most partisan. However, it seems clear that those of the super-rich who take direct political action in this way in Germany, are a very conservative group. The inner circle does not take political action through these channels in Germany, and we can therefore not say anything about their ideology based on the analysis presented here. In the US also, the super-rich clearly lean towards Republicans and therefore conservativism. While there is also a fraction donating more to Democrats indicating more polarization compared to Germany, the tendency is nevertheless clear. The magnitude of individual contributions, the importance of individual contributions for party financing, together with the clear conservative tendency suggests that the super-rich are an important group to push for political success of the Republicans financially. In contrast to that, the inner circle in the US seems a lot more liberal in terms of their individual contributions, and way more pragmatic, as in giving equal shares to both parties, in terms of firm political action. The corporate elite therefore somewhat already seems to play the role of a balanced and responsible corporate elite which some observers would like them to play more (Mizruchi 2013). One way to interpret this diverging behavior of the two groups in the US could also be that the super-rich are only a class-in-itself, in the Marxist sense, that only realize and act individually upon their narrow interests, hence conservative politics that is more profitable to them. The inner circle on the other hand is more connected e.g., through shared board memberships and therefore realize a broader

class interest through the mechanisms discussed above. This might make them more pragmatic and politically action aiming towards less polarization and more sustainable and balanced business politics.[10] This points to interesting hypotheses for future research, one of them would be why and how the US inner circle forms a class for-itself while in Germany they seem to be completely absent politically in the analyzed channels.

The results presented here shed light on a few important questions about powerful individuals and groups with power in the economy. However, it also raises some interesting questions. Possibly most pressing is the question how it is to be interpreted that the magnitude of super-rich and corporate elite money is moderate to low in both countries. While this question cannot be addressed based on the presented analysis, it is however possible to add some speculations as a basis for future research. First of all, it is helpful to constate that the super-rich and the corporate elite are groups made up from individuals who can be seen as the personifications of capital. They represent large fractions of capital by ownership, by control or by both. Their material interests are intertwined with that of capital accumulation. Why would these groups not influence politics directly?

I see two likely interpretations of the presented results. The first is that the super-rich are politically inactive. This could again either be the case for two reasons. First, because they do not feel responsible and disconnected from a political elite as for example argued by Mizruchi (2013) or the concept of stealth politics (Kantola and Vesa 2022; Page et al. 2018). Second, they could also be inactive because they do not have to be active. We know from the literature and classic debates that sometimes "business" does not even need to take political action because it is structurally clear that their economic success is important for the economy of a country – and therefore also for politicians to be re-elected (Block 1977; Lindblom 1982). It might therefore be that politicians simply act in the interest of capital - and thereby the interest of the super-rich and corporate elite whose material interests are tied with it - without capital taking direct action. This could also be the case at this point in history after capital had won significant battles in the 1970s, as argued by Mizruchi (2013).

The second interpretation is that there might also be other institutionalized avenues through which either coordination among politics and business takes place, or through which the super-rich can take influence. Policy boards, for example, are a common and well-studied example of another avenue of influence that was not studied here (e.g. Comet 2019; Domhoff 2013; Luther-Davies et al. 2022). There will also be exclusive social clubs (Cousin and Chauvin 2014; Domhoff 1975), elite meetings such as the Bilderberg conference (Richardson, Kakabadse, and Kakabadse

[10] Credits go to Joshua Murray for pointing out this alternative explanation.

2011), the world economic forum or other venues where coordination takes place. Similarly, lawmakers interested in realizing the interests of individual firms or business groups may simply overtake the opinions of interest groups such as business from comments to laws (Pagliari and Young 2020; Yackee and Yackee 2006). A final example for a way the interest of business may simply be overtaken into politics is the classic channel of revolving doors between the corporate elite and politics, with former politicians using their contacts and social capital to lobby for business interests in exchange for high salaries (Carboni 2017; Lobbycontrol 2022b).

To assess the relevance of these channels, it is of course crucial to identify the relevant level at which agency takes place. Results presented here suggest that at least direct action of the groups of the super-rich, and the corporate elite does have some - but only limited explanatory power for the influence of business. It might however be that business is simply organized in a different way. Another relevant perspective could be to differentiate more between sectors, instead of individuals with economic power in general. Some literature which already does this is for example that focusing on political action of business associations and think tanks. These may make substantial efforts to foster ideas in favor of business in the public discourse, to enable politicians to act on behalf of it. For Germany, this was shown for example for the *Initiative Neue Soziale Marktwirtschaft* (Kinderman 2017). In the US, there are also well known cases including the direct involvement of very conservative super-rich families (Leonard 2019; Skocpol and Hertel-Fernandez 2016). Finally, there can of course be more hidden channels through which coordination between the corporate elite and super-rich coordinate with politics or simply try to take influence. These can likely only be discovered with more in-depth and qualitative methods. One anecdotal example is the hidden party contribution to right-wing extremist party AfD by German billionaire August von Finck, which could only be uncovered by investigative journalism (Amann, Becker, and Röbel 2018). It becomes clear that alternative avenues of influence are numerous. Although this would require a tremendous workload, a systematic overview of all avenues, and importantly also a theoretical structuration of different levels and which groups and actors are relevant at which levels is much needed in the literature.

Finally, some shortcomings of the present analysis need to be disclosed to evaluate the results presented here. First, the group of the super-rich analyzed is not the full population and there are multiple steps in the genesis of the database used in which attrition takes place. According to Credit Suisse (2021), the population of individuals with 100m USD net financial wealth 2020 was roughly 4,000 individuals in Germany, and 34,000 in the US. Of this population, which is larger than the population with 250m aimed at here, 7,193 individuals (18%) were known from the original list and snowballed from ORBIS. While it is hard to state exact numbers, it is likely that especially for the US, a relevant share of the super-rich is not known by name and could therefore

not be analyzed. Also, in later stages of the data generation process, individuals and their firms could possibly not be matched in ORBIS or they are not included in the data. It could therefore be that the presented numbers are lower bounds and coefficient could be biased due to selection. In addition to selection and data issues, only three avenues of influence could be analyzed here. There will also be other channels of influence that matter and therefore results and conclusions concerning power of the super-rich and the corporate elite only refer to these three avenues. As a final limitation to be discussed here, the inference of individual political attitudes from firm action might be problematic as raised by Burris and Staples (2012:326), and defended by Murray (2017:1633). However, since I am using both individual and firm action and compare the two, I think this objection needs to be kept in mind to evaluate the presented results but is not as pressing.

Lay wisdom tells us that the rich have more power than other citizens. It often remains unclear, however, how this larger influence is supposed to unfold and how it can be measured. At least when looking at direct political action taken by these groups, there is some influence, but its magnitude does not seem sufficient to constate a clear overrepresentation - or even oligarchy - of the super-rich, and the corporate elite. The single exception may be party contribution of the richest Americans. The fact that so many firms unrelated to these two groups spend so much more on political action also speaks against the idea that business acts in a coordinated way, or that many firms are free rider, but rather suggests that firms act independently and likely for their own specific interest. To explain the power of business therefore might require looking deeper into the interrelations between politics and capital – those that do not require direct action by individuals with power in the economy.

I. References

Allen, Michael Patrick. 1981. "Power and Privilege in the Large Corporation: Corporate Control and Managerial Compensation." *American Journal of Sociology* 86(5):1112–23.

Allen, Michael Patrick, and Philip Broyles. 1989. "Class Hegemony and Political Finance: Presidential Campaign Contributions of Wealthy Capitalist Families." *American Sociological Review* 54(2):275–87.

Amann, Melanie, Sven Becker, and Sven Röbel. 2018. "Der Milliardär Und Die AfD." *Der Spiegel*. Retrieved March 23, 2023 (https://www.spiegel.de/politik/august-von-finck-und-die-rechten-der-milliardaer-hinter-der-afd-a-00000000-0002-0001-0000-000160960453).

Arndt, H. Lukas R. 2019. "Varieties of Affluence. How Political Attitudes of the Rich Are Shaped by Income or Wealth." *European Sociological Review* 36(1):136–58.

Banerjee, Tarun, and Joshua Murray. 2020. "Class Dominance or Fracturing? Sources of Broad Interest in Lobbying by Fortune 500 Corporations." *Sociological Perspectives* 64(2):157–75. doi: 10.1177/0731121420921883.

Beckert, Jens. 2022a. "Durable Wealth: Institutions, Mechanisms, and Practices of Wealth Perpetuation." *Annual Review of Sociology* 48(1). doi: 10.1146/ANNUREV-SOC-030320-115024.

Beckert, Jens. 2022b. ""Varieties of Wealth. Towards a Comparative Sociology of Wealth Inequality."" MPIfG, Cologne.

Berghoff, Hartmut. 2006. "The End of Family Business? The Mittelstand and German Capitalism in Transition, 1949-2000." *The Business History Review* 80(2):263–95.

Berle, Adolf A., and Gardiner Means. 1991. *The Modern Corporation and Private Property*. New York: Routledge.

Block, Fred. 1977. "The Ruling Class Does Not Rule: Notes on the Marxist Theory of the State." *Socialist Revolution* 33:6–28.

Bonica, Adam. 2016. "Avenues of Influence: On the Political Expenditures of Corporations and Their Directors and Executives." *Business and Politics* 18(4):367–94. doi: 10.1515/bap-2016-0004.

Bonica, Adam, and Howard Rosenthal. 2015. "The Wealth Elasticity of Political Contributions by the Forbes 400." Retrieved July 3, 2022 (https://papers.ssrn.com/abstract=2668780).

Bornefeld, Andreas. 2019. "Vermögensforschung." *Https://Www.Reichsten.de/En/Index*. Retrieved April 26, 2019 (https://www.reichsten.de/en/index).

Broyles, Philip A. 1993. "The Corporate Elite Goes To Washington: Presidential Campaign Contributions of Corporate Officers." *Social Science Research* 22:72–91. doi: 10.1006/ssre.1993.1004.

Bubola, Emma. 2021. "She's Inheriting Millions. She Wants Her Wealth Taxed Away." *New York Times*. Retrieved March 23, 2023 (https://www.nytimes.com/2022/10/21/world/europe/marlene-engelhorn-wealth-tax.html).

Bureau van Dijk. 2020. "Orbis. Powering the Business of Certainty." Retrieved March 4, 2020 (https://www.bvdinfo.com/en-gb/our-products/data/international/orbis).

Burnham, James. 1975. *The Managerial Revolution*. Westport, Connecticut: Greenwood Press, Publishers.

Burris, Val. 1991. "Director Interlocks and the Political Behavior of Corporations and Corporate Elites." *Social Science Quarterly* 72(3):537–51.

Burris, Val. 2000. "The Myth of Old Money Liberalism: The Politics of the Forbes 400 Richest Americans." *Social Problems* 47(3):360–78. doi: 10.2307/3097235.

Burris, Val. 2001. "The Two Faces of Capital: Corporations and Individual Capitalists as Political Actors." *American Sociological Review* 66(3):361. doi: 10.2307/3088884.

Burris, Val. 2005. "Interlocking Directorates and Political Cohesion among Corporate Elites." *American Journal of Sociology* 111(1):249–83. doi: 10.1086/428817.

Burris, Val, and Clifford L. Staples. 2012. "In Search of a Transnational Capitalist Class: Alternative Methods for Comparing Director Interlocks within and between Nations and Regions." *International Journal of Comparative Sociology* 53(4):323–42. doi: 10.1177/0020715212460256.

Cagé, Julia. 2020. *The Price of Democracy: How Money Shapes Politics and What to Do about It*. Cambridge, MA: Harvard University Press.

Carboni, Marika. 2017. *The Financial Impact of Political Connections*. Cham: Springer International Publishing.

Chandler, Alfred D. Jr. 1977. *The Visible Hand*. Cambridge and London: Belknap Press of Harvard University Press.

Chu, Johan S. G., and Gerald F. Davis. 2016. "Who Killed the Inner Circle? The End of the Era of the Corporate Interlock Network." *American Journal of Sociology* 122(3):714–54. doi: 10.2139/ssrn.2061113.

Clawson, Dan, and Alan Neustadtl. 1989. "Interlocks , PACs , and Corporate Conservatism Author." *American Journal of Sociology* 94(4):749–73.

Comet, Catherine. 2019. "How Does the Inner Circle Shape the Policy-Planning Network in France?" *Socio-Economic Review* 17(4):1021–41. doi: 10.1093/SER/MWX051.

Cousin, Bruno, and Sébastien Chauvin. 2014. "Globalizing Forms of Elite Sociability: Varieties of Cosmopolitanism in Paris Social Clubs." *Ethnic and Racial Studies* 37(12):2209–25. doi: 10.1080/01419870.2014.934260.

Credit Suisse. 2021. "Global Wealth Report 2021."

Dahrendorf, Ralf. 1959. *Class and Class Conflict in Industrial Society*. Stanford: Stanford University Press.

De La Cruz, A., A. Medina, and Y. Tang. 2019. *Owners of the World's Listed Companies*. Paris: OECD.

Dedupe.io. 2022. "Dedupe." Retrieved March 20, 2023 (https://github.com/dedupeio/dedupe).

Deutscher Bundestag. 2022. "Lobbyregister." Retrieved March 23, 2023 (https://www.lobbyregister.bundestag.de/startseite).

Domhoff, G. William. 1967. *Who Rules America?* Englewood Cliffs N.J.: Prentice Hall.

Domhoff, G. William. 1975. *The Bohemian Grove and Other Retreats. A Study in Ruling-Class Cohesiveness*. New York, Hagerstown, San Francisco, London: Harper & Row.

Domhoff, G. William. 2013. *Who Rules America?* Boston: McGraw-Hill.

Fink, Alexander. 2017. "Donations to Political Parties: Investing Corporations and Consuming Individuals?" *Kyklos* 70(2):220–55.

Forbes. 2022. "Forbes Billionaires List. The Richest in 2022."

Freeman, Tom C., Sebastian Horsewell, Anirudh Patir, Josh Harling-Lee, Tim Regan, Barbara B. Shih, James Prendergast, David A. Hume, and Tim Angus. 2022. "Graphia: A Platform for

the Graph-Based Visualisation and Analysis of High Dimensional Data." *PLOS Computational Biology* 7(18). doi: https://doi.org/10.1371/journal.pcbi.1010310.

Friese, Ulrich. 2021. "Pharma-Erbe Als Politischer Investor Bei Den Grünen." Retrieved March 23, 2023 (https://www.faz.net/aktuell/wirtschaft/unternehmen/antonis-schwarz-als-investor-engagiert-sich-der-pharma-erbe-bei-den-gruenen-17400619.html).

Gajek, Eva Maria. 2016. "Reichtum Und Reiche in Der Bundesrepublik Der 1960er Jahre. Eine Dialektik von Sichtbarkeit Und Unsichtbarkeit." *Werkstatt Geschichte* 73:51–70.

Gnirke, Kristina, Alexander Kühn, and Simone Salden. 2021. "Drogerie-Erben Roßmann Und Werner Im SPIEGEL-Gespräch. Wer Hat Es Leichter – Der Unternehmensgründer Oder Sein Sohn?" *Spiegel*. Retrieved March 23, 2023 (https://www.spiegel.de/wirtschaft/unternehmen/drogerie-unternehmer-christoph-werner-dm-und-raoul-rossmann-ich-fuehre-das-unternehmen-nicht-als-sohn-a-d507ce8d-fa77-4ae8-8704-66d27f7a14dd).

Goerres, Achim, and Martin Höpner. 2014. "Polarizers or Landscape Groomers? An Empirical Analysis of Party Donations by the 100 Largest German Companies in 1984–2005." *Socio-Economic Review* 12:517–44. doi: 10.1093/ser/mwt027.

Gourevitch, Peter A., and James Shinn. 2005. *Political Power and Corporate Control: The New Global Politics of Corporate Governance*. Princeton, NJ: Princeton University Press.

Heemskerk, Eelke M., and Frank W. Takes. 2016. "The Corporate Elite Community Structure of Global Capitalism." *New Political Economy* 21(1):90–118. doi: 10.1080/13563467.2015.1041483.

Heerwig, Jennifer A. 2018. "Money in the Middle: Contribution Strategies among Affluent Donors to Federal Elections, 1980–2008." *American Journal of Sociology* 123(4):1004–63. doi: 10.1086/694888.

Heerwig, Jennifer A., and Joshua Murray. 2019. "The Political Strategies and Unity of the American Corporate Inner Circle: Evidence from Political Donations, 1982–2000." *Social Problems* 66(4):580–608. doi: 10.1093/socpro/spy014.

Hilferding, Rudolf. 1955. *Das Finanzkapital. Eine Studie Über Die Jüngste Entwicklung Des Kapitalismus*. Berlin: Dietz Verlag.

Huijzer, M. Jouke, and Eelke M. Heemskerk. 2021. "Delineating the Corporate Elite: Inquiring the Boundaries and Composition of Interlocking Directorate Networks." *Global Networks* 21(4):791–820. doi: 10.1111/glob.12316.

James, David R., and Michael Soref. 1981. "Profit Constraints on Managerial Autonomy: Managerial Theory and the Unmaking of the Corporation President." *American Sociological Review* 46(1):1–18. doi: 10.2307/2095023.

Jensen, Michael C., and William H. Meckling. 1976. "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure." *Journal of Financial Economics* 3(4):305–60. doi: 10.1016/0304-405x(76)90026-x.

Kantola, Anu, and Juho Vesa. 2022. "Silence of the Wealthy: How the Wealthiest 0.1% Avoid the Media and Resort to Hidden Strategies of Advocacy." *European Journal of Communication* 02673231221105137. doi: 10.1177/02673231221105137.

Kapalschinski, Christoph. 2019. "Verena Bahlsen Sorgt Mit Äußerung Über Zwangsarbeit in NS-Zeit Für Empörung." *Handelsblatt*. Retrieved November 11, 2022 (https://www.handelsblatt.com/unternehmen/handel-konsumgueter/keks-dynastie-erbin-verena-bahlsen-sorgt-mit-aeusserung-ueber-zwangsarbeit-in-ns-zeit-fuer-empoerung/24335912.html).

Keister, Lisa A., Mingxiang Li, and Hang Young Lee. 2021. "Do You Need Business Assets to Be Rich?" *Socius: Sociological Research for a Dynamic World* 7:1–3. doi: 10.1177/23780231211031684.

Kim, In Song, Jan Stuckatz, and Lukas Wolters. 2020. "Strategic and Sequential Links between Campaign: Donations and Lobbying."

Kinderman, Daniel. 2017. "Challenging Varieties of Capitalism's Account of Business Interests: Neoliberal Think-Tanks, Discourse as a Power Resource and Employers' Quest for Liberalization in Germany and Sweden." *Socio-Economic Review* 15(3):587–613. doi: 10.1093/ser/mww040.

Larsen, Anton Grau, and Christoph Houman Ellersgaard. 2017. "Identifying Power Elites—k-Cores in Heterogeneous Affiliation Networks." *Social Networks* 50:55–69. doi: 10.1016/j.socnet.2017.03.009.

Larsen, Anton Grau, and Christoph Houman Ellersgaard. 2018. "The Inner Circle Revisited: The Case of an Egalitarian Society." *Socio-Economic Review* 16(2):251–75. doi: 10.1093/ser/mwx052.

Lehrer, Mark, and Sokol Celo. 2016. "German Family Capitalism in the 21st Century: Patient Capital between Bifurcation and Symbiosis." *Socio-Economic Review* 14(4):729–50. doi: 10.1093/ser/mww023.

Leonard, Christopher. 2019. *Kochland: The Secret History of Koch Industries and Corporate Power in America*. New York: Simon & Schuster.

Lindblom, Charles E. 1982. "The Market as Prison." *Journal of Politics* 44(2):324–36.

Lobbycontrol. 2022a. "Lobbypedia." Retrieved March 23, 2023 (https://lobbypedia.de/wiki/Spezial:Abfrage_ausf%C3%BChren/Parteispenden).

Lobbycontrol. 2022b. "Seitenwechsel Im Überblick." Retrieved November 22, 2022 (https://lobbypedia.de/wiki/Seitenwechsler_in_Deutschland_im_%C3%9Cberblick).

Luther-Davies, Philip, Kasia Julia Doniec, Joseph P. Lavallee, Lawrence P. King, and G. William Domhoff. 2022. "Corporate Political Power and US Foreign Policy, 1981–2002: The Role of the Policy-Planning Network." *Theory and Society* 51(4):629–52. doi: 10.1007/s11186-022-09475-3.

Marx, Karl. 1887. *Capital. A Critique of Political Economy. Volume I. Book One: The Process of Production of Capital*. Moscow, USSR: Progress Publishers.

Mizruchi, Mark S. 1992. *The Structure of Corporate Political Action : Interfirm Relations and Their Consequences*. Cambridge: Harvard University Press.

Mizruchi, Mark S. 2004. "Berle and Means Revisited: The Governance and Power of Large U.S. Corporations." *Theory and Society* 33:579–617.

Mizruchi, Mark S. 2013. *The Fracturing of the American Corporate Elite*. Cambridge: Harvard University Press.

Mizruchi, Mark S., and Deborah M. Bey. 2003. "Corporate Control, Interfirm Relations, and Corporate Power." Pp. 310–30 in *The Handbook of Political Sociology*, edited by T. Janoski, R. R. Alford, A. M. Hicks, and M. A. Schwartz. Cambridge: Cambridge University Press.

Mosca, Gaetano. 1939. *The Ruling Class*. McGraw-Hill Book Company.

Murray, Joshua. 2014. "Evidence of a Transnational Capitalist Class-for-Itself: The Determinants of PAC Activity among Foreign Firms in the Global Fortune 500, 2000-2006." *Global Networks* 14(2):230–50. doi: 10.1111/glob.12037.

Murray, Joshua. 2017. "Interlock Globally, Act Domestically: Corporate Political Unity in the 21st Century." *American Journal of Sociology* 122(6):1617–63. doi: 10.1086/691603.

Neßhöver, Christoph, and Andreas Bornefeld. 2018. "Die 1001 Größten Vermögen [The 1001 Largest Fortunes]." *Manager Magazin*. Retrieved March 7, 2019 (https://www.manager-magazin.de/premium/reichste-deutsche-liste-der-1001-reichsten-a-00000000-0002-0001-0000-000159721570).

OpenSecrets. 2022. "Bulk Data." Retrieved November 8, 2022 (https://www.opensecrets.org/bulk-data).

Page, Benjamin I., Larry M. Bartels, and Jason Seawright. 2013. "Democracy and the Policy Preferences of Wealthy Americans." *Perspectives on Politics* 11(1):51–73. doi: 10.1017/S153759271200360X.

Page, Benjamin I., Jason Seawright, and Matthew J. Lacombe. 2018. *Billionaires and Stealth Politics*. Chicago and London: The University of Chicago Press.

Pagliari, Stefano, and Kevin L. Young. 2020. "Exploring Information Exchange among Interest Groups: A Text-Reuse Approach." *Journal of European Public Policy* 27(11):1698–1717. doi: 10.1080/13501763.2020.1817132.

Piketty, Thomas. 2014. *Capital in the Twenty-First Century*. Cambridge: Belknap Press of Harvard Univ. Press.

Richardson, Ian, Andrew Kakabadse, and Nada Kakabadse. 2011. *Bilderberg People: Elite Power and Consensus in World Affairs*. London and New York: Routledge.

Rühle, Alex. 2010. "Aldi Brüder. Reich, Reicher, Unsichtbar." *Süddeutsche Zeitung*. Retrieved November 11, 2022 (https://www.sueddeutsche.de/politik/aldi-brueder-reich-reicher-unsichtbar-1.786283).

Scott, John, ed. 1990. *The Sociology of Elites*. Vol. I–III. Aldershot, UK; Brookfield, USA: Edward Elgar.

Scott, John. 1997. *Corporate Business and Capitalist Classes*. Oxford; New York: Oxford University Press.

Scott, John. 2008. "Modes of Power and the Re-Conceptualization of Elites." *The Sociological Review* 56(1_suppl:26–43. doi: https://doi.org/10.1111/j.1467-954X.2008.00760.x.

Skocpol, Theda, and Alexander Hertel-Fernandez. 2016. "The Koch Network and Republican Party Extremism." *Perspectives on Politics* 14(3):681–99. doi: 10.1017/S1537592716001122.

Stamm, Isabell K. 2016. "Coordination Tasks and Negotiation Modes of Linked Lives in Entrepreneurial Families." *Journal of Marriage and Family* 78(4):939–56. doi: 10.1111/JOMF.12304.

Stokman, Frans N., Rolf Ziegler, and John Scott, eds. 1985. *Networks of Corporate Power a Comparative Analysis of Ten Countries*. 1. publ. Cambridge: Polity Press.

Sweezy, Paul M. 1951. "The American Ruling Class. Part I." *Monthly Review* 3(1):10–17.

Useem, Michael. 1984. *The Inner Circle: Large Corporations and the Rise of Business Political Activity in the U.S. and U.K.* Oxford; New York: Oxford University Press.

Veblen, Thorstein. 1899. *The Theory of the Leisure Class: An Economic Study in the Evolution of Institutions*. New York: Macmillan.

Wealth-X. 2022. "World Ultra Wealth Report 2022." Retrieved March 21, 2023 (https://www.wealthx.com/report/world-ultra-wealth-report-2022/).

Windolf, Paul. 2002. *Corporate Networks in Europe and the United States*. Oxford; New York: Oxford University Press.

Winters, Jeffrey A., and Benjamin I. Page. 2009. "Oligarchy in the United States?" *Perspectives on Politics* 7(4):731–51. doi: DOI: 10.1017/S1537592709991770.

Würzer, Julian. 2019. "Lidl-Gründer Dieter Schwarz: Der Verborgene Milliardär Wird 80." *Augsburger Allgemeine*. Retrieved March 23, 2023 (https://www.augsburger-allgemeine.de/wirtschaft/Portraet-Lidl-Gruender-Dieter-Schwarz-Der-verborgene-Milliardaer-wird-80-id55500996.html).

Yackee, Jason Webb, and Susan Webb Yackee. 2006. "A Bias Towards Business? Assessing Interest Group Influence on the U.S. Bureaucracy." *The Journal of Politics* 68(1):128–39. doi: 10.1111/j.1468-2508.2006.00375.x.

Zeitlin, Maurice. 1989. *The Large Corporation and Contemporary Classes*. New Brunswick, New Jersey: Rutgers University Press.

4th Article:

Finance Capital 2.0?

Global Varieties of Shareholder Concentration

Lukas Arndt[1,2], Benjamin Braun[1] and Sebastian Kohl[3]

[1] Max Planck Institute for the Study of Societies

[2] Sciences Po Paris

[3] John F. Kennedy Institute, Free University Berlin

Abstract

The structure of share ownership is a key variable for the comparative study of capitalism. The canonical distinction is between concentrated shareholder structures dominated by families and non-financial companies, and dispersed structures dominated by institutional capital pools. However, this distinction has been upended by the rise of giant asset managers—an institutional source of shareholder concentration. At the same time, the rise of extreme wealth inequality points to increased shareholder concentration also amongst traditional blockholders, individuals and families. This paper marshals global, firm-level share ownership data to study these global varieties of shareholder concentration. In addition to finding support for the dual concentration hypothesis, we also offer a detailed quantitative overview of who owns what in the global corporate economy.

Submission status:

To be submitted after a final round of comments by colleagues (journal to be determined).

1 Introduction

Students of capitalism tend to care deeply about the structure of share ownership—the best measurable indicator for the distribution of power over the means of production.[1] In comparative corporate governance as well as in comparative political economy (CPE), the canonical distinction is between concentrated and dispersed shareholder structures (Hall and Soskice, 2001; Amable, 2003; Aguilera and Jackson, 2003). In coordinated market economies, concentrated shareholder structures with dominant financial institutions and extensive cross-shareholdings protected firms from short-term capital market pressures; whereas dispersed shareholder structures in liberal market economies exposed firms to hostile takeovers. This variation in share ownership structures aligned with differences in labor market institutions and innovation systems, as well as with different stakeholder coalitions between shareholders, managers, and workers (Höpner, 2003; Gourevitch and Shinn, 2005). Until recently, the expectation in the comparative literature on governance was that, following the dissolution of the dense networks of cross-shareholdings in many coordinated market economies, advanced economies were on a path towards convergence with the shareholder structure of the United States (Aguilera and Jackson, 2010; Goyer, 2011; Callaghan, 2018). The observation that "ownership of large listed companies is dispersed [...] in the U.S. and concentrated in most other countries" was considered "one of the best established stylized facts" in the corporate finance literature (Franks, Mayer and Rossi, 2009, 4009).

The purpose of this paper is to challenge this stylized fact. Since the turn of the century, the shareholding landscape has undergone a radical transformation, driven by the rapid growth of institutional capital pools. This category, comprises both asset owners—such as pension funds, endowments, or insurers—and asset managers, primarily providers of mutual funds and index funds (Fichtner, Heemskerk and Garcia-Bernardo, 2017; Braun, 2021). Their growth has been fueled, in turn, by the expansion of funded

[1] Mindful of the critique of the ideologically charged and legally misleading description of shareholders as 'owners' of the corporation (Stout, 2012; Robé, 2011), we consistently avoid the term 'corporate ownership' and instead refer to 'share ownership' and 'shareholder structure'. Similarly, Palladino (2022) has cautioned against the use of 'investors' as a synonym for 'shareholders', which is why we consistently stick to the latter term.

pension systems. Today, these institutional capital pools have become a source of shareholder concentration—in stark contrast to their customary association, in the CPE literature, with dispersed share ownership. At the same time, the rise of extreme wealth inequality points to increased shareholder concentration also amongst individuals and families, the traditional blockholders. As several authors have observed, this configuration resembles the "finance capital" configuration diagnosed by Rudolf Hilferding, under which large banking conglomerates, together with a small number of industrial tycoons (in the United States: 'robber barons') dominated the US and German corporate sectors at the turn of the 19th century (Davis, 2008; Maher and Aquanno, 2022; Braun, 2022). Are we living through the emergence of a finance capital 2.0 configuration? To examine this question, we marshal the Bureau van Dijk's *ORBIS* dataset, which provides global, firm-level information on share ownership for both listed and unlisted companies. ORBIS is notoriously patchy and difficult to handle for research purposes (Garcia-Bernardo and Takes, 2018; Kalemli-Ozcan et al., 2022; Liu, 2020; Bajgar et al., 2020). This paper is therefore based on a dataset that takes into account the state of the art of the ORBIS literature and extends the data in several regards to make it usable for our analysis of ownership concentration. Among other data preparation steps, detailed in section 4, we have assembled a novel, more reliable dataset of institutional shareholder categories, as well as a dataset of super-rich families with net wealth in excess of USD 250 million.

Our descriptive results show that liberal market economies have switched from quintessential dispersed-ownership societies to concentrated shareholder structures. The dominance of foreign shareholders, meanwhile, is most extreme in continental Europe—largely because the centers of the global asset management sector lie outside of Europe. Across the advanced world, the effect of controlling stakes held by super-rich individuals on the size of minority holdings by diversified asset managers is neutral. Direct holdings by institutional asset owners, such as pension funds, are minimal, while having large funded pension systems does not afford countries domestic shareholder majorities.

A principal component analysis shows that ownership concentration comes in at least four different varieties (finance-dominated English-speaking countries; foreign-dominated,

small and open economies; rich oligarchies; and developmental economies). We uncover that only a minority of variance plays at the country level (when compared to the firm- or sector level). Common-Law regimes are associated with all kinds of institutional ownership, particularly asset manager capitalism, Scandinavians mostly with banks and pension funds, and Roman-law systems mostly with public ownership. Super-rich families and individual corporate owners are also systematically different from asset managers (and other institutional shareholders) because their assets are much more concentrated, have a stronger focus on manufacturing and carbon-emission-intensive sectors.

In theoretical terms, our results challenge the 'Gerschenkronian' assumption, widespread in CPE, that shareholder structure is determined by national developmental paths and the composition and financing needs of the non-financial sector (Gerschenkron, 1962; Zysman, 1983; Hall and Soskice, 2001). Instead, we advance a finance-centered explanation, according to which shareholder structure is increasingly a function of developments internal to the financial system in general, and of the growth of institutional capital pools in particular.

2 Share ownership in CPE: Three problems

The canonical varieties-of-capitalism distinction between coordinated market economies with concentrated shareholder structures and liberal market economies with dispersed shareholders aligned neatly with differences in labor market institutions and innovation systems (Hall and Soskice, 2001; Amable, 2003). Although insightful and conceptually productive at the time, this distinction introduced three enduring problems into CPE scholarship on shareholder structures. First, they mistook a snapshot view for a stable equilibrium (Streeck, 2009). While scholars were well aware that concentrated shareholder structures encountered in coordinated market economies were being eroded by the late 1990s, dispersed structures were—at least implicitly—seen as a stable characteristic of liberal market economies. If anything, this view has been reinforced by the more recent work of corporate finance scholars, which has shown shareholder structures to converge

towards greater holdings by foreign institutional shareholders (Franks and Mayer, 2017; Aminadav and Papaioannou, 2020). Overall, this story was consistent with the static view of the dominant theory in the law-and-finance literature, which was centered on minority shareholder protection and legal origins. Whereas common-law countries provided strong protections for minority shareholders, shareholders in civil-law countries had to protect themselves by holding large blocks of shares to counterbalance the power of corporate management (La Porta et al., 1998; La Porta, Lopez-de-Silanes and Shleifer, 1999). A quarter century later, minority shareholder protections have converged globally towards the high level of common-law jurisdictions (Katelouzou and Siems, 2015), while the size of share blocks held by institutional capital pools appears to have *increased* everywhere. Indeed, given what we know about global asset managers' geographical footprint, we expect that share blocks larger than 5 percent are more prevalent in common-law jurisdictions.

Second, and more fundamentally, the CPE literature reflects a Gerschenkronian theory of finance as a second-order sector that serves the non-financial sector. Although the state plays an important role in mediating this relationship, it is the financing needs of firms that determine whether a country develops a market-based or a bank-based financial system (Zysman, 1983). Against this 'financing view', we emphasize the financial sector's asset management function. Marxist scholars of finance have long focused on the accumulation of financial balances by the wealthy, and thus on the tendency for claims on the corporate sector to accumulate in the financial sector (Hilferding, 1985; Aglietta, 1979). Indeed, several authors have highlighted the apparent return of the "finance capital" configuration, with asset managers rather than banks now the dominant actors (Davis, 2008; Maher and Aquanno, 2022; Braun, 2022). Theoretically, this 'asset management view' shifts the source of variation from the relationship between financial and non-financial firms to developments *within the investment chain* (Braun, 2016).[2] From this perspective, the growth of institutional capital pools should impact shareholder structures everywhere, and in ways that have relatively little to do with developments in

[2]As Mark Roe Roe (1994, xv) has noted, corporate governance is "partly just the tail to the larger kite of the organization of savings."

corporate finance.

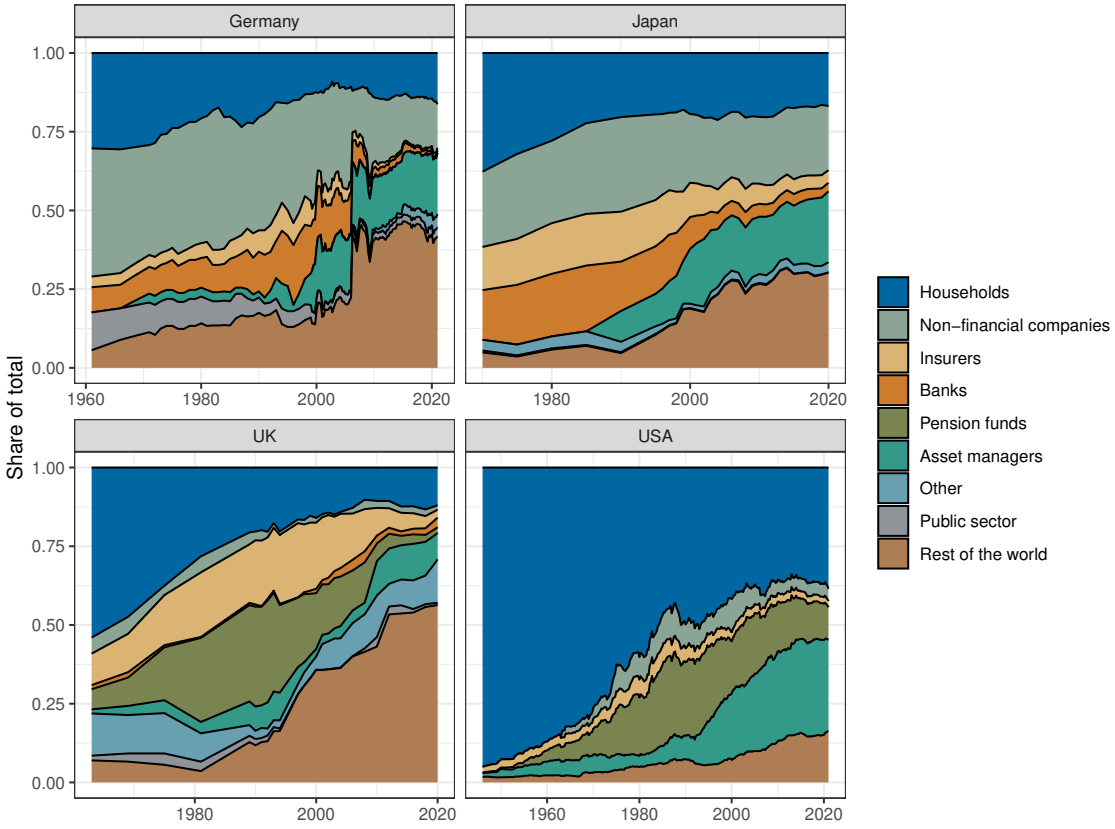
The third theoretical problem is closely related—the focus on the nation-state as the unit of analysis. The approach is arguably a child of the Bretton Woods period, characterized by financial repression and greatly reduced transnational financial activity. Even during the first era of financial globalization, however, the finance capital configuration studied by Hilferding was organized primarily within national borders. Under conditions of contemporary global financial integration, however, methodological nationalism has outlived its purpose. The dominance of the United States in the global asset management sector (Haberly and Wójcik, 2022) makes it so that US asset managers are overrepresented in shareholder registers across the world, relative to the holdings belonging to beneficial owners located in the US. For instance, with roughly 40 percent of its assets managed for clients outside of the United States, BlackRock is one of the largest managers of European money (BlackRock, 2022, 10). This has consequences for the economic geography of shareholder structures. Whereas asset manager capitalism is synonymous with a rising foreign share for all other countries, the share of foreign shareholders of US companies remains low even as the dominance of asset managers increases.

3 From aggregate to firm-level data

How has share ownership developed over the past half-century? Figure 1 combines data from various sources and aggregates shareholder categories so as to render them commensurable across the two quintessential coordinated (Germany and Japan) and liberal (UK and US) market economies. The two most visible changes both cut across the CME-LME distinction. The first is the rise of foreign shareholdings to around 50% in Germany and UK, and just over 25% in Japan. By contrast, the foreign share remains much lower in the United States. The second change is the decline of shares held directly by households and non-financial companies, and the rise of holdings by institutional capital pools, which comprise both asset owners, such as pension funds and insurers, and asset managers. However, the data understates the size of institutional capital pools, much of which

is hidden in the category 'Rest of the world', in the form of foreign (mainly: US) asset managers. Thus, while Figure 1 points to common trends across the four countries, the aggregate nature of the data tells us little about the concentration of shareholdings, or about the prevalence of blockholdings. It also obscures differences between different types of firms or sectors, as well as between the portfolios of different categories of institutional capital pools.

Figure 1: Shareholder structures of listed companies in the long run



Data: Germany: Fohlin (2005, 233), European Commission, Eurostat; Japan: Japan Exchange Group; UK: ONS; USA: Federal Reserve, Financial Accounts.

In order to study these questions, firm-level data is needed. For some years now, large-scale commercial databases enable to examine the shareholder structure at the micro level, meaning at the level of the individual company and the individual shareholder. For studies with a comparative or global scope, ORBIS, provided by Bureau van Dijk, has been the dataset of choice. The pioneering study of global share ownership was by Vitali, Glattfelder and Battiston (2011, 4), who found that the network of corporate control was "highly concentrated in the hands of few top holders", and that corporate control

was "much more unequally distributed than wealth". Among the top 50 holders in the control network, all but two were financial companies, of which half were based in the US. First ranked was Barclays, whose asset management arm was acquired by BlackRock in 2009 (Wigglesworth, 2021). This acquisition consolidated what Fichtner, Heemskerk and Garcia-Bernardo (2017) called the "hidden power of the Big-3"—the dominant position in the global shareholder landscape of the world's three largest asset managers, BlackRock, Vanguard, and State Street.

From the investment-chain-centered perspective outlined above, patterns of shareholder concentration should increasingly be shaped by the growth of institutional capital pools in general, and by the rise of index-tracking asset managers in particular (Braun, 2016; Petry, Fichtner and Heemskerk, 2019). In 2020, membership in the German DAX, M-DAX, and S-DAX was associated with mean holdings by BlackRock and Vanguard of 3, 2, and 1%, respectively (Steuer, 2022, 89). This suggests that inclusion in a country's blue-chip index—a variable unrelated to the business model or financing situation of individual companies—should be a strong predictor of the size of the holdings of the world's largest asset managers. We, therefore, expect that holdings of institutional shareholders, which in the CPE literature have long been associated with stock ownership *dispersion*, have become drivers of stock ownership *concentration* (Braun, 2021). This hypothesis goes against the "stylized fact" of high share ownership dispersion in the United States (Franks, Mayer and Rossi, 2009, 4009), which the work of Aminadav and Papaioannou (2020, 1194) has very recently re-affirmed.

4 Data and methods

To analyze share ownership regimes in the OECD, we use data from the ORBIS database (Bureau van Dijk, 2020). For 2022, ORBIS provides data on more than 400 million firms, including balance sheet information and other financials, data on management, as well as detailed ownership data for many firms in the sample. We use yearly snapshots of the data for the month of June for all years from 2010 to 2021 which we downloaded

from the ORBIS portal. While this data is, to our knowledge, the best choice for global analyses in terms of geographical coverage and the number of firms covered, there are several problems with the ORBIS database specifically (Garcia-Bernardo and Takes, 2018; Kalemli-Ozcan et al., 2022; Liu, 2020; Bajgar et al., 2020), and with such big off-the-shelf datasets in general (Heemskerk et al., 2018). Arndt (2023) gives an overview of known and remaining problems with the ORBIS data. We are confident to have overcome several of these issues in our data preparation steps. However, unresolved problems remain, and in the following we summarize the choices we made in preparing the data and selecting the sample, and briefly discuss the main unresolved issues.

4.1 Data preparation

Our data preparation follows best practice guidelines for ORBIS and addresses the major known problems, including BvD identifiers that change over time, the reporting lag, duplicates, missing values, and incorrectly assigned shareholder types.

Ownership data and shares

ORBIS provides direct ownership shares as well as total shares, which are supposed to include all indirectly held shareholdings. We use the total shares of ownership as a basis and interpolate gaps of missing data between reported values. There are issues with missing values over time for which our solution is to linearly interpolate ownership shares from 2010 to 2021, for a maximum gap of up to 2 years. From our assessment, there are several cases in which direct and total shares are mixed up. In these cases, only one of the two is provided and we circumvent this issue by complementing total shares after interpolation of gaps with direct shares.

Multilevel data: Firms and corporate groups

Firms in ORBIS may be from all levels of complex corporate hierarchies. For example, they may be national subsidiaries with a specific purpose for the whole corporate group (Reurink and Garcia-Bernardo, 2020), or they may be the top holding company. In

ORBIS, the latter are called Global Ultimate Owners (GUO) which may be firms or individuals. Unfortunately, we could only download the 2022 cross-sectional GUO data, and therefore corporate group membership as well as mergers and acquisitions at this point in time. For listed firms of which we analyze shareholders, we keep subsidiaries as well as GUOs if both are listed independently. For shareholding firms, we only keep total shares by GUOs to avoid double counting of direct and indirect shareholdings. The GUO variable reported by BvD is known to be erroneous and in some cases to not actually reflect the correct global ultimate owner. This is a bias we need to acknowledge but could not resolve. Some scholars have tried to achieve this by manually coding ultimate owners (Aminadav and Papaioannou, 2020).

Deduplication of firms and shareholders

Duplicate firms are a known problem in ORBIS (e.g., Garcia-Bernardo and Takes, 2018) and we discovered that the problem is at least as pressing for shareholders that are not firms.[3] Duplicates may for example be caused by different data providers contributing to the database. Although an external data engineering company apparently removed a share of such duplicates in ORBIS, relevant numbers remain in the database. We further remove duplicates in three ways. First, we check for simple duplicate firm names. Second, we apply the procedure suggested by (Garcia-Bernardo and Takes, 2018) based on the network topology of the network of interlocking directorates. Third, we trained a supervised learning deduplication algorithm based on firm names and other properties with the DeDupe library in Python and aggregated 186,737 duplicate shareholders of all listed and non-listed firms [4].

Missing values and interpolation

A reporting lag of 1-2 years in ORBIS is known from the literature (e.g. Kalemli-Ozcan et al., 2022, p. 4) which further introduces inevitable bias in those of our results that

[3]Another type of duplicate is duplicate financial reports of the same firm, as discussed in depth in (Kalemli-Ozcan et al., 2022).

[4]This step was not performed for Chinese shareholders due to ambiguous names in ORBIS translated from Mandarin.

depend on market capitalization. We take a simple and coarse approach to tackle this problem by interpolating market capitalization for firms with yet missing data in 2020 by taking the mean value of the past three years. The main reason for missing financial data for listed firms is that some company reports have not been published yet, e.g. due to varying accounting regulations by country.

Shareholder categories

Grouping shareholders into meaningful categories is essential to making sense of share ownership data. Although ORBIS assigns categories to most shareholders, its classifications are often questionable, if not wrong. For instance, ORBIS classifies BlackRock and Vanguard, the world's largest asset managers, as a bank and as a non-financial corporation, respectively. This classification error alone would severely distort the results of any analysis seeking to distinguish between the holdings of different types of financial firms.

For our own typology, we use the following categories: Asset managers, banks, endowments, family offices, foundations, hedge funds, insurers, private equity firms, private pension funds, public pension funds, sovereign wealth funds, and wealth managers. We assigned financial institutions to these categories based on a mix of manual searches and classifications made by other data providers. These include the PI/Thinking Ahead Institute's list of the 500 largest asset managers, as well as lists of institutional asset owners used by Preqin and the Sovereign Wealth Fund Institute. Since many more shareholders than the 7000 financial institutions on our list appear in ORBIS, we then use the ORBIS categories to fill up the missing entries in the newly created shareholder category variables. In specific cases, we override ORBIS categories based on a combination of string searches and manual checks. For instance, shareholders whose names contain the words 'authority' or 'national investment' could be assigned to the 'Public sector' category. Compared to the categorization provided by ORBIS, our enhanced shareholder groupings constitute a major improvement for scholars seeking to work with ORBIS shareholder data.

Finally, we used string searches to exclude all shareholder relationships that ORBIS

assigns to aggregate groups of unidentified shareholders.[5] Since ORBIS—presumably depending on incongruous information from national regulatory authorities—reports these free float shareholdings inconsistently across countries, we instead calculate our own, internationally consistent 'Free float' variable by subtracting the sum of all reported stakes held by identified shareholders from 100.

Identification of super-rich families

We classify holdings of super-rich families as a distinct shareholder type. We define super-rich individuals as individuals with an estimated family net worth of 250m USD, and as being listed on a national journalistic rich list in the past 10 years. We use a consolidated rich list including 10,831 of such families globally. We apply a 10-step matching procedure mainly based on individual and firm names to identify individuals from these families.[6] We include their firms, as well as direct and indirect shareholdings as super-rich shareholdings. About 5,884 (54%) of families and related shareholding entities could be identified in the ORBIS database. It must be stressed that not the full population of the super-rich defined in this way could be identified in ORBIS. According to data provider Wealth-X (2021, p. 5), the full population of individuals with USD 250m or more was around 19,640 (without distinguishing the number of families), while we identify around 19,982 family members only from the 5,884 super-rich families that we could match to the ORBIS data[7]. There will be selection on both, individuals that do not occur as shareholders in the database at all, and additional shareholdings of identified individuals that could not be matched correctly to the respective family.

All other rich individuals and the financial institutions serving them are subsumed under the broader category of very high net-worth (VHNW) individuals. This category comprises individual shareholders that ORBIS lists by name ('Named shareholders'), as

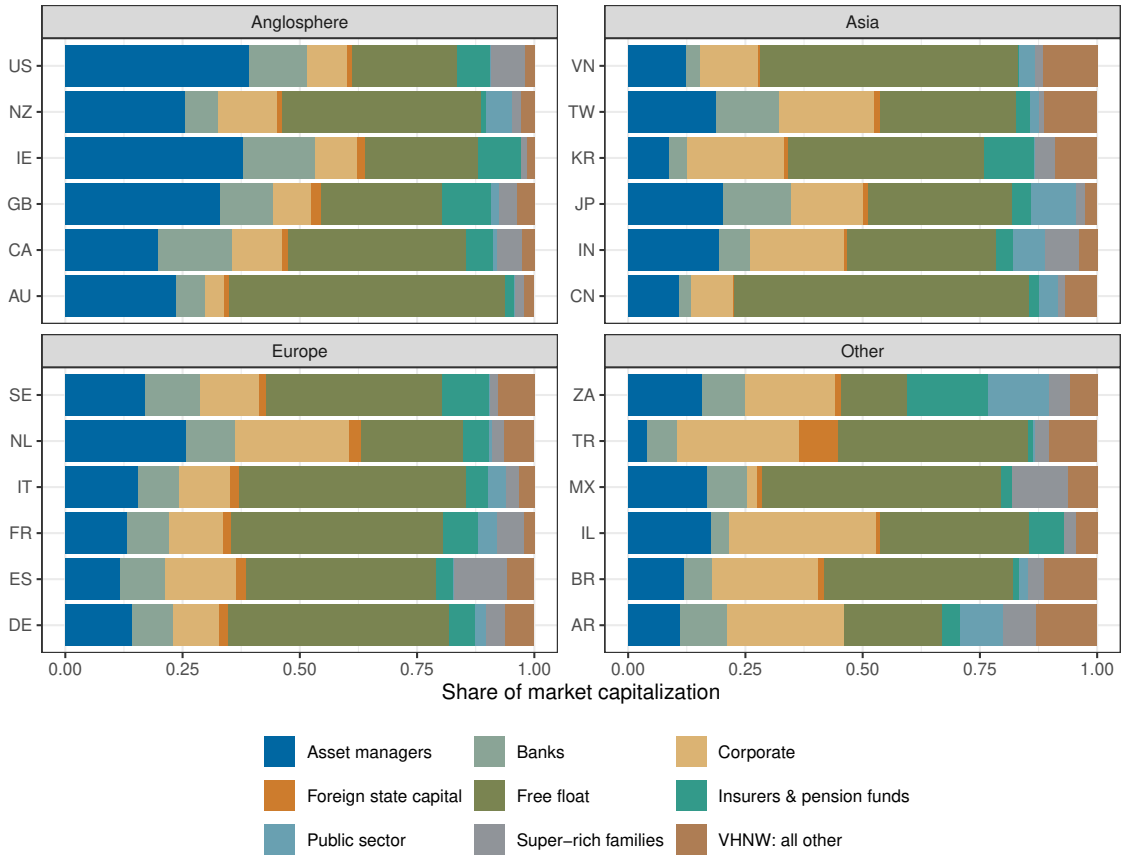
[5]These have ORBIS-assigned shareholder names such as 'Public', 'Institutional investors', or 'Unnamed shareholders, aggregated'.

[6]For a detailed description of the matching process and its results see Appendix 2 to the third article of this dissertation: Linking Wealth and Power.

[7]This number does, however, include a significant number of duplicates. The reason for this is that distinct individuals were inferred from the data and in several cases different ways of spelling led to duplicates which could not be removed but were counted as distinct individuals. This is not relevant for our analysis but to determine the accurate number of distinct individuals that we could identify

well as those family offices, wealth managers, and foundations that have not previously been matched with a super-rich family.[8]

Figure 2: Share ownership by main shareholder categories



Note: Weighted by market capitalization. See the text for information on aggregation methods. Data: ORBIS.

Figure 2 presents the data in its simplest, most easily grasped form, for a selection of large economies. Each bar represents one country and displays the distribution of share ownership, weighted by market capitalization, among the main categories of shareholders. Within the financial sector, we distinguish between asset managers, banks, and pension funds and insurers.[9] Sovereign wealth funds' domestic shareholdings are subsumed under public ownership ('Public sector'), whereas their cross-border shareholdings are assigned to the separate category of 'Foreign state capital'.

[8]Foundations that serve as holdings for super-rich families are assigned to those families. Since we have no way of distinguishing between foundations serving as holdings and foundations that manage their endowments independently, we assign all remaining foundations to the category of very high net-worth individuals.

[9]The asset manager category is dominated by conventional asset managers. Hedge funds and private equity firms account but for a tiny fraction of shareholdings, and only in the Anglophone countries.

The data indicates both variation between (regional) country groups, and significant similarities within those groups. Institutional capital pools dominate share ownership in both English-speaking countries and in Europe. In the former, asset managers account for roughly one-third of total market capitalization. Financial institutions together have taken control of more than 50% of the market capitalization in the US, UK, and Ireland, and close to that in the Netherlands, Japan, and Taiwan. In Asia and selected large economies in the rest of the world, VHNW individuals and corporations have a larger footprint. Of total VHNW shareholdings, roughly half are assigned to identified super-rich families. Foreign state capital accounts for only a small slice of shareholdings, with the exception of Turkey, whereas domestic public ownership plays a sizable role only in India and New Zealand. The seeming absence of public ownership in China is a strong indicator of problems with ORBIS data on shareholders in Chinese companies.

4.2 Sample

Following the implementation of the data preparation steps listed above, we exclude all firms that are inactive or delisted in 2022. We start with a sample of 38 OECD member countries, plus 16 large economies and financial centers that are not OECD members. We reduce this sample by excluding countries with fewer than 50 listed companies, to avoid large country-level effects driven by outliers or data problems.[10] Excluding stock-market listed financial vehicles, such as exchange-traded funds (ETFs) or real estate investment trusts (REITs), which we identify via string searches, eliminates 1,607 entities from the sample. The resulting sample contains 51,222 unique listed firms and 674,738 unique shareholding entities. Using a cut-off for firms for which ORBIS under-reports (less than 10%) or over-reports (more than 150%) shareholder stakes, we arrive at a sample containing 50,125 firms (see Appendix A.1 for details).

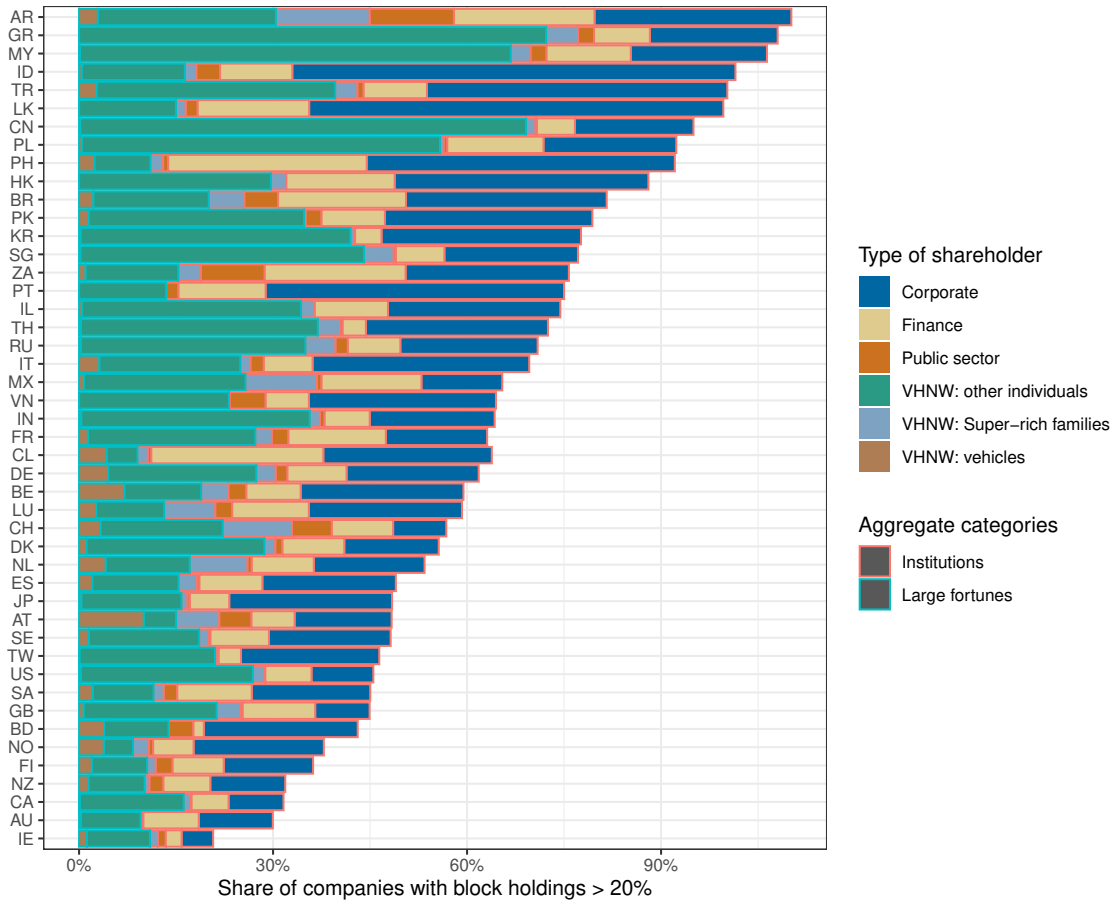
[10]Generally, ORBIS data quality is known to be better for larger firms, and for countries with higher GDP (e.g. Kalemli-Ozcan et al., 2022; Bajgar et al., 2020). The list of countries with fewer than 50 listed companies comprises are Czech Republic, Estonia, Iceland, Hungary, Lithuania, Latvia, Slovakia, and Slovenia.

5 Descriptive results on shareholder concentration

In most countries, stock markets are highly concentrated, and dominated by a relatively small number of (very) large corporations. Where appropriate, this section therefore displays data only for the top fifth of companies, as measured by market capitalization.

5.1 Who are the blockholders?

Figure 3: Share of companies with blockholdings of 20% or larger



Note: Data: ORBIS.

Shareholder concentration is the central outcome variable the literature on comparative shareholder structures seeks to explain. The most widely used indicator is the prevalence of blockholdings, commonly defined as share blocks of 20% or more (Edmans and Holderness, 2017; Aminadav and Papaioannou, 2020). Figure 3 displays the percentage of companies in each country that have at least one 20-percent blockholder. The

ranking shows Argentina, Greece, and Malaysia at the top, followed by Indonesia, Turkey, Sri Lanka, and China. Western and Northern European countries are located in the middle, whereas in the US and other English-speaking countries, only 30 to 40% of companies have a blockholder. Who are these blockholders? The largest group are individuals with large fortunes, marked by the turquoise line. In the interest of transparency, this category is further broken down into VHNW individuals (i.e., shareholders listed by name in ORBIS), super-rich families (identified by the method described above), as well as family offices and foundations.[11] The remaining share blocks—between half and two-thirds of the total in most countries—represent strategic blockholdings by the financial and the non-financial sectors, both domestic and in the form of foreign direct investment.

5.2 Three worlds of shareholder concentration?

The 'stylized fact' of shareholder dispersion in the United States and the other Anglophone countries has recently been re-affirmed by the most sophisticated study of the global shareholder landscape. Using ORBIS data, Aminadav and Papaioannou (2020, 1205-1208) report a concentration measure—the share of voting rights controlled by the one, three, and five largest shareholders—that places the US, the UK, and Australia at the very bottom of the concentration ranking. However, this perspective ignores that concentration is likely to become economically meaningful at a much lower threshold when the one, three, and five largest *institutional* shareholders are *the same* across large parts of the corporate economy.

This is what has happened with the growth of institutional capital pools, which themselves become a *source of shareholder concentration*. As shown in Table 1, which updates the count by Fichtner, Heemskerk and Garcia-Bernardo (2017) of the big-3 asset managers' blockholdings, the number of those has remained broadly similar, with the exception of BlackRock's and Vanguard's blockholdings larger than 10 percent, which have doubled and tripled, respectively. Figure 4 shows the holdings of the big-2 asset man-

[11]Many foundations invest their endowments in diversified portfolios, usually via asset managers. However, when a foundation holds a 20-percent stake in a listed company, that is usually because it serves the function of a family holding. Family offices and foundations are particularly prominent in Austria and Belgium.

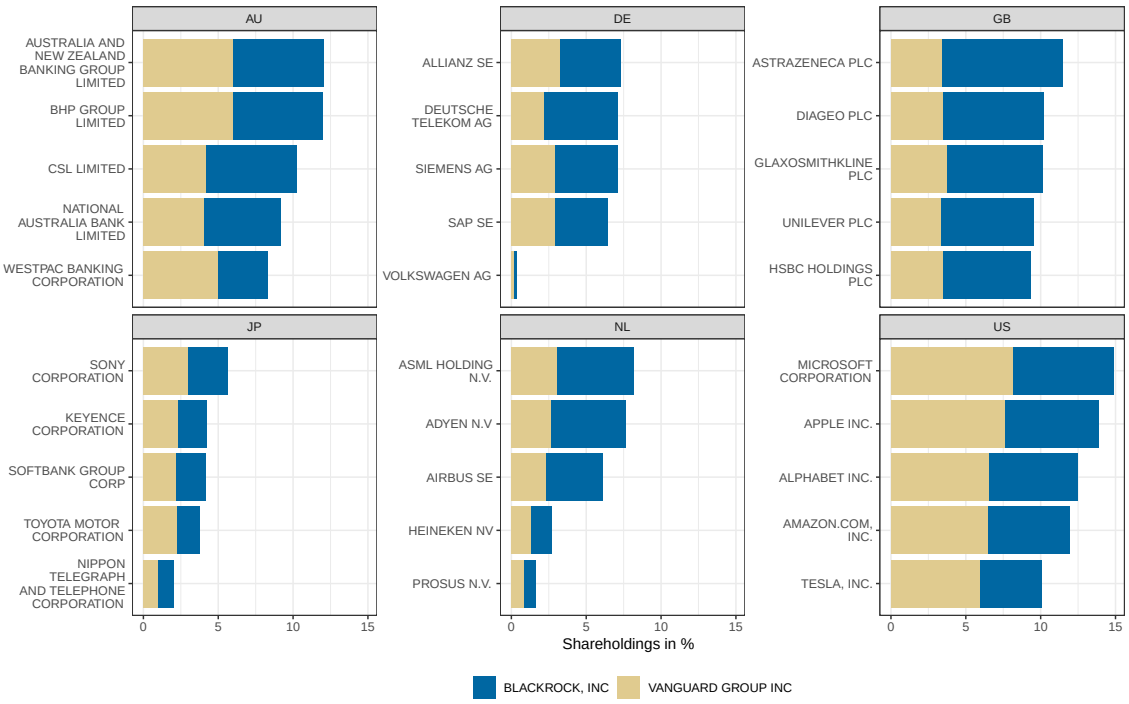
Table 1: Share blocks held by the Big-5 asset managers

	2016			2021		
	>3%	>5%	>10%	>3%	>5%	>10%
BlackRock	3648	2632	375	3352	2540	768
Vanguard	2821	1855	163	3366	1894	582
Fidelity	1956	1309	506	1402	870	230
State Street	1113	281	13	1150	238	16
Dimensional	1708	590	4	1195	491	2

Note: Data for 2016 from Fichtner, Heemskerk and Garcia-Bernardo (2017, 312).

agers in the five largest companies in six selected countries. Their combined holdings are largest in the US, where they range from 10% to 15%, and smallest in Japan, where they mostly remain below 5%.

Figure 4: Big-2 shareholdings in the top-5 companies in selected countries

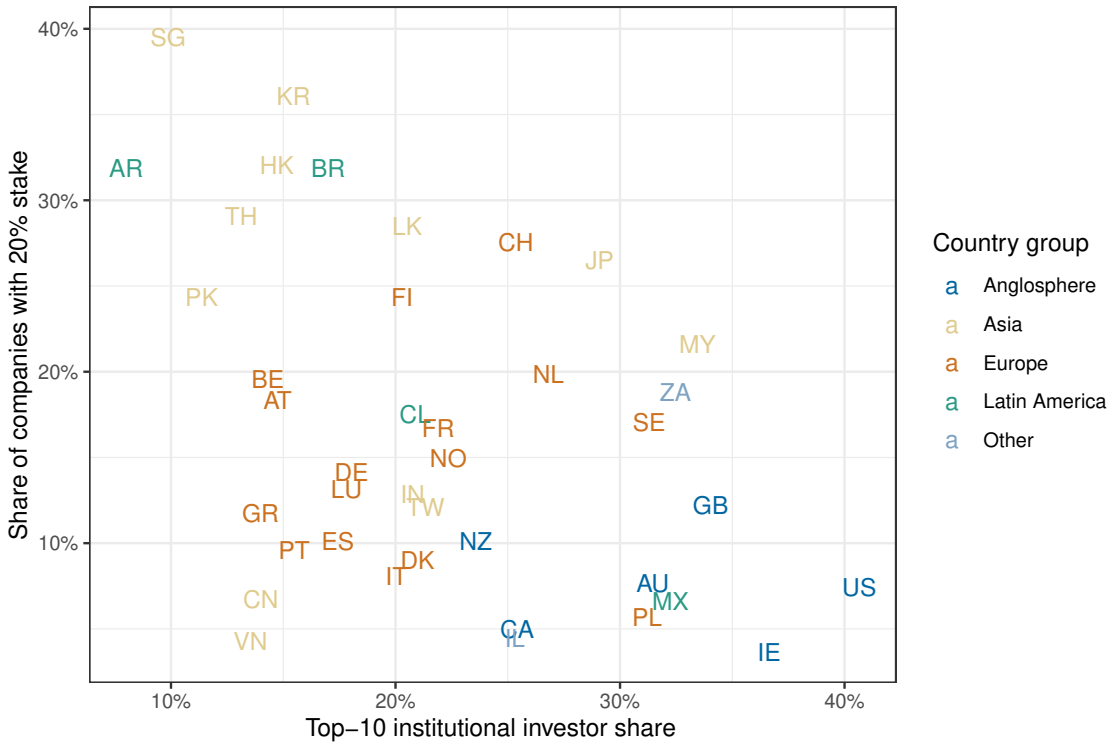


Note: Top quintile of firms per country by market capitalization. Data: ORBIS.

Given this dominant position of the largest asset managers, attempts to measure shareholder concentration need to capture both varieties of concentration—strategic blockholders with concentrated portfolios and global asset managers with diversified portfolios. Figure 5 does precisely that, plotting the share of companies with a 20% stake on the

y-axis and the share held by the ten largest institutional shareholders on the x-axis (i.e., any shareholder that is not an individual, corporation, or state entity). The plot shows data only for the largest quintile of companies, but results look similar for the full sample. The results suggest that there are three distinct worlds of share ownership concentration: strategic blockholder dominance in countries such as Argentina, Brazil, Korea, and Hong Kong; institutional capital pool dominance in English-speaking countries; and a combination of moderate levels of both in Western European countries. In other words, while it is true that those (Anglophone) countries whose shareholder structures are conventionally described as dispersed have few blockholders, they have the *highest* degree of institutional shareholder concentration. European countries combine moderate levels of both types of concentration.

Figure 5: Share of companies with a 20% blockholder versus share held by the top-10 institutional shareholders

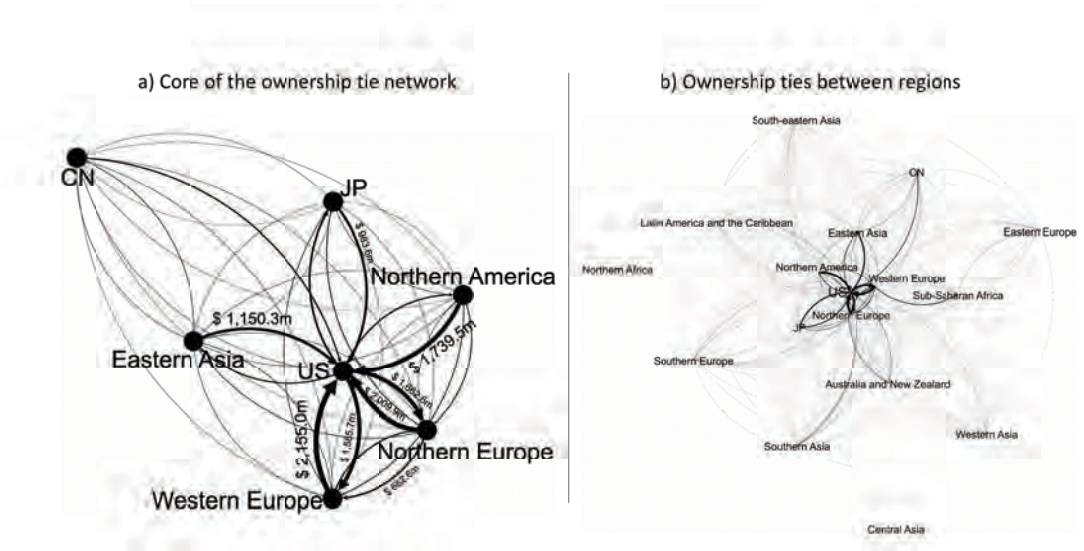


Note: Top quintile of firms per country by market capitalization. Data: ORBIS.

5.3 Global ownership relations

Figure 6 shows the network of ownership relations between the three largest global stock markets US, Japan, and China, and the other world regions' aggregated stock markets. The core of the network and the largest ownership ties are illustrated to the left (6a), and the larger network to the right (6b). Ties denote capital held by shareholders from one region in listed firms of another region. Ties have to be read clockwise, e.g. shareholders from Eastern Asia holding $ 1.15bn in US listed firms. The network clearly centers around the US as the largest capital market of listed firms. The largest ties are accordingly shareholders from other regions holding capital in US listed firms. For example, shareholders from Western Europe hold $ 2.1bn of listed firm capital in US listed firms, and $ 2.0bn are held by shareholders from Northern Europe. The largest share of capital held by US shareholders is $ 1.6bn in listed firms from Northern Europe.

Figure 6: Ownership relations between shareholders and listed firms from the largest three stock markets (US, JP, CN) and the other aggregated world regions.



Note: Direction of shareholdings to be read clockwise. Example: North America holds a larger share in US listed capital than the reverse. The larger network b) excludes small peripheral regions like Polynesia for better readability The graph clearly centers around the US. Data: ORBIS. Visualization tool: Gephi (Bastian, Heymann and Jacomy, 2009)

.

5.4 Inductive typology

It does appear that existing top-down typologies of capitalism capture the financialized capitalism of firms only with difficulties. An alternative typological approach is to generate typologies inductively (Amable, 2003), given the advantage of fine-grained firm-ownership micro-data. To this purpose, we turn to principal-component-analysis which is frequently used to reduce the number of data dimensions describing many individuals. As the most important variables informing potential dimensions, we consider the different features describing the ownership structure of firms: market capitalization, foreign-held share, asset manager share, top-10 institutions share, and top-2 institutions share.

The PCA produces two main dimensions whose joint explained variance approximates 50% or more. The first dimension produces high correlations with any kind of institutional ownership shares (bank, insurance, pension or sovereign wealth funds) and with the absolute size of the capital markets. We interpret this dimension as the degree of institutional capture of deep financial markets. The second dimension, almost vertical to the first, correlates strongly with the degree of rich individuals' and non-profit ownership shares. Uncorrelated with the first two dimensions and in between are firm shares owned by foreign shareholders.

When combining the two dimensions and aggregating on the country level, the resulting Figure 7 yields three to four stylized varieties of firm ownership.

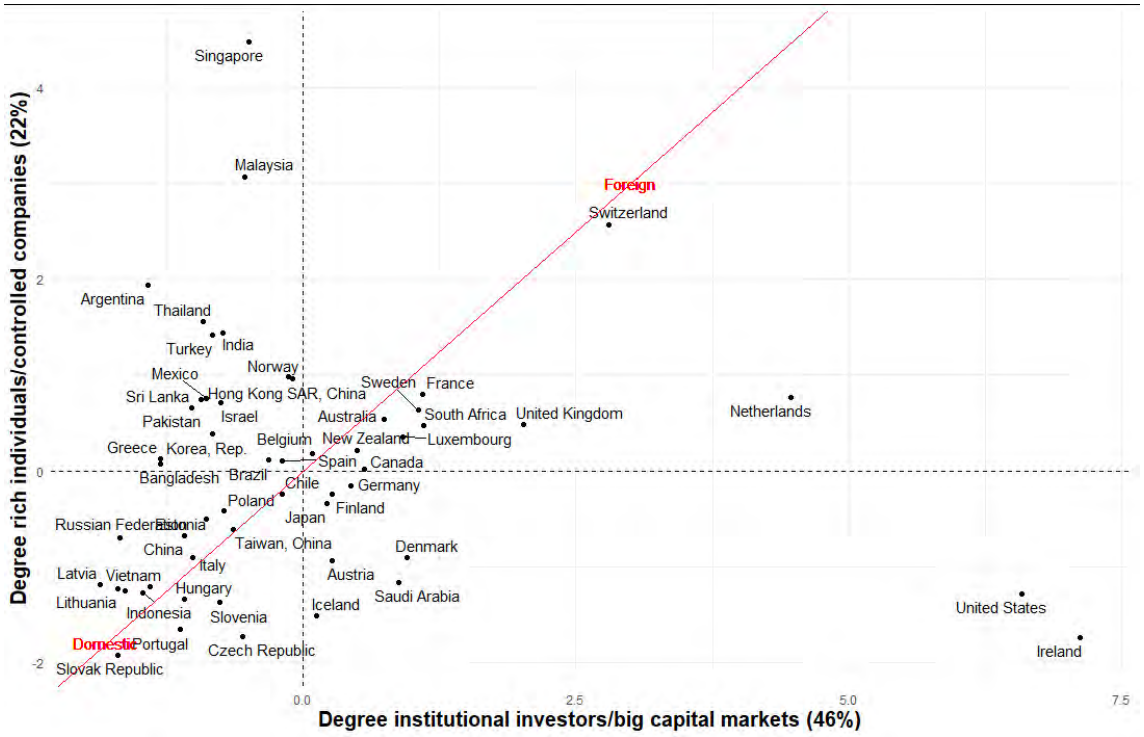
1. Countries with deep capital markets and high rates of institutional share ownership (particularly asset managers, and pension funds). The US or Ireland are extreme cases in this group which otherwise contains some continental and Northern European countries.

2. Countries with low shares of institutional, but high shares of non-profit/individual share ownership (financial oligarchies). In this group, there are many larger and richer emerging economies (e.g. the BRICs).

3. Countries with domestic shareholders and low concentration of owners, no matter whether institutional or not. In this group, there are many Eastern European, Asian, and post-socialist economies.

4. Countries with high foreign ownership shares and moderate levels of both indi-

Figure 7: PCA of ownership structures



.

vidual/institutional shareholder concentration. In this group, there are many Anglo-countries, remaining European countries, particularly the tax-haven countries as extreme cases.

What emerges from this typology is that the simple distinction between concentrated and dispersed ownership is too undercomplex to make sense of how countries group by different dimensions of concentrated ownership.

5.5 Regression results: shareholder perspective

In a final step, we inquire about the determinants of concentrated ownership in two types of multivariate regressions, first on the degree of control and size of top shareholders (controlled stakes of top-1, top-2, and top-5) regardless of who the shareholder is; second, on the shares held by six different shareholder types (banks, insurers, super-rich, asset managers, pension funds, public sector). Given the hierarchical nature of our firm-level data, we use multi-level models with individual firms being nested in countries, which

allows us to include both country-level and firm-level variables. The dependent variable is the sum of all stakes the different shareholders (top-2, banks, insurers, etc.) hold in a given firm. For many of the smaller firms in particular, there is not even one institutional shareholder of every type reported, such that the different shareholder-ownership variables show an inflation of zero-values. We, therefore, conduct two kinds of regressions, a logistic regression is on the binary variable of whether a firm has shareholders of a particular type or not. This likelihood of being owned by a particular shareholder measures how *extensive* or widespread share ownership is. The second OLS regression then focuses on the subsample of firms with a non-zero ownership share which rather follows a normal distribution. The variable has a limited range, but proxies control within firms better than using absolute dollar sums.

As predictor variables, we use (La Porta et al., 1997)'s definition of countries' jurisdictions and also use Maddison's GDP per capita (scaled) as a proxy for economic development (Bolt et al., 2018). We also include institutional variables such as indices for labor/employment regulation and union power (Aminadav and Papaioannou, 2020) as well as the size of the financial sector, proxied through a country's pension fund assets and the financial value added share, both taken from OECD. On the firm level, we use logarithmized market capitalization (dollars) to measure company size and use firms' foundation years to measure their age. Finally, we code the 20-sector NACE code into finance, manufacturing, and other sector to control for firms' sector-specific ownership structures. All firms are nested into the 54 different countries.

Table 2 presents the result for different definitions of how concentrated share ownership is, regardless of who the shareholders are. Column 1 is a logistic regression on whether firms have a shareholder with a controlling stake of more than 20%, whereas columns 2-3 are linear regressions on the non-zero sums of the top-2 and top-5 biggest shareholders of firms, respectively. Controlling stakes follow a different logic than the shares of top owners: firms in French (rather than Common-Law) jurisdictions, in sectors other than finance or manufacturing, and with a low market capitalization are more likely to have a controlling owner. Holdings of the biggest 2 or 5 owners, by contrast, are

not significantly driven by any jurisdiction and are higher for older, bigger, and financial sector firms in countries that are financialized. Union relations or employment regulation do not have significant effects. Overall, the country level explains between 24 and 27 percent of the total variance, the residual being at the firm level.

Table 2: Regressions on controlling stake, top-2 and top-5 shareholdings

	Controlling stake	Big-2 share	Big-5 share
	(1)	(2)	(3)
French (ref. Common Law)	1.142** (0.390)	0.055 (0.327)	0.187 (0.343)
German law	0.433 (0.361)	−0.570 (0.301)	−0.450 (0.317)
Scandinavian	0.376 (0.586)	−0.146 (0.490)	−0.173 (0.514)
Manufacturing (ref. finance)	0.019 (0.048)	−0.172*** (0.030)	−0.193*** (0.032)
Other sectors	0.521*** (0.046)	−0.253*** (0.030)	−0.228*** (0.032)
Founding year	0.015 (0.012)	−0.072*** (0.009)	−0.089*** (0.009)
GDP pc	−0.175 (0.172)	0.314* (0.130)	0.348** (0.134)
Log market cap	−0.123*** (0.004)	0.103*** (0.003)	0.121*** (0.003)
Employment regulation	−0.152 (0.138)	−0.111 (0.121)	−0.123 (0.128)
Collective relation	0.133 (0.144)	0.071 (0.132)	0.088 (0.143)
Finance value added	−0.201 (0.150)	0.294** (0.111)	0.388** (0.122)
Pension fund assets	−0.169 (0.194)	−0.218 (0.172)	−0.176 (0.175)
Constant	0.290 (0.212)	0.601*** (0.171)	0.558** (0.180)
Observations	40,145	10,635	13,209
Log Likelihood	−23,502.580	−14,204.800	−19,407.990
Akaike Inf. Crit.	47,033.160	28,439.600	38,845.980
Bayesian Inf. Crit.	47,153.570	28,548.670	38,958.310
Note:			*p<0.05; **p<0.01; ***p<0.001

While these results help us understand some determinants for why there is control and concentration in share ownership, they do not yet answer the question of what explains the different varieties of concentration. To this purpose, Table 3 shows the results of the logistic regressions on whether firms, nested in countries, are owned by shareholders of different types. The country level explains between 1 and 24 percent of the total variance, the residual variance playing at the firm level. Operating in Common-law jurisdictions significantly increases the likelihood of having one or more asset managers, corporate or public sector entities, or VHNW individuals among the shareholders when compared to other German or French jurisdictions. Firms in Scandinavian jurisdictions, in turn, are much more likely to have pension funds or banks among their owners. A country's GDP per capita does not significantly increase the likelihood of its firms having large institutional capital pools as shareholders. This may be due to the very high significance of the firm-level variable market capitalization which increases the likelihood of having any of the large institutional capital pools as shareholders. Generally, firms in the manufacturing sector are also more likely than financial or any other firm to have large shareholders. Firms of younger age are more likely to count rich individuals among their owners, while older firms have a significantly higher probability of insurers, pension

Table 3: Logistic regression on firm-level holdings by different types of shareholders

	Banks	Insurers	AMs	Rich individuals	Pension funds	Corporates	Public sector
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
French (ref. Common Law)	−0.364 (0.396)	−0.337 (0.398)	−2.085*** (0.467)	0.037 (0.237)	−0.072 (0.325)	−1.573*** (0.389)	−1.238* (0.573)
German law	0.086 (0.478)	0.255 (0.482)	−1.449** (0.559)	−0.470 (0.284)	0.019 (0.390)	−1.358** (0.469)	−0.014 (0.689)
Scandinavian	1.418* (0.641)	0.662 (0.643)	−0.277 (0.754)	−0.558 (0.375)	0.970 (0.518)	0.574 (0.628)	−0.842 (0.937)
Manufacturing (ref. finance)	0.889*** (0.084)	0.573*** (0.076)	0.401*** (0.093)	0.557*** (0.071)	0.609*** (0.079)	1.059*** (0.081)	0.346*** (0.073)
Other sectors	0.632*** (0.083)	0.263*** (0.075)	−0.118 (0.091)	0.219** (0.070)	0.342*** (0.077)	0.416*** (0.078)	0.116 (0.072)
Founding year	0.027 (0.047)	−0.065* (0.027)	0.004 (0.041)	−0.033 (0.032)	−0.094*** (0.019)	0.080 (0.052)	−0.095*** (0.022)
GDP pc	0.354 (0.185)	0.089 (0.186)	0.251 (0.218)	0.277* (0.109)	0.148 (0.151)	−0.133 (0.181)	−0.426 (0.267)
Log market cap	1.076*** (0.012)	0.981*** (0.013)	1.281*** (0.014)	0.766*** (0.010)	0.960*** (0.013)	1.034*** (0.010)	0.755*** (0.010)
Constant	−5.831*** (0.293)	−6.285*** (0.296)	−3.896*** (0.336)	−5.808*** (0.186)	−7.037*** (0.252)	−3.774*** (0.282)	−5.859*** (0.414)
Observations	44,994	44,994	44,994	44,994	44,994	44,994	44,994
Log Likelihood	−7,638.978	−7,866.481	−6,858.393	−8,329.066	−7,367.470	−8,919.011	−8,438.042
Akaike Inf. Crit.	15,297.960	15,752.960	13,736.780	16,678.130	14,754.940	17,858.020	16,896.080
Bayesian Inf. Crit.	15,385.100	15,840.100	13,823.930	16,765.280	14,842.080	17,945.170	16,983.230

Note: *p<0.05; **p<0.01; ***p<0.001

Table 4: OLS regression on firm-level holding percentages by different shareholders

	Dependent variable:						
	Banks	Insurers	Ass Man	VHNW	Pensions	Corporate	Public
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
French (Ref. English)	−0.146 (0.100)	−0.815 (0.899)	−0.113 (0.492)	−0.602 (4.032)	−0.273** (0.117)	−2.189 (2.656)	1.807*** (0.630)
German	−0.126 (0.098)	−1.134** (0.551)	−0.047 (0.488)	−0.049 (3.867)	−0.216 (0.133)	−1.464 (2.214)	0.166 (0.493)
Scandinavian	−0.123 (0.135)	−0.564 (1.111)	−0.247 (0.678)	−0.570 (5.601)	−0.113 (0.173)	−1.282 (3.454)	−0.208 (1.073)
Manufac	−0.162*** (0.062)	0.471 (0.796)	−1.278*** (0.286)	−8.248*** (2.604)	−0.049** (0.022)	−3.703* (2.069)	−0.169 (0.503)
Other (ref. Financial)	−0.153** (0.062)	−0.028 (0.790)	−0.879*** (0.285)	−7.877*** (2.586)	−0.048** (0.022)	−1.616 (2.060)	0.165 (0.500)
Founding year	−0.035* (0.019)	−0.224 (0.243)	−0.246*** (0.086)	−1.688** (0.741)	−0.003 (0.007)	−1.760*** (0.605)	−0.147 (0.164)
GDP pc	−0.057 (0.037)	−0.482* (0.254)	−0.192 (0.196)	−1.267 (1.600)	−0.100** (0.043)	−1.554 (0.947)	0.046 (0.214)
Log market cap	−0.166*** (0.009)	−0.501*** (0.123)	−0.671*** (0.039)	−4.598*** (0.336)	−0.036*** (0.003)	−2.344*** (0.269)	−0.372*** (0.086)
Constant	1.476*** (0.100)	4.234*** (1.188)	5.747*** (0.442)	37.501*** (3.618)	0.505*** (0.092)	19.323*** (2.672)	2.772*** (0.822)
Observations	11,452	8,113	13,931	15,057	7,490	14,844	3,547
Log Likelihood	−23,316.640	−36,208.220	−50,863.950	−88,295.170	−6,050.880	−83,831.660	−12,944.570
Akaike Inf. Crit.	46,655.280	72,438.430	101,749.900	176,612.300	12,123.760	167,685.300	25,911.130
Bayesian Inf. Crit.	46,736.080	72,515.450	101,832.900	176,696.200	12,199.900	167,769.000	25,979.040

Note: *p<0.1; **p<0.05; ***p<0.01

funds, and the public sector among their owners.

Table 4 shows a linear regression on holdings, again for the different types of shareholders. Firms in Common-law jurisdictions generally have higher shares of large shareholders, significantly for insurers and pension funds, whatever their kind, with the exception of public shareholders which have significantly higher shares in French jurisdictions. This holds when controlling for the size of firms in the denominator and also in form of log market capitalization as the independent variable (which has a negative effect throughout as it is more difficult for shareholders to gain higher percentages of larger firms). Note again the difference to table 3: firms in French jurisdictions are less likely in terms of numbers to have public shareholders, but when the state is present, it holds particularly large stakes. This inversion of the relationship between *extensive* and *intensive* share ownership can also be seen for the remaining predictors: now it is firms of old age and particularly in the financial sector which have the highest shares of large shareholders almost throughout all ownership categories and most significantly so. More concentrated share ownership in most categories is also easier in economies with lower GDP per capita.

5.6 Regression results: investee perspective

The results up to now focused on the question of which factors determine companies' shareholder structures. We now switch the perspective to examine the portfolio structure of different types of shareholders. We first investigate the classical sectoral distinction between manufacturing, finance, and other sectors and then turn to sectors particularly intensive in climate-change pushing carbon emissions.

To answer this investee-focused question, we identified the sectoral portfolio composition of the around 600k investing entities in our data set, i.e. which share of their total assets under management went into the two very salient sectors of manufacturing and finance (vs. all others). We use again the binary variable of whether shareholders hold any shares in these sectors as the dependent variable in the first set of logistic regressions which proxies the extensive importance of these investments. In a second set of regressions, we use the non-zero portfolio share as a limited dependent variable in

OLS-regressions for different sector shares - results for using the log of absolute dollar investments are similar - while generally using the shareholder size (log assets), the legal origin (reference: Common-law jurisdiction) of their headquarter country as well as their type (reference category: Asset Managers) as main independent variables.

Table 6 shows that shareholders in Common-law jurisdictions are generally more likely to be invested in firms of different sectors than in other jurisdictions, with the exception of German-law shareholders' preference for manufacturing firms and French/Scandinavian-law shareholders for financial and other service sector firms. In comparison to asset managers, super-rich families and VHNW individuals as well as corporates are much more likely to be invested in manufacturing, whereas banks, pension funds, insurers, or the public sector is much less invested in firms of the manufacturing sector. This sectoral preference is almost exactly reversed for the financial sector. While this describes how *extensive* shareholders of different types are invested across sectors, Table 5 shows that shareholders outside of Common-law jurisdictions tend to have more concentrated portfolio shares in the three different sectors of manufacturing, finance, and other sectors. Only Scandinavian jurisdictions have portfolios less concentrated in these sectors. These results hold for the fact that larger shareholders have significantly lower portfolio shares in general. Relative to asset managers, only foundations, corporates, rich families, and rich individuals have more concentrated portfolios in manufacturing, finance, and mostly also in other sectors.

Economic sector destinations of investments have also been classified by how relevant they are for climate-change-related processes on the Nace-2 level (Battiston et al., 2022; Solutions, 2022). Table 7 and Table 8 present the determinants of whether shareholders' portfolios contain investments in any of the climate-sensitive sectors (fossil fuels, energy-intensive, utility, and oil sectors) and what the cumulative portfolio share of those investing in these sectors is, respectively. These two different approaches do again justice to the zero-inflated variable distribution, but can also be interpreted as how extensive vs. how intensive climate-sensitive investments have been. Large shareholders, for instance, are much more likely to hold oil shares, but less likely to hold energy-intensive firms.

Table 5: OLS on sector portfolio share of shareholders

	Manufacturing	Finance	Other
	(1)	(2)	(3)
Log(AUM)	−0.958*** (0.019)	−0.859*** (0.032)	−2.396*** (0.023)
French (ref. Common Law)	2.049*** (0.374)	4.242*** (0.459)	4.288*** (0.392)
German law	5.705*** (0.162)	3.903*** (0.401)	6.242*** (0.226)
Scandinavian	−3.102*** (0.139)	−2.567*** (0.201)	−12.651*** (0.183)
Banks (ref. AM)	−12.747*** (0.837)	3.556*** (0.915)	−24.152*** (0.980)
Corporate	0.820*** (0.208)	−1.190*** (0.323)	0.217 (0.258)
Employees	11.077*** (2.739)	6.936* (3.131)	14.702*** (3.178)
Endowments	−11.464 (11.103)		−13.042* (6.535)
Family offices	−2.144 (1.947)	8.127* (3.269)	2.801 (2.322)
Foreing state capital	−3.599 (7.024)	−86.326*** (10.804)	−33.716*** (6.988)
Foundations	9.490*** (0.849)	−2.563 (1.342)	9.235*** (1.136)
Hedge funds	−25.167*** (2.186)	−35.807*** (2.716)	−38.814*** (1.632)
Insurers	−4.051** (1.378)	−12.315*** (1.088)	−29.219*** (1.311)
Private equity	3.073* (1.228)	−12.025*** (1.929)	−8.229*** (1.335)
Private pension funds	13.669 (11.108)	5.843 (7.645)	−21.784*** (5.575)
Public pension funds	−24.063*** (4.968)	9.258 (10.802)	−42.127*** (2.863)
Public sector	−3.653* (1.573)	−1.183 (1.372)	−9.556*** (1.315)
Super-rich families	9.582*** (0.540)	4.780*** (0.932)	16.281*** (0.713)
VNNW	10.685*** (0.200)	5.510*** (0.317)	16.956*** (0.251)
Wealth managers	−23.854*** (5.554)		−72.828*** (18.475)
Constant	87.908*** (0.190)	94.135*** (0.286)	85.089*** (0.231)
Fixed effects	Country	Country	Country
Observations	122,509	46,932	139,994
R^2	0.129	0.101	0.254
Adjusted R^2	0.129	0.101	0.254
Residual Std. Error	15.699 (df = 122488)	10.798 (df = 46913)	18.474 (df = 139973)
F Statistic	907.846*** (df = 20; 122488)	293.194*** (df = 18; 46913)	2,377.990*** (df = 20; 139973)

Note: *p<0.05; **p<0.01; ***p<0.001

Table 6: Logistic regression on investment in the sector (vs. none)

	Manufacturing	Finance	Other
	(1)	(2)	(3)
Log(AUM)	0.012*** (0.0004)	0.002*** (0.0003)	−0.018*** (0.0004)
French (ref. Common Law)	−0.101*** (0.007)	0.010* (0.005)	0.095*** (0.007)
German law	0.142*** (0.003)	−0.077*** (0.003)	−0.001 (0.004)
Scandinavian	−0.055*** (0.003)	0.084*** (0.002)	0.046*** (0.003)
Banks (ref. AM)	−0.116*** (0.015)	0.064*** (0.011)	−0.016 (0.015)
Corporate	0.099*** (0.004)	−0.009** (0.003)	−0.043*** (0.004)
Employees	−0.071 (0.054)	0.109** (0.039)	0.070 (0.055)
Endowments	−0.246 (0.133)	−0.093 (0.097)	0.274* (0.136)
Family offices	−0.014 (0.037)	−0.036 (0.027)	0.022 (0.038)
Foreing state capital	−0.112 (0.124)	−0.045 (0.091)	0.182 (0.127)
Foundations	0.024 (0.018)	0.037** (0.013)	0.024 (0.019)
Hedge funds	−0.201*** (0.031)	−0.029 (0.023)	0.236*** (0.032)
Insurers	−0.166*** (0.022)	0.141*** (0.016)	0.083*** (0.023)
Private equity	−0.084*** (0.022)	−0.008 (0.016)	0.074** (0.023)
Private pension funds	−0.268* (0.116)	0.018 (0.085)	0.284* (0.119)
Public pension funds	−0.269*** (0.064)	−0.093* (0.047)	0.470*** (0.066)
Public sector	−0.226*** (0.024)	0.110*** (0.018)	0.174*** (0.025)
Super-rich families	0.071*** (0.011)	−0.007 (0.008)	0.036** (0.012)
VNNW	0.069*** (0.004)	0.022*** (0.003)	−0.026*** (0.004)
Wealth managers	−0.013 (0.110)	−0.105 (0.080)	−0.234* (0.113)
Constant	0.363*** (0.004)	0.086*** (0.003)	0.398*** (0.004)
Fixed effects	Country	Country	Country
Observations	318,478	318,478	318,478
Log Likelihood	−217,316.600	−117,835.600	−225,873.500
Akaike Inf. Crit.	434,675.100	235,713.200	451,789.000

Note: *p<0.05; **p<0.01; ***p<0.001

Shareholders in Scandinavian jurisdictions are more likely to be invested in any of the climate-relevant sectors when compared to their Common-Law counterparts but do so with significantly lower portfolio shares.

In comparison to asset managers, rich families and individuals are significantly more likely to invest in carbon-intensive sectors (with the exception of oil), whereas most other shareholder groups are less likely to be invested in these sectors. This picture is reinforced by Table 8 where rich families and individuals also hold bigger shares in the climate-sensitive sectors when compared to asset managers. Many other shareholder groups are, however, also much more present in these sectors, e.g. corporations, private pensions, hedge or sovereign wealth funds (but not banks).

6 Conclusion

The comparative capitalism literature has long suffered from a lack of up-to-date, global data on corporate shareholdings. This paper is the first to marshal a large-scale dataset of firm-level data to study shareholder structures across the globe. Our first contribution to the fields of economic sociology and political economy is the application of a data processing chain that takes account of the existing problems and shortcomings of the ORBIS data off-the-shelf and makes the shareholder data usable for comparative academic research. As detailed in the data section above, this chain consists of a large number of links and required many rounds and manually correcting numerous issues by trial-and-error. In addition, we have complemented the data by two additional external data sources (one for shareholder types and one for super-rich individuals) that allow us to meaningfully and reliably categorize shareholders—something that is not possible with the original ORBIS dataset. Both the code for data processing and these additional datasets will be shared publicly upon completion of the project.

Substantively, our goals are primarily exploration and documentation—to give readers an overview of what corporate share ownership looks like today across the globe. We have deliberately refrained from presenting too strong of a narrative, in order to let the

Table 7: Logistic regression on investments (or not) in climate-sensitive sectors

	Fossil fuel	Energy-intensive	Utility	Oil
	(1)	(2)	(3)	(4)
Log(AUM)	−0.053 (0.030)	−0.037** (0.013)	0.013 (0.074)	0.503*** (0.018)
French (ref. Common Law)	−0.292 (0.306)	−0.316** (0.116)	−0.042 (0.422)	0.990*** (0.134)
German law	0.329 (0.265)	0.354*** (0.063)	0.436 (0.431)	−1.573*** (0.187)
Scandinavian	1.146*** (0.183)	2.400*** (0.096)	5.280*** (1.048)	0.284* (0.123)
Banks (ref. AM)	−0.512 (0.366)	−0.744*** (0.155)	0.347 (1.065)	−0.605* (0.300)
Corporate	0.256 (0.176)	0.386*** (0.065)	0.565 (0.375)	−0.669*** (0.143)
Employees	13.894 (2,399.545)	12.062 (179.529)	14.820 (2,787.940)	−12.671 (544.946)
Endowments	14.215 (1,385.007)	12.325 (623.837)		2.553* (1.066)
Family offices	13.819 (890.071)	0.572 (0.602)	−1.455 (1.224)	−1.881 (1.018)
Foundations		11.965 (492.030)	15.033 (3,956.180)	−16.836 (1,427.898)
Free float	13.812 (559.391)	1.615** (0.507)	14.191 (1,312.403)	−1.757 (1.014)
Hedge funds	−0.326 (0.625)	0.193 (0.201)	−0.214 (1.081)	−12.426 (157.134)
Insurers	14.277 (639.999)	1.151 (0.722)	14.986 (1,317.397)	−0.493 (0.601)
Private equity	−0.500 (0.631)	−0.540* (0.254)	−0.919 (1.127)	−2.661** (1.018)
Private pension funds	14.252 (690.132)	−0.198 (0.260)	14.108 (1,823.948)	−1.827 (1.010)
Public pension funds	13.927 (1,658.784)	−0.824 (1.161)	15.040 (3,956.180)	−14.524 (1,420.849)
Public sector	13.950 (1,348.829)	−0.816 (0.798)		−1.837 (1.037)
Sovereign wealth	−0.437 (0.649)	−0.258 (0.340)	0.285 (1.078)	−1.333 (0.732)
Super-rich	1.883 (1.021)	0.638*** (0.183)	−0.821 (0.602)	−1.140*** (0.324)
VNNW	2.250*** (0.214)	2.158*** (0.086)	2.472*** (0.647)	−0.002 (0.136)
Wealth managers	14.258 (1,195.744)	12.333 (509.199)		0.258 (1.073)
Constant	2.608*** (0.153)	2.369*** (0.058)	2.505*** (0.332)	−7.014*** (0.146)
Fixed effects	Country	Country	Country	
Observations	33,021	95,358	12,277	413,795
Log Likelihood	−1,529.768	−6,106.069	−172.344	−4,147.246
Akaike Inf. Crit.	3,101.536	12,256.140	382.688	8,338.491

Note: *p<0.05; **p<0.01; ***p<0.001

Table 8: OLS regression on investment share in climate-sensitive sectors

	Fossil fuel	Energy-intensive	Utility	Oil
	(1)	(2)	(3)	(4)
Log(AUM)	−9.606*** (0.159)	−4.900*** (0.061)	−5.633*** (0.263)	−6.560*** (0.578)
French (ref. Common Law)	4.752** (1.466)	7.406*** (0.577)	12.243*** (1.758)	9.348** (3.543)
German law	10.388*** (1.063)	11.273*** (0.245)	12.625*** (1.311)	15.531** (4.999)
Scandinavian	−19.670*** (0.659)	−5.393*** (0.200)	−10.732*** (0.949)	−10.194** (3.341)
Banks (ref. AM)	−3.638 (2.577)	−19.358*** (1.190)	−10.301* (4.589)	−2.422 (6.902)
Corporate	6.655*** (0.808)	2.173*** (0.303)	−1.919 (1.240)	24.382*** (3.436)
Employees	49.008* (21.129)	20.164*** (4.137)	22.120 (13.681)	
Endowments	1.170 (12.202)	−19.223 (14.303)		4.534 (24.371)
Family offices	5.712 (8.009)	−1.187 (2.825)	−27.122* (11.254)	−7.943 (24.402)
Foundations		13.949 (11.681)	21.461 (19.318)	
Free float	29.707*** (5.022)	14.176*** (1.275)	−13.630* (6.908)	40.372 (24.446)
Hedge funds	34.242*** (4.122)	8.463*** (1.024)	−11.442 (5.932)	
Insurers	−14.566* (5.699)	−29.251*** (2.721)	−8.509 (6.596)	−11.887 (14.205)
Private equity	−3.840 (4.544)	−17.195*** (1.821)	−1.204 (7.961)	−9.001 (24.391)
Private pension funds	18.186** (6.137)	7.020*** (1.644)	−12.206 (9.746)	−2.734 (24.385)
Public pension funds	−5.465 (14.954)	−0.398 (11.691)	18.518 (19.319)	
Public sector	22.794 (12.200)	−30.833*** (6.112)		6.963 (24.504)
Sovereign wealth	19.196*** (4.917)	−10.924*** (2.149)	16.129*** (4.074)	2.030 (17.339)
Super-rich	32.876*** (2.586)	20.571*** (0.830)	25.272*** (3.883)	43.637*** (7.810)
VNNW	29.738*** (0.804)	18.013*** (0.293)	20.050*** (1.234)	43.883*** (3.489)
Wealth managers	−28.946** (10.578)	−14.990 (11.679)		−4.388 (24.416)
Constant	88.365*** (0.804)	86.032*** (0.302)	90.345*** (1.397)	64.202*** (4.356)
Fixed effects	Country	Country	Country	
Observations	32,697	93,857	12,234	628
R^2	0.310	0.232	0.260	0.661
Adjusted R^2	0.310	0.232	0.259	0.651
Residual Std. Error	21.096 (df = 32676)	20.223 (df = 93835)	19.278 (df = 12215)	24.262 (df = 610)
F Statistic	735.344*** (df = 20; 32676)	1,350.902*** (df = 21; 93835)	238.600*** (df = 18; 12215)	69.919*** (df = 17; 610)

Note: *p<0.05; **p<0.01; ***p<0.001

data speak for itself as much as possible. That said, we show that the case for a finance capital 2.0 configuration is quite strong, in the sense that shareholder concentration again comes in two guises, namely as blockholdings by VHNW individuals and as holdings concentrated in the hands of few, very large institutional capital pools.

The regression results presented here are meant to demonstrate what kind of research questions the data allow us to ask, and thus to open up new research questions rather than to give definitive answers. Notably, we see considerable potential in using ORBIS data to compare portfolio structures across different categories of shareholders, both within and across jurisdictions. Particularly promising in this regard is the result that VHNW individuals are significantly more likely to be shareholders in carbon emission-intensive sectors than institutional capital pools. This line of research promises to make an important contribution to the fast-growing literature on climate policy support coalitions.

References

Aglietta, Michel. 1979. *A theory of capitalist regulation: The US experience*. London: Verso.

Aguilera, Ruth V. and Gregory Jackson. 2003. "The Cross-National Diversity of Corporate Governance: Dimensions and Determinants." *The Academy of Management Review* 28(3):447–465.

Aguilera, Ruth V. and Gregory Jackson. 2010. "Comparative and International Corporate Governance." *The Academy of Management Annals* 4(1):485–556.

Amable, Bruno. 2003. *The diversity of modern capitalism*. Oxford: Oxford University Press.

Aminadav, Gur and Elias Papaioannou. 2020. "Corporate control around the world." *The Journal of Finance* 75(3):1191–1246. ISBN: 0022-1082 Publisher: Wiley Online Library.

Arndt, Lukas. 2023. "When Should We Believe Research Using ORBIS Firm Data?" Instititution: MPIfG.

Bajgar, Matej, Giuseppe Berlingieri, Sara Calligaris, Chiara Criscuolo and Jonathan Timmis. 2020. Coverage and representativeness of Orbis data. Technical report OECD Paris: .
URL: *https://www.oecd-ilibrary.org/science-and-technology/coverage-and-representativeness-of-orbis-data*$_c$*7bdaa*03 − *en*

Bastian, Mathieu, Sebastien Heymann and Mathieu Jacomy. 2009. "Gephi: an open source software for exploring and manipulating networks.".

Battiston, Stefano, Irene Monasterolo, Bas van Ruijven and Volker Krey. 2022. "The NACE – CPRS – IAM mapping: A tool to support climate risk analysis of financial portfolio using NGFS scenarios.".
URL: *https://papers.ssrn.com/abstract=4223606*

BlackRock. 2022. "Annual Report 2021.".

Bolt, Jutta, Robert Inklaar, Herman de Jong and Jan Luiten van Zanden. 2018. "Rebasing 'Maddison': new income comparisons and the shape of long-run economic development." *GGDC Research Memorandum* 174.

Braun, Benjamin. 2016. "From performativity to political economy: index investing, ETFs and asset manager capitalism." *New Political Economy* 21(3):257–273.

Braun, Benjamin. 2021. Asset manager capitalism as a corporate governance regime. In *The American Political Economy: Politics, Markets, and Power*, ed. Jacob Hacker, Alexander Hertel-Fernandez, Paul Pierson and Kathleen Thelen. Cambridge: Cambridge University Press pp. 270–294.

Braun, Benjamin. 2022. "Exit, Control, and Politics: Structural Power and Corporate Governance under Asset Manager Capitalism." *Politics & Society* 50(4):630–654.
URL: *https://journals.sagepub.com/doi/abs/10.1177/00323292221126262*

Bureau van Dijk. 2020. "Orbis. Powering the business of certainty.".
URL: *https://www.bvdinfo.com/en-gb/our-products/data/international/orbis*

Callaghan, Helen. 2018. *Contestants, Profiteers, and the Political Dynamics of Marketization: How Shareholders gained Control Rights in Britain, Germany, and France*. Oxford University Press.

Davis, Gerald F. 2008. "A new finance capitalism? Mutual funds and ownership reconcentration in the United States." *European Management Review* 5(1):11–21.

Edmans, Alex and Clifford G. Holderness. 2017. Blockholders: A survey of theory and evidence. In *The handbook of the economics of corporate governance*, ed. Benjamin Hermalin and Michael Weisbach. Vol. 1 Elsevier pp. 541–636.

Fichtner, Jan, Eelke M. Heemskerk and Javier Garcia-Bernardo. 2017. "Hidden power of the Big Three? Passive index funds, re-concentration of corporate ownership, and new

financial risk." *Business and Politics* 19(2):298–326.

Fohlin, Caroline. 2005. The history of corporate ownership and control in Germany. In *A history of corporate governance around the world: Family business groups to professional managers*. University of Chicago Press pp. 223–282.

Franks, Julian and Colin Mayer. 2017. Evolution of Ownership and Control Around the World: The Changing Face of Capitalism. In *The Handbook of the Economics of Corporate Governance*. Vol. 1 Elsevier pp. 685–735.

Franks, Julian, Colin Mayer and Stefano Rossi. 2009. "Ownership: Evolution and regulation." *The Review of Financial Studies* 22(10):4009–4056.

Garcia-Bernardo, Javier and Frank W. Takes. 2018. "The effects of data quality on the analysis of corporate board interlock networks." *Information systems* 78:164–172. ISBN: 0306-4379 Publisher: Elsevier.

Gerschenkron, Alexander. 1962. Economic backwardness in historical perspective. In *Economic backwardness in historical perspective: A book of essays*, ed. Alexander Gerschenkron. Cambridge, MA: Belknap Press pp. 5–30.

Gourevitch, Peter A. and James Shinn. 2005. *Political power and corporate control: The new global politics of corporate governance*. Princeton, NJ: Princeton University Press.

Goyer, Michel. 2011. *Contingent capital: Short-term investors and the evolution of corporate governance in France and Germany*. Oxford: Oxford University Press.

Haberly, Daniel and Dariusz Wójcik. 2022. *Sticky Power: Global Financial Networks in the World Economy*. Oxford, New York: Oxford University Press.

Hall, Peter A. and David Soskice. 2001. *Varieties of Capitalism. The Institutional Foundations of Comparative Advantage*. Oxford: Oxford University Press.

Heemskerk, Eelke M., Kevin Young, Frank W. Takes, Bruce Cronin, Javier Garcia-Bernardo, Lasse F. Henriksen, William Kindred Winecoff, Vladimir Popov and Audrey Laurin-Lamothe. 2018. "The promise and perils of using big data in the study of corporate networks: problems, diagnostics and fixes." *Global Networks* 18(1):3–32.

Hilferding, Rudolf. 1985. *Finance capital: a study of the latest phase of capitalist development*. Routledge.

Höpner, Martin. 2003. *Wer beherrscht die Unternehmen? Shareholder Value, Managerherrschaft und Mitbestimmung in Deutschland*. Campus Verlag.

Kalemli-Ozcan, Sebnem, Bent Sorensen, Carolina Villegas-Sanchez, Vadym Volosovych and Sevcan Yesiltas. 2022. "How to construct nationally representative firm level data from the ORBIS global database: New facts and aggregate implications." Publisher: National Bureau of Economic Research Series: NBER Working Paper Series.

Katelouzou, Dionysia and Mathias Siems. 2015. "Disappearing Paradigms in Shareholder Protection: Leximetric Evidence for 30 Countries, 1990–2013." *Journal of Corporate Law Studies* 15(1):127–160.

La Porta, Rafael, Florencio Lopez-de Silanes, Andrei Shleifer and Robert W. Vishny. 1997. "Legal Determinants of External Finance." *The Journal of Finance* 52(3):1131–1150.

La Porta, Rafael, Florencio Lopez-de Silanes, Andrei Shleifer and Robert W. Vishny. 1998. "Law and Finance." *The Journal of Political Economy* 106(6):1113–1155.

La Porta, Rafael, Florencio Lopez-de-Silanes and Andrei Shleifer. 1999. "Corporate ownership around the world." *The journal of finance* 54(2):471–517. ISBN: 0022-1082 Publisher: Wiley Online Library.

Liu, Grace. 2020. "Data quality problems troubling business and financial researchers: A literature review and synthetic analysis." *Journal of Business & Finance Librarianship* 25(3-4):315–371. Publisher: Routledge.
URL: *https://www.tandfonline.com/doi/abs/10.1080/08963568.2020.1847555*

Maher, Stephen and Scott M. Aquanno. 2022. "The New Finance Capital: Corporate Governance, Financial Power, and the State." *Critical Sociology* 48(1):55–73. Publisher:

SAGE Publications Ltd.
URL: *https://doi.org/10.1177/0896920521994170*

Palladino, Lenore. 2022. "The Myth that Shareholders are "Investors".".
URL: *https://lpeproject.org/blog/the-myth-that-shareholders-are-investors/*

Petry, Johannes, Jan Fichtner and Eelke Heemskerk. 2019. "Steering capital: the growing private authority of index providers in the age of passive asset management." *Review of International Political Economy* 28(1):152–176.

Reurink, Arjan and Javier Garcia-Bernardo. 2020. "Competing for capitals: the great fragmentation of the firm and varieties of FDI attraction profiles in the European Union." *Review of International Political Economy* pp. 1–34.

Robé, Jean-Philippe. 2011. "The Legal Structure of the Firm." *Accounting, Economics, and Law* 1(1). Publisher: De Gruyter Section: Accounting, Economics, and Law: A Convivium.
URL: *https://www.degruyter.com/document/doi/10.2202/2152-2820.1001/html*

Roe, Mark J. 1994. *Strong Managers, Weak Owners. The Political Roots of American Corporate Finance*. Princeton, NJ: Princeton University Press.

Solutions, FFI. 2022. The Carbon Underground 100: Oil and gas. Technical report.

Steuer, Sebastian. 2022. *Passive Investoren*. Nomos Verlagsgesellschaft mbH & Co. KG.
URL: *https://www.nomos-elibrary.de/10.5771/9783748936176/passive-investoren*

Stout, Lynn A. 2012. *The shareholder value myth: How putting shareholders first harms investors, corporations, and the public*. San Francisco, CA: Berrett-Koehler Publishers.

Streeck, Wolfgang. 2009. *Re-forming capitalism: Institutional change in the German political economy*. Oxford: Oxford University Press.

Vitali, Stefania, James B. Glattfelder and Stefano Battiston. 2011. "The network of global corporate control - Appendix." *PloS one* 6(10):1–6.

Wealth-X. 2021. "World Ultra Wealth Report.".
URL: *https://go.wealthx.com/world-ultra-wealth-report-2021*

Wigglesworth, Robin. 2021. *Trillions: How a Band of Wall Street Renegades Invented the Index Fund and Changed Finance Forever*. Penguin.

Zysman, John. 1983. *Governments, Markets, and Growth: Financial Systems and the Politics of Industrial Change*. Ithaca: Cornell University Press.

APPENDICES

Peaks of Capital - Appendix

Additional Analyses and Supplementary Information

Contents

Introduction

In this appendix, I provide several additional analyses referenced in the main text. The goal of presenting these additional analyses is to give confidence in the robustness and plausibility of results, and to provide more detailed information on data and analyses, as well as some complementary analyses.

1. Original list of the super-rich

a. Sources of the list

As explained in the main text, the original list was provided by a private researcher (Bornefeld 2019). It is based overall on ca. 370 sources such as journalistic rich lists, mostly from the past 10 years, and additional individual research on individual and family names. Table 1 presents the top sources as examples. The full list cannot be disclosed in line with the copyright agreement.

Table 1: Top 15 sources on which the original list is based

	Number of families
Forbes 2019	1137
Hurun 2019	979
Man.Mag. 2018	882
Hurun 2020	742
CRL	682
Forbes 2018	435
Hurun 2018	415
Sunday Times 2019	404
2019 Challenges	372
Forbes Families 2015	219
BRW 2019	193
MM2019 Manual	193
El Mundo 2019	187
Market Cap Dec. 2019	129
Bilanz 2017	116

b. Geographic distribution of families

Figure 1 in the main text presents the geographic distribution of the findings. Figure 1 in this appendix presents the geographic distribution of the families in the original list. For those countries where families were found, the relations seem comparable: Those countries where most families were found are also the ones where there are the most families in the original list. This speaks at least against a visibly evident selection bias in this step.

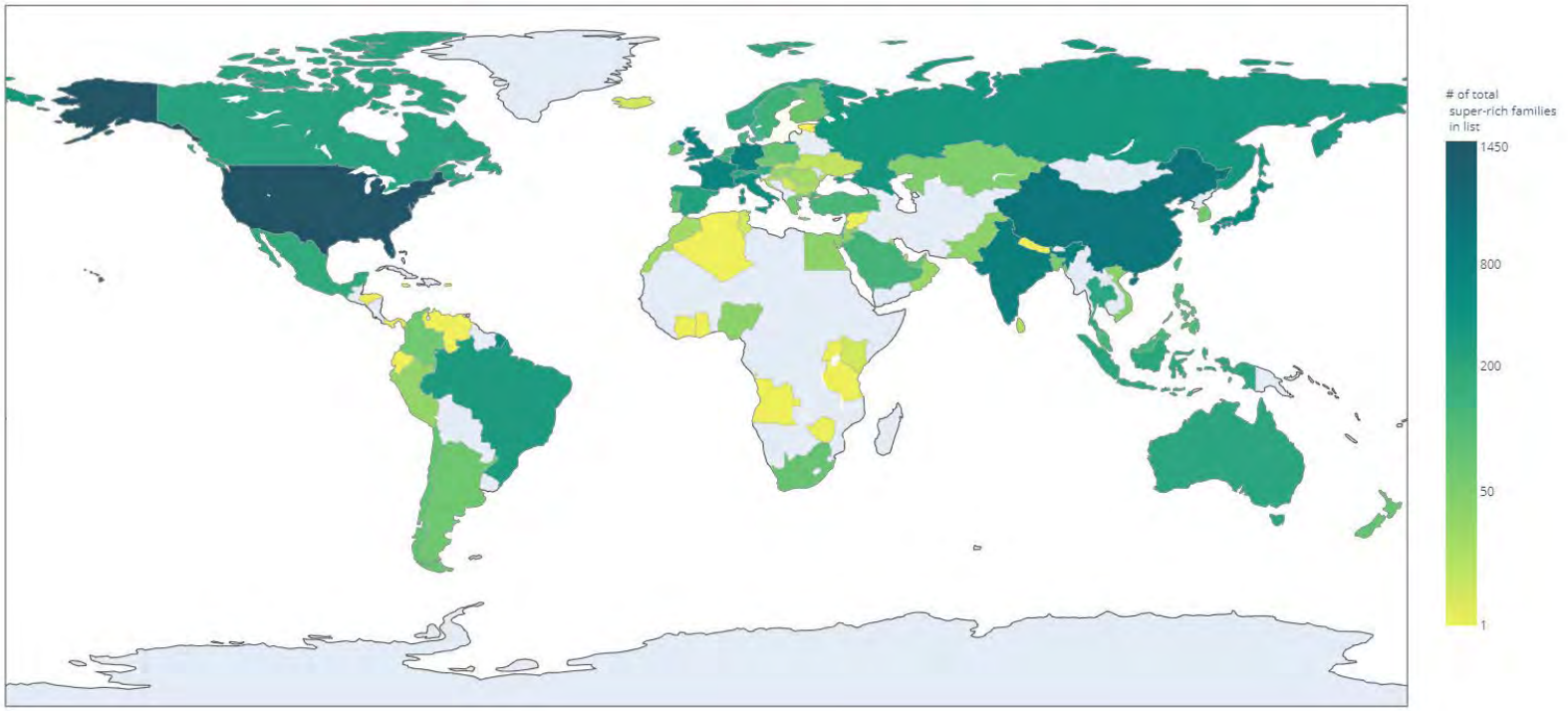


Figure 1: Geographic distribution of families with estimated net worth >= US$ 250m.

2. Matching between the rich list and ORBIS[1]

Matching between the original rich list and the ORBIS data was conducted in eight steps. In principle, the process has two stages: a first one in which the coverage is driven up, i.e., the share of families from the original list for which there is at least one match in the ORBIS data, and a second stage in which family members of the identified individuals and family shares with the same last name were identified. The steps are briefly described in the following.

As preparation for all steps, some preprocessing steps were performed. For example, parts of shareholder and director names that do not contain relevant information were deleted such as "MRS", "Prof.", "Dr.", or " & Family". All names were converted to lowercase. For some steps, the strings were then read in with a nameparser that automatically determined first, mid, and last names from a name string (Gulbranson 2022). In English-speaking countries it is common that names are given in shortened or varied forms in different data sources. For example, "Lawrence Ellison" becomes "Larry Ellison" and William Henry Gates III becomes "Bill Gates". Therefore a collected list of common shortened forms of names was used to add all common English shortened names to English first names (Rémy 2022). For many steps, names that included a name from the top 1% most frequent first or last names (28,827 names) were excluded because these led to too many false-positive matches. If an individual's last name is "Martin" or "Smith", it occurs too frequently that another individual with the same last name who is not related to that individual is involved in the same firm.

Including variant names and different ways of spelling umlauts, the records in the original list amount to about 30,000. These are matched to 4,505,212 distinct shareholder names, and 2,036,257 senior manager and director names (including previous directors before 2020). The challenge of the whole process is that with about 6.5 million names of entities, almost every imaginable name occurs multiple times and the tricky bit is to make sure it is the correct individual based only on the name of that individual, in some cases only the last name, firm name, and country. All of these might be spelled differently, firm names might have changed, or new firms might have been founded to reshuffle the holding structure, etc. The final procedure led to the identification of 44% of the list with a reasonable level of confidence that the match was to the same individual (mainly defined as a fuzzy token set ratio between firm names $>= 80$), and not having too many false positives. Not all additional preparation steps can be explained here, but some are mentioned in the description of the individual steps when necessary.

[1]This article is based on a first version of the matching algorithm results, hence version 1.0. A second version builds the basis for the third article and includes a second download from ORBIS as well as some other advancements. These are described in the appendix to the third article as V2.0.

1st Phase: Driving up coverage of original list

1st Step: Search names among the largest individual global ultimate owners

The first step started with shareholders that are individual or family global ultimate owners (GUO) of a corporation. For all the names in our original list, it was checked whether the first name *and* lastname were included in the family or individual name of global ultimate shareholders of a company. For example, a match is when a GUO first name includes " bill " and the last name " gates " or the full name " henckel " and " fami" in the case of family shares. The main part of this was carried out using similar lines:

```
[(" " + firstname + " " in " " + n + " ") if
  pd.notnull(n) and pd.notnull(firstname) else
  False for n in chunk_dum["GUO - First name"].str.lower().to_numpy()],


([" " + lastname + " " in " " + n + " " if pd.notnull(n)
   else False for n in
   chunk_dum["GUO - Last name"].str.lower().to_numpy()])) &\
   (firstname!=lastname) & (str(firstname)!="nan") & (str(lastname)!="nan"),
   dtype=bool)
```

Obviously, depending on how common a name was, not every match was the correct super-rich individual. Therefore, for all steps matches, needed to be confirmed by using the name of a firm in which the individuals were involved either as shareholders or as directors. This was done by fuzzy matching with the package fuzzywuzzy (Cohen 2020). In more detail for this and several of the other steps, a match was classified as correct when the firm names fulfilled at least one of the following requirements:

1. The firm in which the matched shareholder or director is involved has a very similar name (Token Set Ratio >= 80) compared to the firm name in our original list

2. The firm in which the matched director is involved has a very similar name (Token Set Ratio >= 80) to the most similar firm in ORBIS (ORBIS batch search), to the firm that was listed for this individual in our original list.

In addition to this, the last name of the individual we are looking for may not be among the 1% (28,827) most common names - i.e. the name can be considered as not frequent.

Since these matches are likely to be correct because we can assume the majority of the largest individual and family shareholders in the world also to be among the richest in the world, the firms of matches were added as firm names to confirm matches in subsequent steps.

In total, 8% of families were identified in this step.

2nd Step: Perfect matches

This step simply matched perfect matches, without middle names or lowercase names from the original list. For example if the name "bill gates" from the original list is equal to a shareholder "bill gates", it is kept as a match. The relevant line of code is:

```
rich_names = rich_names[rich_names["compare_names"].isin(chunk["compare_names"])]
```

This was done iteratively for chunks of the millions of shareholder and director names both with and without middle names. For China, this step was only performed disregarding middle names because the automatic determination of first and mid names.

In all, 23% of families were identified in this step, which led to a total of 31% up to this step.

3rd Step: Match names in data from firms matched by ORBIS batch search

The ORBIS portal offers a handy tool to automatically find lists of firms: the firm batch search. In addition to the firm names from our original list, we also used firm names matched in this way. This tool implemented in the ORBIS online portal also takes into account previous and similar names. It therefore offers a flexible matching procedure that is more powerful than simple name matching. Only the director and shareholder data for firms matched in this way were used in this step. For every shareholder and director of these firms, it was checked whether any of them contained the same last name as the individual or family associated with this firm from our list. This was implemented to the implementation of step one, with one of the main lines being:

```
cond1 = np.array([" " + family + " " in " " + lastname + " "
                  if pd.notnull(lastname) else False
                  for lastname,
                  family in zip(families, lastnames)],
                  dtype=bool)
```

To make this more concrete: If anyone with the name Klatten or Quandt was a senior manager or shareholder at BMW or one of the known other companies, he or she would be classified as a match. This leads to more false positives the more common the name is. Therefore this was only done for names that did not contain any of the top 5% most frequent first and last names (144,135 names). Since this was only done for matched firms, firm names were not confirmed again with fuzzy matching in this step.

This step led to 5% additional findings and therefore to overall 36% coverage up to this step.

4th Step: Search first and last names in data

This step is close to the procedure for step 1 but this time it was applied to the full shareholder and director data. This comparison is computationally heavy because 30,000 names from the original list have to be compared to millions of shareholder and director names. The matching was carried out separately for countries. This means records were compared by country, which may lead to unmatched cases where individuals are only a shareholder or director in another country than that which was coded as home country in the original list.

In short, for every shareholder and director name it was checked whether the first and last name of the original list were in the first and last names of directors and shareholders. Alternatively, shareholders may contain only the last name and the word stem "fami" to be classified as a match. Matches were then confirmed again by comparing the names of firms in which the shareholder or director was involved with the firm name from the original list, from the ORBIS batch search, or from the firm names matched in step 1.

This step led to 8% additional findings and therefore to overall 43% coverage up to this step.

5th Step: Plain fuzzy matching

In this step, plain fuzzy comparison of the full name from the original list and the full shareholder and director names was performed. This procedure can detect slight deviations such as misspellings or different conversion of umlauts or other characters in either data source. This step again analyzed the full shareholder and director data, only for families that were not matched yet. The main comparison was done with the cdist function from rapidfuzz, which is a more efficiently implemented fuzzy matching package that can easily be computed in parallel (Bachmann 2022).

```
matches = process.cdist(lower_ind,
            dirnames, scorer=fuzz.WRatio,
            score_cutoff=80, workers=-1)
```

This step led to 2% additional findings and therefore to overall 45% coverage.

6th Step: Looking for foundations

Step 6 was analyzing foundation and trust data specifically. All shareholders that were classified as either "Foundation, research Institute" or "Mutual and pension fund, nominee, trust, trustee" were extracted from the shareholder data. For all names from the original list, it was checked whether there was a foundation or trust that includes the last name from the list *and* held shares in one of the companies that were related to the family. The first part was again done similarly to previous steps by checking whether the name occurred in the foundation or trust name:

```
cond1 = [(" " + lastname + " ") in
(" " + str(x) + " ")
                    for x in country_founds["compare_names"].to_numpy()]
```

The second part was again confirmed by checking fuzzy matching of firm names as described for step 1.

This step led to 2% additional findings and therefore to overall 47% coverage up to this step.

7th Step: Search names among the largest individual global ultimate owners

In this final step of the first phase, only the largest firms in ORBIS were taken as a start. From all firms classified as "Very large company" in ORBIS, those where the GUO was an individual or family were kept. Assuming that the owners of the largest corporations are also likely to be among the global super-rich, this step allowed a looser matching procedure. It was checked whether the last name of any name from the original list was included. Then, the firms owned by the matched names were again compared by fuzzy matching as explained for step 1.

This step led to 3% additional findings and therefore to overall 50% coverage up to this step. Since this was the final step of the first phase to drive up coverage, this share is also equal to the final coverage of families. As a final step to decrease the number of false positive matches, those cases were dropped where there was a shareholder and director last name, but the family name was not in this last

name. In some of the previous steps, it could have happened that last names such as “Martin” could have let to wrong matches. This decreased the degree of found families to 44% but increased its accuracy.

2nd Phase: Snowball family members in the same firms

8th Step: Snowballing of family members

The logic of this main snowballing step is the following. We now have a list of directors and shareholders as a product of the past seven steps. Each of these roles is linked to one or more companies. In all of these companies, it was checked whether there was another individual shareholder or director with the same last name. This procedure was performed iteratively. Every iteration worked roughly like this:

- All firms in which they were involved were merged to directors and shareholders
- In all of these firms it was checked whether there is another directors or shareholder with the same last name in this firm. If yes, they were added as family members.

The procedure was repeated for new findings, until none were found. For this step, as for some previous steps, dashes were removed from the last names, so double names such as Rueggeberg-Buettner become ” rueggeberg buettner ” and the family last name ” rueggeberg ” could therefore be matched successfully. This way at least last names that contain the original family name could be matched which reduces the bias excluding women with different married names.

Findings were then de-duplicated and cleaned to get a final cleaned list of directors and managers, as well as shareholding individuals, trusts, and foundations. As already seen in the main text, more than 9,000 additional family members could be identified in this way, adding up to 13,597 individual members of the 4,250 identified families.

3. Data and preparation

a. Descriptive statistics

Table 2: Descriptive statistics of variables used in regression analysis of the full sample.

	Not super-rich	Super-rich	Overall
	(N=1601425)	(N=6820)	(N=1608245)
Family home continent			
Africa	15979 (1.0%)	88 (1.3%)	16067 (1.0%)
Eastern Europe	158367 (9.9%)	130 (1.9%)	158497 (9.9%)
Far East and Central Asia	447866 (28.0%)	2159 (31.7%)	450025 (28.0%)
Middle East	23098 (1.4%)	261 (3.8%)	23359 (1.5%)
North America	391046 (24.4%)	1117 (16.4%)	392163 (24.4%)
Oceania	28636 (1.8%)	100 (1.5%)	28736 (1.8%)
South and Central America	56971 (3.6%)	458 (6.7%)	57429 (3.6%)
Western Europe	473470 (29.6%)	2500 (36.7%)	475970 (29.6%)
Missing	5992 (0.4%)	7 (0.1%)	5999 (0.4%)
Sex			
F	312028 (19.5%)	840 (12.3%)	312868 (19.5%)
M	1190490 (74.3%)	5918 (86.8%)	1196408 (74.4%)
Missing	98907 (6.2%)	62 (0.9%)	98969 (6.2%)
Age			
Mean (SD)	55.5 (10.6)	60.3 (13.1)	55.5 (10.6)
Median [Min, Max]	55.0 [10.0, 100]	60.0 [24.0, 97.0]	55.0 [10.0, 100]
Missing	929269 (58.0%)	1520 (22.3%)	930789 (57.9%)
Position			
Executive	693173 (43.3%)	1007 (14.8%)	694180 (43.2%)
Supervisory	908252 (56.7%)	5813 (85.2%)	914065 (56.8%)
Board size			
Mean (SD)	203 (786)	39.3 (232)	202 (785)
Median [Min, Max]	12.0 [1.00, 6980]	14.0 [1.00, 6980]	12.0 [1.00, 6980]
Missing	7562 (0.5%)	21 (0.3%)	7583 (0.5%)

Table 2: Descriptive statistics of variables used in regression analysis of the full sample. *(continued)*

	Not super-rich	Super-rich	Overall
Firm founding year			
Mean (SD)	1980 (34.1)	1980 (33.3)	1980 (34.1)
Median [Min, Max]	1990 [1200, 2020]	1990 [1650, 2020]	1990 [1200, 2020]
Missing	40538 (2.5%)	131 (1.9%)	40669 (2.5%)
NACE Sector			
A - Agriculture, forestry and fishing	18073 (1.1%)	79 (1.2%)	18152 (1.1%)
B - Mining and quarrying	19107 (1.2%)	153 (2.2%)	19260 (1.2%)
C - Manufacturing	436404 (27.3%)	2160 (31.7%)	438564 (27.3%)
D - Electricity, gas, steam and air conditioning supply	31927 (2.0%)	76 (1.1%)	32003 (2.0%)
E - Water supply; sewerage, waste management and remediation activities	13960 (0.9%)	31 (0.5%)	13991 (0.9%)
F - Construction	66776 (4.2%)	222 (3.3%)	66998 (4.2%)
G - Wholesale and retail trade; repair of motor vehicles and motorcycles	195255 (12.2%)	747 (11.0%)	196002 (12.2%)
H - Transportation and storage	60087 (3.8%)	175 (2.6%)	60262 (3.7%)
I - Accommodation and food service activities	25146 (1.6%)	165 (2.4%)	25311 (1.6%)
J - Information and communication	76248 (4.8%)	429 (6.3%)	76677 (4.8%)
K - Financial and insurance activities	287439 (17.9%)	1333 (19.5%)	288772 (18.0%)
L - Real estate activities	44022 (2.7%)	306 (4.5%)	44328 (2.8%)
M - Professional, scientific and technical activities	137404 (8.6%)	394 (5.8%)	137798 (8.6%)
N - Administrative and support service activities	58139 (3.6%)	187 (2.7%)	58326 (3.6%)
O - Public administration and defence; compulsory social security	5747 (0.4%)	13 (0.2%)	5760 (0.4%)
P - Education	9613 (0.6%)	36 (0.5%)	9649 (0.6%)
Q - Human health and social work activities	43380 (2.7%)	63 (0.9%)	43443 (2.7%)
R - Arts, entertainment and recreation	20263 (1.3%)	85 (1.2%)	20348 (1.3%)
S - Other service activities	16435 (1.0%)	112 (1.6%)	16547 (1.0%)

Table 2: Descriptive statistics of variables used in regression analysis of the full sample. *(continued)*

	Not super-rich	Super-rich	Overall
T - Activities of households as employers; undifferentiated goods- and services-producing activities of households for own use	37 (0.0%)	0 (0%)	37 (0.0%)
U - Activities of extraterritorial organisations and bodies	19 (0.0%)	0 (0%)	19 (0.0%)
Missing	35944 (2.2%)	54 (0.8%)	35998 (2.2%)
List status			
Delisted	34408 (2.1%)	138 (2.0%)	34546 (2.1%)
Listed	541156 (33.8%)	3662 (53.7%)	544818 (33.9%)
Unlisted	1018834 (63.6%)	2998 (44.0%)	1021832 (63.5%)
Missing	7027 (0.4%)	22 (0.3%)	7049 (0.4%)
Revenue			
Mean (SD)	5720000 (27700000)	3890000 (17600000)	5710000 (27700000)
Median [Min, Max]	117000 [-14800000, 514000000]	453000 [-480000, 514000000]	118000 [-14800000, 514000000]
Missing	135684 (8.5%)	356 (5.2%)	136040 (8.5%)
Employees			
Mean (SD)	15300 (100000)	12500 (56800)	15300 (100000)
Median [Min, Max]	472 [1.00, 2200000]	2010 [1.00, 2200000]	474 [1.00, 2200000]
Missing	402644 (25.1%)	1935 (28.4%)	404579 (25.2%)
Largest component?			
Mean (SD)	0.542 (0.498)	0.833 (0.373)	0.543 (0.498)
Median [Min, Max]	1.00 [0, 1.00]	1.00 [0, 1.00]	1.00 [0, 1.00]

Figure 2 presents the distribution of the different executive and supervisory positions of the super-rich. These numbers are the basis of the individual positions used in the first regression analysis. The categories can be read as an interaction between being a member of a super-rich family, super-rich ownership involvement in a firm, and type of position.

The continuous variables board size, revenue, and employees, as well as estimated family wealth and family size, were logged because of their right skewness. Table 3 presents descriptive statistics of the variables used in regression analysis of the super-rich sample.

Positions of directors in firms in the sample		
	Abs.	**%**
Non-rich director holding BoD position in firm not owned by super-rich family	877,424	54,69
Non-rich executive position in firm not owned by super-rich family	662,325	41,28
Non-rich director holding BoD position in firm where super-rich family has controlling share	30,723	1,91
Non-rich executive position in firm where super-rich family has controlling share	30,806	1,92
Super-rich family is owner-manager	786	0,05
Super-rich family member on BoD in firm not controlled by super-rich family	2,422	0,15
All positions	1,604,486	100

Figure 2: Positions of directors in firms. Interaction between position, family involvement and being super-rich.

Table 3: Descriptive statistics of additional variables used in regression analysis of the super-rich sample.

	Overall
	(N=6927)
Family size	
Mean (SD)	2.62 (2.06)
Median [Min, Max]	2.00 [1.00, 12.0]
Missing	2067 (29.8%)
Estimated net wealth US$m	
Mean (SD)	4390 (10700)
Median [Min, Max]	1400 [250, 191000]
Family ownership	
Controlling share	2924 (42.2%)
Minority share	1462 (21.1%)
No ownership	2541 (36.7%)

As a final description of the used variables, Figure 3 presents correlation of the used continous variables before imputation and after logging.

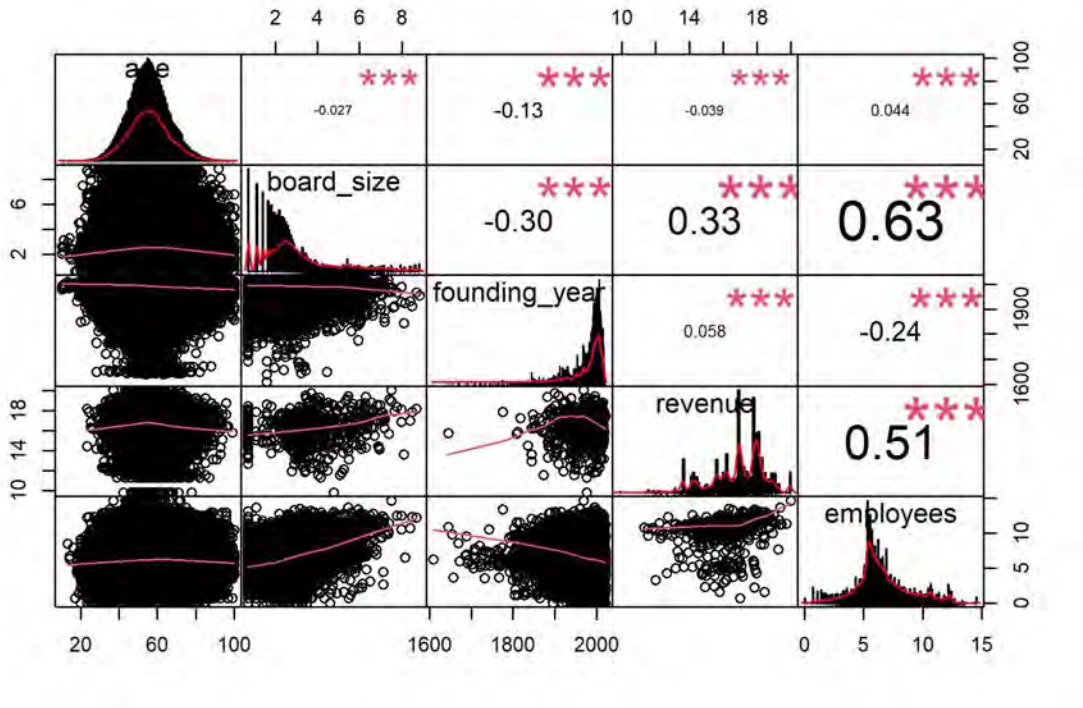


Figure 3: Correlations among continuous variables used in regression analysis of the full sample.

b. Multiple imputation

As presented in the previous section, some variables have missing data. A multiple imputation procedure was applied to estimate plausible values for these missing records instead of dropping the whole record and therefore all information. All presented variables were used for multiple imputation with random forests in Python (Wilson 2021). The resulting distribution of imputed values in comparison to observed values is presented in the following for the full sample (Figure 4) and only for the super-rich (Figure 5).

c. ORBIS sample and subsidiaries

It is known from previous work on interlocking directorates, as well as work using the ORBIS data, that the decision to include or exclude subsidiaries is very relevant. Theoretically, it is possible that they include relevant ties and interlocks that stem

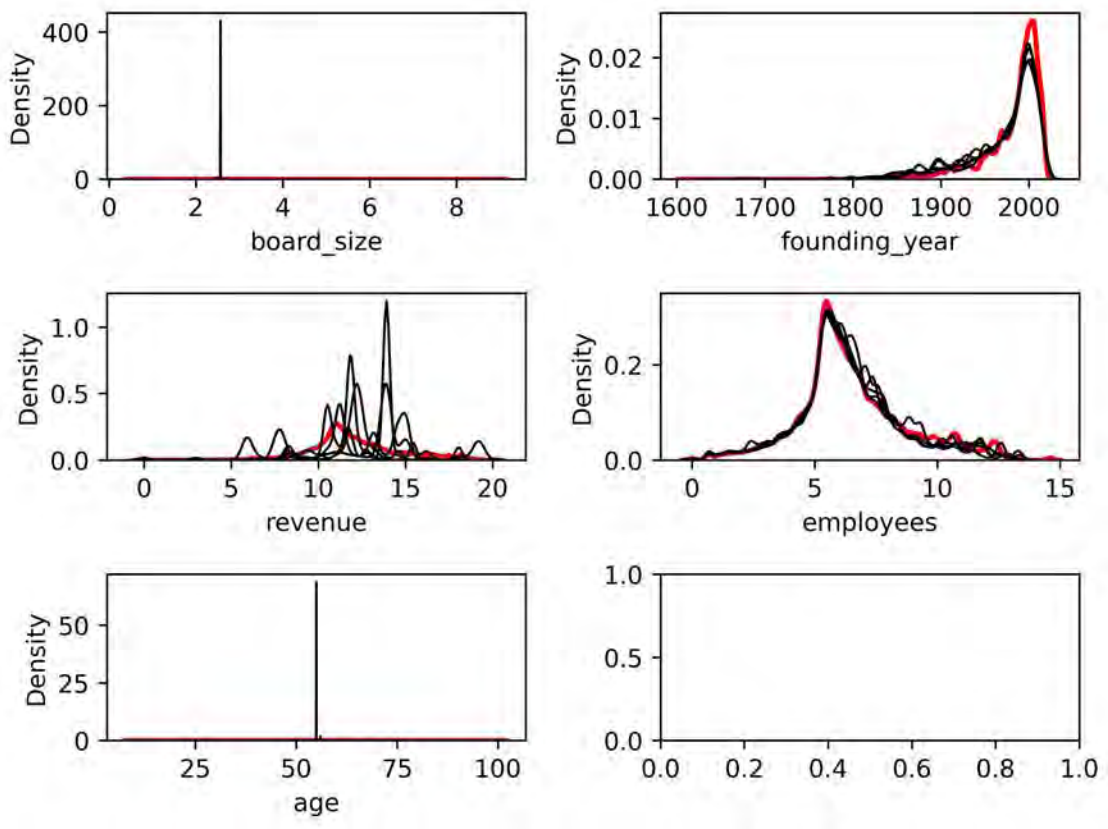


Figure 4: Distribution of imputed values in the full sample in the five imputed datasets (black) vs. observed values (red).

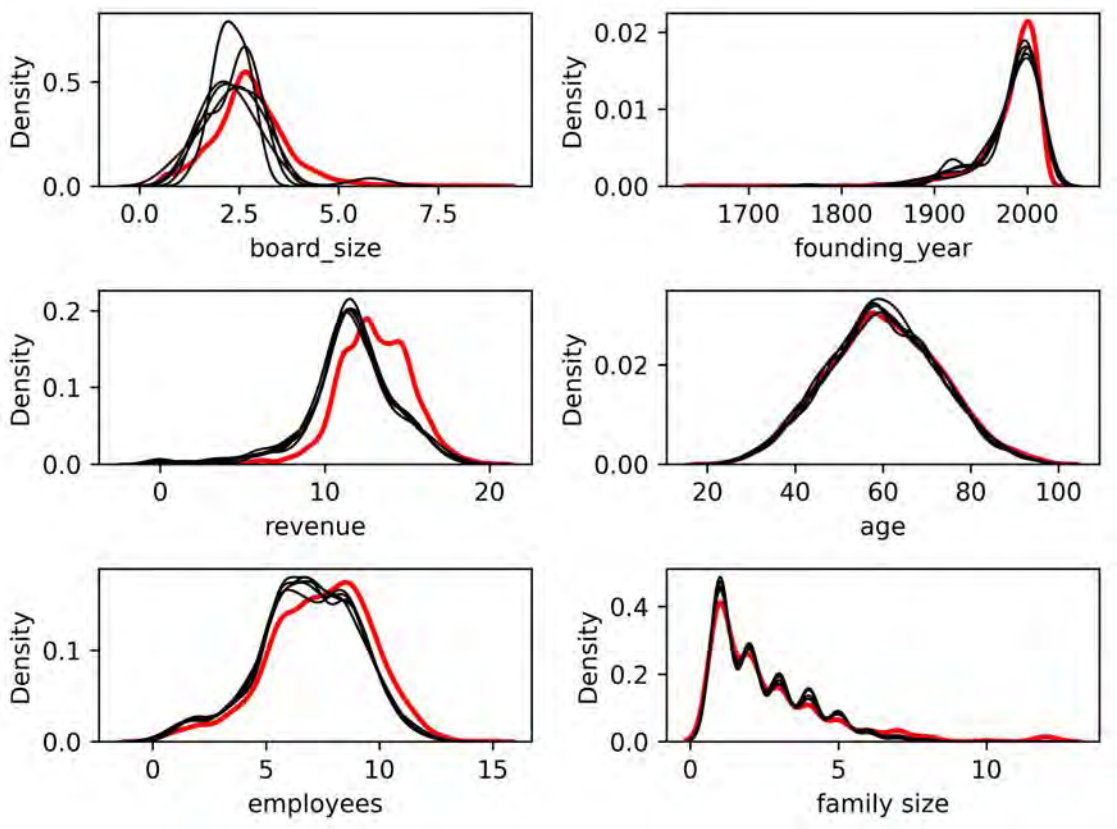


Figure 5: Distribution of imputed values in the super-rich sample's five imputed datasets (black) vs. observed values (red).

from subsidiaries to other corporate groups. On the other side they also include of theoretically irrelevant administrative ties within corporate groups. When it comes to the ORBIS data it is also not straightforward to exclude subsidiaries as Heemskerk and Takes (2016) discuss. They therefore decide to include subsidiaries in their studies. I decided to go the opposite way for two reasons. First, the handling of administrative ties does not seem manageable in ORBIS and different attempts I made to structure them or at least assign them to the correct corporate group failed. Second, the ORBIS portal search seems to have improved in comparison to what Heemskerk and Takes (2016) report. It is now possible to exclude subsidiaries of 50% or more of other *firms* only and therefore to keep those firms that have a majority share from a government or individual. To mention what this means for the examples the authors name: Shell PLC is included because it is not owned by another firm with more than 50%. Although Volkswagen is not included in the sample used for this study, this is because it is a subsidiary of Porsche Holding which is included. Theoretically I think this is justified because it assumes economic power over Volkswagen to be actually in the Porsche Holding Board and senior management. This is possibly debatable but it is a straightforward decision of sampling to include only the head of corporate groups. The final example, Aegon, is also included in this sample according to the definition.

I also briefly note that it is important to exclude non-profit firms, because otherwise many cultural institutions are included in the sample that cannot be meaningfully classified as firms in the sense analyzed here.

4. Main and additional analyses

a. Network comparison with Heemskerk and Takes (2016)

Figure 6 presents a comparison between the network characteristics found in Heemskerk and Takes (2016) and those of the network analyzed in this study. While the sample used in the 2016 study is larger, the selection criteria are comparable. While the comparability of measures gives confidence in the validity of the network and how it was constructed, one can also clearly see evidence of a decrease in the global network of interlocking directorates and its density. For example, the share of directors with more than one board position, the percentage of firms active in the graph, edges per firm, the clustering coefficient, and the share of firms in the largest component have all decreased considerably.

Largest firms for the present study were defined as ca. 50% largest firms in terms of revenue, which is equal to all firms with a revenue of at least US$ 50m, and 50% largest firms in terms of number of employees, which is equal to firms with at least 200 employees. The sample excludes all subsidiaries and nonprofit firms as well as duplicates. Heemskerk and Takes (2016:98) used the ORBIS classifications of

	Heemskerk and Takes (2016)	This study
Sample	2013 Cross-section from ORBIS database for 1,000,000 largest firms including subsidiaries	2020 Cross-section from ORBIS database for 250,000 firms without subsidiaries
Firms	968,409	254,208
Directors	3,262,413	1,473,212
Average directors per firm	3.37	5.79
Directors with >1 affiliation	366,871	94,519
Per cent of directors with >1 affiliation	11.25%	6%
Average number of interlockers per firm	0.38	0.37
Average number of interlockers per active firm	0.94	1.25
Per cent of firms active in graph	40.5%	28.9%
Nodes (companies)	391,992	73,675
Edges (interlocks)	1,712,060	209,290
Edges per active firm	4.37	2.84
Clustering coefficient (active firms)	0.530	0.08
Connected components (Ccs) with size >1	55,620	11,530
Firms in largest Cc	238,866	46,045
Per cent of firms in largest Cc	61%	18%
Edges in largest Cc	1,533,080	176,556
Per cent of edges in largest Cc	89.55%	84.35%
Clustering coefficient in largest Cc	0.62	0.33
Average distance in largest Cc	7.75	6.87

Figure 6: Network Characteristics in Comparison to Heemskerk and Takes (2016)

"very large companies" and "large companies". The major difference might be that subsidiaries were excluded here based on an ORBIS portal search application that was apparently improved over time, while they are part of the sample in Heemskerk and Takes (2016).

Some of the differences presented in Figure 5 might be explained by different samples or other differing definitions, and some measures are not comparable over networks with different sizes. Nevertheless, the overall impression is quite striking that there seem to be far fewer connections between the largest global firms in 2020 compared to 2013. Despite this decrease, the world's richest are part of the remaining network.

b. Full regression tables

Table 4: Full regression models presented in the main text.

	Model 1	Model 2	Model 3
(Intercept)	0.00***	0.00***	3.43
	(0.00)	(0.00)	(13.81)
dm_roletypeSupervisory	1.46***		1.70***
	(0.03)		(0.20)
rich	3.68***		
	(0.25)		
genderM	1.09***	1.09***	1.30*
	(0.01)	(0.01)	(0.15)
age	1.01***	1.01***	1.00
	(0.00)	(0.00)	(0.00)
home_regionEastern Europe	1.57	1.63*	0.42
	(0.36)	(0.38)	(0.36)
home_regionFar East and Central Asia	1.12	1.16	0.84
	(0.26)	(0.27)	(0.68)
home_regionMiddle East	2.98***	2.99***	2.19
	(0.83)	(0.84)	(2.40)
home_regionNorth America	0.90	0.92	0.34
	(0.23)	(0.24)	(0.27)
home_regionOceania	4.95***	4.97***	1.15
	(1.33)	(1.35)	(1.33)
home_regionSouth and Central America	4.53***	4.63***	0.53
	(1.07)	(1.10)	(0.44)
home_regionWestern Europe	3.69***	3.70***	0.73
	(0.83)	(0.84)	(0.57)
board_size	3.10***	3.08***	2.66***
	(0.11)	(0.11)	(0.23)
regionEastern Europe	0.37***	0.35***	0.44
	(0.09)	(0.08)	(0.42)
regionFar East and Central Asia	0.63*	0.62*	0.42
	(0.15)	(0.14)	(0.39)
regionMiddle East	0.51*	0.50*	1.93
	(0.15)	(0.14)	(2.48)
regionNorth America	0.80	0.77	1.13
	(0.21)	(0.20)	(1.04)
regionOceania	0.67	0.67	0.91
	(0.19)	(0.19)	(1.10)

	Model 1	Model 2	Model 3
regionSouth and Central America	0.35***	0.34***	1.42
	(0.08)	(0.08)	(1.35)
regionWestern Europe	0.77	0.76	0.83
	(0.18)	(0.18)	(0.76)
founding_year	1.00**	1.00***	1.00
	(0.00)	(0.00)	(0.00)
NACE_mainB - Mining and quarrying	1.99***	1.97***	0.84
	(0.20)	(0.20)	(0.57)
NACE_mainC - Manufacturing	1.09	1.08	0.54
	(0.08)	(0.08)	(0.23)
NACE_mainD - Electricity, gas, steam and air conditioning supply	1.72***	1.73***	0.93
	(0.17)	(0.17)	(0.73)
NACE_mainE - Water supply; sewerage, waste management and remediation activities	1.07	1.07	1.68
	(0.12)	(0.12)	(1.26)
NACE_mainF - Construction	0.96	0.95	0.61
	(0.08)	(0.08)	(0.30)
NACE_mainG - Wholesale and retail trade; repair of motor vehicles and motorcycles	0.89	0.88	0.42*
	(0.07)	(0.07)	(0.18)
NACE_mainH - Transportation and storage	1.25**	1.25**	0.89
	(0.10)	(0.10)	(0.45)
NACE_mainI - Accommodation and food service activities	1.01	0.99	0.59
	(0.10)	(0.10)	(0.30)
NACE_mainJ - Information and communication	1.31**	1.28**	1.47
	(0.11)	(0.10)	(0.75)
NACE_mainK - Financial and insurance activities	1.59***	1.58***	0.89
	(0.13)	(0.12)	(0.38)
NACE_mainL - Real estate activities	1.24	1.23	0.56
	(0.16)	(0.16)	(0.29)
NACE_mainM - Professional, scientific and technical activities	0.73**	0.73**	0.85
	(0.08)	(0.08)	(0.40)
NACE_mainN - Administrative and support service activities	0.88	0.88	0.36*
	(0.08)	(0.08)	(0.18)
NACE_mainO - Public administration and defence; compulsory social security	1.12	1.13	0.85
	(0.47)	(0.47)	(0.97)
NACE_mainP - Education	1.00	1.01	30359.67
	(0.13)	(0.13)	(163124.69)
NACE_mainQ - Human health and social work activities	0.74**	0.75**	0.85
	(0.08)	(0.08)	(0.53)
NACE_mainR - Arts, entertainment and recreation	1.32*	1.30*	0.78
	(0.16)	(0.16)	(0.47)
NACE_mainS - Other service activities	1.47***	1.47***	0.59
	(0.17)	(0.17)	(0.35)

	Model 1	Model 2	Model 3
NACE_mainT - Activities of households as employers; undifferentiated goods- and services-producing activities of households for own use	2.37	2.37	
	(1.02)	(1.01)	
NACE_mainU - Activities of extraterritorial organisations and bodies	3.02	3.00	
	(2.52)	(2.48)	
listedUnlisted	0.22***	0.22***	0.40***
	(0.01)	(0.01)	(0.07)
revenue	1.11*	1.11*	1.05
	(0.04)	(0.04)	(0.04)
employees	1.05*	1.05	0.88**
	(0.02)	(0.02)	(0.04)
involvement_typeCEO of managerial firm		0.68***	
		(0.01)	
involvement_typeNon-rich BoD of super-rich firm		2.00***	
		(0.17)	
involvement_typeNon-rich CEO of super-rich firm		1.28	
		(0.16)	
involvement_typeOwner-manager		2.53***	
		(0.32)	
involvement_typeRich BoD without controlling share		8.70***	
		(0.99)	
involvement_typeRich Executive without controlling share		1.73	
		(0.55)	
involvement_typeRich Owner BoD		2.20***	
		(0.20)	
firmsumshareMinority share			1.48*
			(0.27)
firmsumshareNo ownership			2.11***
			(0.24)
net_wealth			1.49***
			(0.08)
family.size			1.23***
			(0.05)
AIC	1335195.91	1332247.79	4867.88
BIC	1335736.69	1332850.03	5175.82
Log Likelihood	−667553.95	−666074.89	−2388.94
Deviance	1335107.91	1332149.79	4777.88
Num. obs.	1608244	1608244	6927

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

c. National k-cores

The following figure presents the share of different types of super-rich involvement in the cores of national networks. For this purpose, only firms from the respective countries were analyzed.

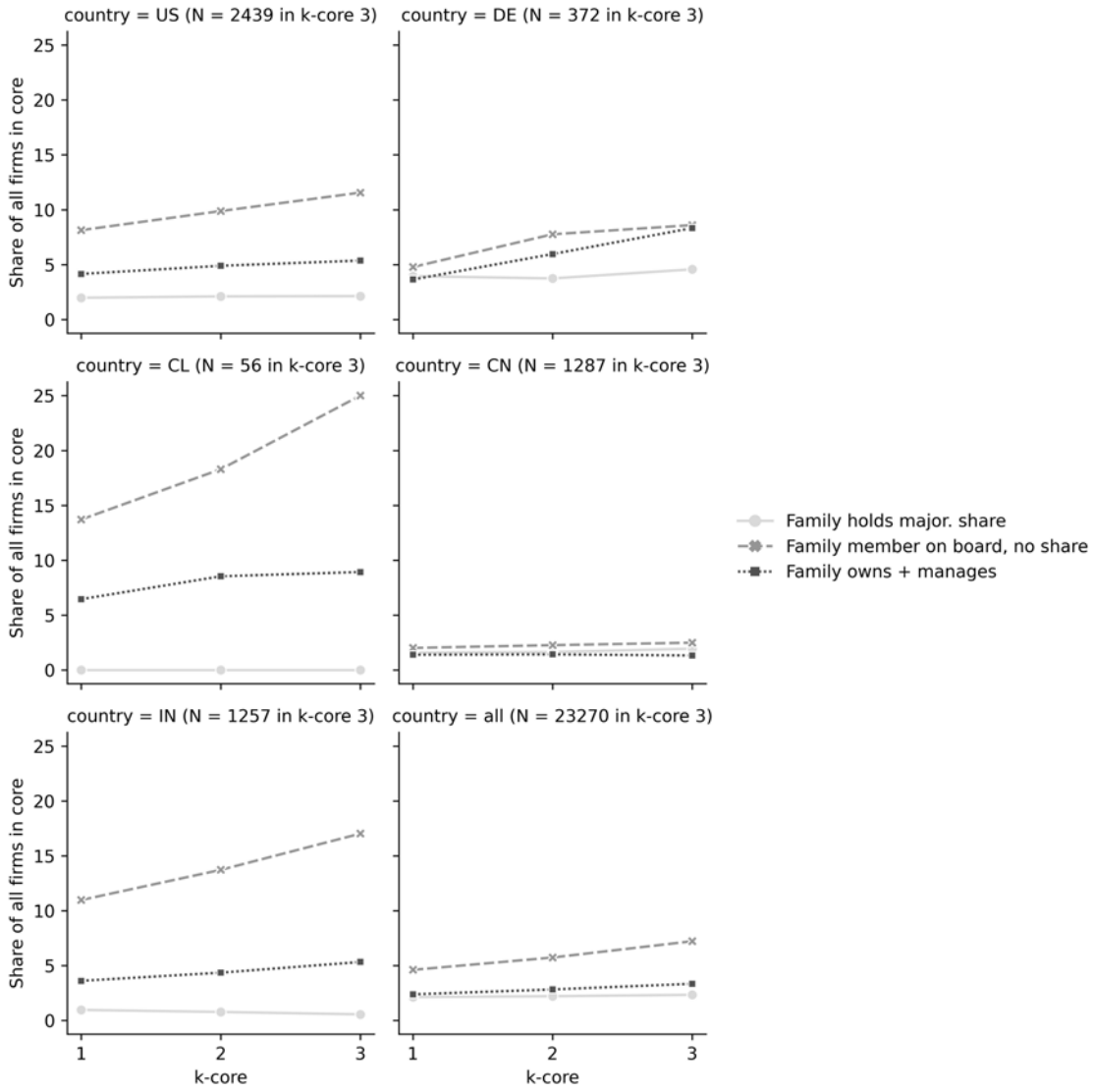


Figure 7: Super-rich involvement in firms in the k-cores for five different countries and all countries combined.

d. Firm dyadic analysis

An alternative explanation is that different super-rich families may be connected not by joining other boards themselves, but by sending employees from their own corporate board. This would be an indirect tie because it means two super-rich family members would not sit on the same board, but would sit on two boards that were interlocked by at least one other board member. If this was a common way for the super-rich to be connected, their firms would be more likely to be interlocked. To test this, I also ran the dyadic regression research design for all firms in the largest component of the firm network. Results show that firms of the super-rich do not show odds ratios to be connected which can be distinguished from zero with enough confidence. This alternative explanation is therefore also rejected here.

Table 5: Weighted logistic regression

	INTERLOCK	
	1	2
rich	4.434 (5.635)	3.872 (4.818)
sector		3.209 (4.142)
country		64.855** (86.999)
size		1.653 (2.368)
Constant	0.000*** (0.000)	0.000*** (0.000)
Observations	32	32
Notes:	*p<0.05; **p<0.01; ***p<0.001	

e. Dyadic analysis without control variables

A common problem in regression analysis is overcontrol bias. It may occur when a control variable is introduced which is correlated with an independent and the dependent variable, and it may bias regressors. To rule out that results presented in the main text are to a significant extent influenced by such biases, the dyadic regression analysis is presented in the following without control variables. One could assume a potential overcontrol bias, for example, if the super-rich who sit on the same boards also tend to cluster in the same sectors and countries, while these two variables are also controlled for. However, the estimated odds ratio for sitting on the same boards as two family members from different dynasties is also less than one and is insigniicant when not including control variables.

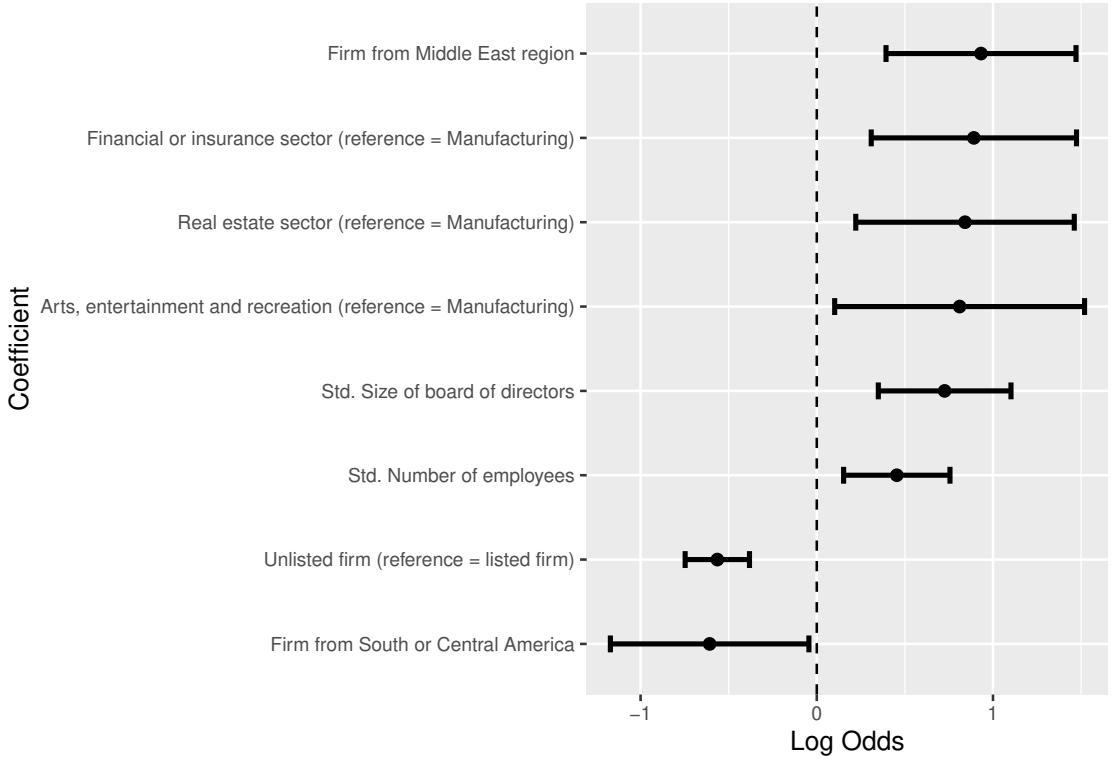
Table 6: Weighted logistic regression

	INTERLOCK	
	All directors	Super-rich family members
	1	2
rich	.306	
rich:wealthdiff		.261*
rich:famdiff		.193**
Constant	0.000***	.004***
Observations	95	1,673
Notes:	*p<0.05; **p<0.01; ***p<0.001	

f. Additional firm analysis

To understand what predicts the presence of the members of super-rich families on Boards of Directors, the main text presents results from a small additional analyis. Super-rich membership of a firm's supervisory board is used as a dependent variable. All firm variables presented in the main text are used as regressors. Figure 5 in the main text presents those coefficients which are significantly different from zero. However, the variable "Another super-rich individual or firm already involved in firm" might be problematic because it is likely to be related to the outcome variable, as well as independent variables. To rule out that this potential problem biases the results presented in Figure 5, results are presented here for the same model without this variable.

Results remain the same with only one relevant difference. This is that the effect of a firm being in the arts and recreation sector, compared to the manufacturing sector, are not significantly different anymore. This questions the robustness of this difference in the main analysis and more detailed analysis of what is going on within this sector in comparison to others is necessary to understand this better. This, however, will be left for future studies and would exceed the scope of this project.



I. References

Bachmann, Max. 2022. "Rapidfuzz." Available at: https://pypi.org/project/rapidfuzz/.

Bornefeld, Andreas. 2019. "Vermögensforschung." Available at: https://www.reichsten.de/en/index. [Accessed: April 26, 2019].

Cohen, Adam. 2020. "fuzzywuzzy Python Package." Available at: https://pypi.org/project/fuzzywuzzy/. [Accessed: February 28, 2020].

Gulbranson, Derek. 2022. "Python-Nameparser." Available at: https://github.com/derek73/python-nameparser.

Heemskerk, Eelke M., and Frank W. Takes. 2016. "The Corporate Elite Community Structure of Global Capitalism." *New Political Economy* 21(1):90–118.

Rémy, Philippe. 2022. "Name-Dataset." Available at: https://github.com/philipperemy/name-dataset/.

Wilson, Samuel. 2021. "Miceforest." Available at: https://github.com/AnotherSamWilson/miceforest.

Linking Wealth and Power - Appendix

Additional Analyses and Supplementary Information

Contents

Introduction

In this appendix, I provide several additional analyses referenced in the main text. The goal of presenting these additional analyses is to give confidence in the robustness and plausibility of results, and to provide more detailed information on data and analyses, as well as some complementary analyses.

1. Original list of the super-rich

a. Sources of the list

As explained in the main text, the original list was provided by a private researcher (Bornefeld 2019). It is based overall on ca. 370 sources such as journalistic rich lists, mostly from the past 10 years, and additional individual research on individual and family names. In addition to that, the list was complemented by many other publicly available rich lists scraped from the internet. Table 1 presents the top sources for the US and German families as examples. The full list cannot be disclosed in line with the copyright agreement.

Table 1: Top 15 sources on which the original list is based

Source	Number of families
Man.Mag. 2018	881
Forbes 2019	548
CRL	318
Forbes Families 2015	219
MM2019 Manual	193
ManMag 2019	59
Man.Mag. 2017	54
Forbes 1988	39
Forbes 1996	39
Forbes 1998	30
Forbes 2018	30
Bloomberg Family 2019	29
Orange County Richest 2019	28
Forbes 2007	24
Market Cap Oct. 2018	21

2. Matching between the rich list and ORBIS (V2.0)

Matching between the original rich list and the ORBIS data was conducted in ten steps. The process has two stages: a first one in which the coverage is driven up, i.e., the share of families from the original list for which there is at least one match in the ORBIS data, and a second stage in which family members of the identified individuals and family shares with the same last name were identified. Next, I give a briefly overview of differences between two versions of the matching process used in this dissertation which were implemented at two different time points (the first in the first months of 2022, and the second in late summer 2022). After that, the ten steps are briefly described in the following.

There are a few differences between the improved matching V2.0 performed for the article "Linking Wealth and Power" in comparison to the matching V1.0 as applied for the article "Peak of Capital?". Besides several minor improvements in the code that are not discussed here in detail, there are four major improvements. First, V1.0 did not include ca. 1000 individuals from the original list that only had an estimated fortune in another currency than dollars. These are now included without further adjustment for the exact size of the fortune. Second, the original list was complemented by additional publicly available rich lists. This way, around 300 additional names were added to the list. Third and most importantly, the matching was performed on a second download from ORBIS that enabled to find a larger share of the super-rich. As basis for V2.0, I used ORBIS data from the first ORBIS download from 2020 in addition to a second download from 2022 with a much larger sample of firms (roughly 4m largest firms and their shareholders where available). While in V1.0 13,597 individual family members were identified, V2.0 includes more than 21,000 matches in the ORBIS data. It is not easy to say how many of these matches are actually individuals because they also include corporate bodies such as foundations or trusts. However, the fourth improvement is also that a the deduplication algorithm DeDupe (Dedupe.io 2022) was trained and used to identify and cluster duplicate individuals among families. Deduplication of individual family members is therefore also improved in comparison to V1.0.

As preparation for all steps, some preprocessing steps were performed. For example, parts of shareholder and director names that do not contain relevant information were deleted such as "MRS", "Prof.", "Dr.", or " & Family". All names were converted to lowercase. For some steps, the strings were then read in with a nameparser that automatically determined first, mid, and last names from a name string (Gulbranson 2022). In English-speaking countries it is common that names are given in shortened or varied forms in different data sources. For example, "Lawrence Ellison" becomes "Larry Ellison" and William Henry Gates III becomes "Bill Gates". Therefore a collected list of common shortened forms of names was used to add all common English shortened names to English first names (Rémy 2022). For many

steps, names that included a name from the top 1% most frequent first or last names (28,827 names) were excluded because these led to too many false-positive matches. If an individual's last name is "Martin" or "Smith", it occurs too frequently that another individual with the same last name who is not related to that individual is involved in the same firm.

Including variant names and different ways of spelling umlauts, the records in the original list amount to about 30,000. These are matched to 4,505,212 distinct shareholder names, and 2,036,257 senior manager and director names (including previous directors before 2020). The challenge of the whole process is that with about 6.5 million names of entities, almost every imaginable name occurs multiple times and the tricky bit is to make sure it is the correct individual based only on the name of that individual, in some cases only the last name, firm name, and country. All of these might be spelled differently, firm names might have changed, or new firms might have been founded to reshuffle the holding structure, etc. The final procedure in V2.0 led to the identification of 56% of the list with a reasonable level of confidence that the match was to the same individual (mainly defined as a fuzzy token set ratio between firm names >= 80), and not having too many false positives. Not all additional preparation steps can be explained here, but some are mentioned in the description of the individual steps when necessary.

1st Phase: Driving up coverage of original list

1st Step: Search names among the largest individual global ultimate owners

The first step started with shareholders that are individual or family global ultimate owners (GUO) of a corporation. For all the names in our original list, it was checked whether the first name *and* lastname were included in the family or individual name of global ultimate shareholders of a company. For example, a match is when a GUO first name includes " bill " and the last name " gates " or the full name " henckel " and " fami" in the case of family shares. The main part of this was carried out using similar lines:

```
[(" " + firstname + " " in " " + n + " ") if
  pd.notnull(n) and pd.notnull(firstname) else
  False for n in chunk_dum["GUO - First name"].str.lower().to_numpy()],


([" " + lastname + " " in " " + n + " " if pd.notnull(n)
   else False for n in
   chunk_dum["GUO - Last name"].str.lower().to_numpy()])) &\
```

```
    (firstname!=lastname) & (str(firstname)!="nan") & (str(lastname)!="nan"),
    dtype=bool)
```

Obviously, depending on how common a name was, not every match was the correct super-rich individual. Therefore, for all steps matches, needed to be confirmed by using the name of a firm in which the individuals were involved either as shareholders or as directors. This was done by fuzzy matching with the package fuzzywuzzy (Cohen 2020). In more detail for this and several of the other steps, a match was classified as correct when the firm names fulfilled at least one of the following requirements:

1. The firm in which the matched shareholder or director is involved has a very similar name (Token Set Ratio >= 80) compared to the firm name in our original list
2. The firm in which the matched director is involved has a very similar name (Token Set Ratio >= 80) to the most similar firm in ORBIS (ORBIS batch search), to the firm that was listed for this individual in our original list.[1]

In addition to this, the last name of the individual we are looking for may not be among the 1% (28,827) most common names - i.e. the name can be considered as not frequent.

Since these matches are likely to be correct because we can assume the majority of the largest individual and family shareholders in the world also to be among the richest in the world, the firms of matches were added as firm names to confirm matches in subsequent steps.

In total, 10% of families were identified in this step.

2nd Step: Perfect matches

This step simply matched perfect matches, without middle names or lowercase names from the original list. For example if the name "bill gates" from the original list is equal to a shareholder "bill gates", it is kept as a match. The relevant line of code is:

```
rich_names = rich_names[rich_names["compare_names"].isin(chunk["compare_names"])]
```

[1]This makes a difference because the ORBIS batch search may for example identify new names of firms such as the current match "Meta" when searching for "Facebook".

This was done iteratively for chunks of the millions of shareholder and director names both with and without middle names. For China, this step was only performed disregarding middle names because the automatic determination of first and mid names fails here among other things because order of names is more ambiguous in the original data: sometimes the order follows western norms of firstname lastname, and sometimes the original mandarin order of lastname firstname is retained. A second reason is that the human name parser used only takes account of the western norm.

In all, 29% of families were identified in this step, which led to a total of 39% up to this step.

3rd Step: Match names in data from firms matched by ORBIS batch search

The ORBIS portal offers a handy tool to automatically find lists of firms: the firm batch search. In addition to the firm names from our original list, we also used firm names matched in this way. This tool implemented in the ORBIS online portal also takes into account previous and similar names. It therefore offers a flexible matching procedure that is more powerful than simple name matching. Only the director and shareholder data for firms matched in this way were used in this step. For every shareholder and director of these firms, it was checked whether any of them contained the same last name as the individual or family associated with this firm from our list. This was implemented similarly to the implementation of step one, with one of the main lines being:

```
cond1 = np.array([" " + family + " " in " " + lastname + " "
                  if pd.notnull(lastname) else False
                  for lastname,
                  family in zip(families, lastnames)],
                  dtype=bool)
```

To make this more concrete: If anyone with the name Klatten or Quandt was a senior manager or shareholder at BMW or one of the known other companies, he or she would be classified as a match. This leads to more false positives the more common the name is. Therefore this was only done for names that did not contain any of the top 5% most frequent first and last names (144,135 names). Since this was only done for matched firms, firm names were not confirmed again with fuzzy matching in this step.

This step led to 6% additional findings and therefore to overall 45% coverage up to this step.

4th Step: Search first and last names in data

This step is close to the procedure for step 1 but this time it was applied to the full shareholder and director data. This comparison is computationally heavy because 30,000 names from the original list have to be compared to millions of shareholder and director names. The matching was carried out separately for countries. This means records were compared by country, which may lead to unmatched cases where individuals are only a shareholder or director in another country than that which was coded as home country in the original list.

In short, for every shareholder and director name it was checked whether the first and last name of the original list were in the first and last names of directors and shareholders. Alternatively, shareholders may contain only the last name and the word stem "fami" to be classified as a match. Matches were then confirmed again by comparing the names of firms in which the shareholder or director was involved with the firm name from the original list, from the ORBIS batch search, or from the firm names matched in step 1.

This step led to 6% additional findings and therefore to overall 51% coverage up to this step.

5th Step: Plain fuzzy matching

In this step, plain fuzzy comparison of the full name from the original list and the full shareholder and director names was performed. This procedure can detect slight deviations such as misspellings or different conversion of umlauts or other characters in either data source. This step again analyzed the full shareholder and director data, only for families that were not matched yet. The main comparison was done with the cdist function from rapidfuzz, which is a more efficiently implemented fuzzy matching package that can easily be computed in parallel (Bachmann 2022).

```
matches = process.cdist(lower_ind,
            dirnames, scorer=fuzz.WRatio,
            score_cutoff=80, workers=-1)
```

This step led to 1% additional findings and therefore to overall 52% coverage.

6th Step: Looking for foundations

Step 6 was analyzing foundation and trust data specifically. All shareholders that were classified as either "Foundation, research Institute" or "Mutual and pension fund, nominee, trust, trustee" were extracted from the shareholder data. For all names from the original list, it was checked whether there was a foundation or trust

that includes the last name from the list *and* held shares in one of the companies that were related to the family. The first part was again done similarly to previous steps by checking whether the name occurred in the foundation or trust name:

```
cond1 = [(" " + lastname + " ") in
(" " + str(x) + " ")
                        for x in country_founds["compare_names"].to_numpy()]
```

The second part was again confirmed by checking fuzzy matching of firm names as described for step 1.

This step led to 1% additional findings and therefore to overall 53% coverage up to this step.

7th Step: Search names among the largest individual global ultimate owners

In this final step of the first phase, only the largest firms in ORBIS were taken as a start. From all firms classified as "Very large company" in ORBIS, those where the GUO was an individual or family were kept. Assuming that the owners of the largest corporations are also likely to be among the global super-rich, this step allowed a looser matching procedure. It was checked whether the last name of any name from the original list was included. Then, the firms owned by the matched names were again compared by fuzzy matching as explained for step 1.

This step led to 3% additional findings and therefore to overall 56% coverage up to this step. Since this was the final step of the first phase to drive up coverage, this share is also equal to the final coverage of families.

2nd Phase: Snowball family members in the same firms

8th Step: Snowballing of family members

The logic of this main snowballing step is the following. We now have a list of directors and shareholders as a product of the past seven steps. Each of these roles is linked to one or more companies. In all of these companies, it was checked whether there was another individual shareholder or director with the same last name. This procedure was performed iteratively. Every iteration worked roughly like this:

- A datsat was generated including all found super-rich directors and shareholders linked to all of the companies they are related to

- In all of these firms it was checked whether there is another directors or shareholder with the same last name in this firm. If yes, they were added as family members.

The procedure was repeated for new findings, until none were found. For this step, as for some previous steps, dashes were removed from the last names, so double names such as Rueggeberg-Buettner become " rueggeberg buettner " and the family last name " rueggeberg " could therefore be matched successfully. This way at least last names that contain the original family name could be matched which reduces the bias excluding women with different married names.

Compared to the previous step, this step snowballed additionally 11,151 shareholding entities and 6,917 managers that belong to the known super-rich families with enough certainty.

Final steps 9 and 10: Cleaning and deduping findings

Findings were then cleaned (step 9) and de-duplicated (step 10) to get a final cleaned list of directors and managers, as well as shareholding individuals, trusts, and foundations. The cleaning steps are not presented here in detail but include distinguishing between individuals, corporate bodies, and family shares, corrected parsing of first and last names where necessary, and predicting gender based on the first name and country. The deduplication was performed with DeDupe (Dedupe.io 2022) based on different string comparison metrics, belonging to the same family, and gender. Deduplication led to the identification of 1,734 duplicate records of family members (e.g. the same name spelled differently or with or without middle names). However, deduplication is also not perfect and some individuals may still occur multiple times. For all countries, more than 21,000 family members from 6,069 families could be identified in this way. For the analysis presented in the main text, only the known German (1,272) and US (1,978) families were used as presented in Table 1 in the main text.)

3. Matching between ORBIS and political action data

A presented in the main text, six external sources for political action data were matched to the ORBIS data: individual party donations of 10,000 € or more in Germany, corporate party donations in Germany, lobbying costs to Bundestag, individual PAC contributions of $200 or more in the US, firm PACs, and lobbying costs to Congress in the US. In the following, I briefly describe how the matching took place between ORBIS and these six sources using DeDupe (Dedupe.io 2022).

Germany

Individual party contributions

Matching was performed for names of individual shareholders and names of individual managers separately. This is because in ORBIS names of managers and names of shareholders are distinct variables from distinct modules of the ORBIS database. The first rows of the director data used for matching are shown in the following table:

Table 2: Director data from ORBIS to be matched.

DM UCI Unique Contact Identifier	DM Full name	DM Last name	DM First name	City	DM Gender	German	US	frequent name	Postcode	lat	lon	latlong
P296958406	JOERG HOFFMANN	HOFFMANN	JOERG	WOLFSBURG	M	Yes	No	True	38436	NA	NA	NA
P183688004	HUSSAIN AL ABDULLA	AL ABDULLA	HUSSAIN	WOLFSBURG	M	Yes	No	False	38436	NA	NA	NA
P355641526	HESSA AL JABER	AL JABER	HESSA	WOLFSBURG	F	Yes	No	False	38436	NA	NA	NA
P144115127	BERND WALTER ALTHUSMANN	ALTHUSMANN	BERND WALTER	WOLFSBURG	M	Yes	No	False	38436	NA	NA	NA
P070822856	KAI BLIESENER	BLIESENER	KAI	WOLFSBURG	M	Yes	No	False	38436	NA	NA	NA
P415002917	DANIELA CAVALLO	CAVALLO	DANIELA	WOLFSBURG	F	Yes	No	True	38436	NA	NA	NA
P216722401	HANS PETER FISCHER	FISCHER	HANS PETER	WOLFSBURG	M	Yes	No	True	38436	NA	NA	NA
P173209626	MARIANNE HEISS	HEISS	MARIANNE	WOLFSBURG	F	Yes	No	True	38436	NA	NA	NA
P038957413	LARS JAERVKLO	JAERVKLO	LARS	WOLFSBURG	M	Yes	No	False	38436	NA	NA	NA
P389503246	ULRIKE JAKOB	JAKOB	ULRIKE	WOLFSBURG	F	Yes	No	True	38436	NA	NA	NA

Names, city, gender, and postcode stem from ORBIS. German and US are dummies whether the firm is a German or a US firm respectively. frequent_name is a dummy indicating whether the first or the last name is among the top 1% most frequent names among all names in ORBIS. This indicator is included as a predictor for a correct match because more frequent names are more likely to be false positive matches of different persons with the same name. Lat, long, and latlong stem from Amstadt (2017) who provides a large list of many German postcodes with latitude and longitude coordinates. The postcode of the firm is matched with the respective coordinates of that postcode. Unfortunately, the list is not complete and for example this specific postcode for a part of Wolfsburg is missing.

The data for German party contributions stem from Lobbycontrol (2022). The following table gives an insight into the data.

Table 3: Examples of matched ORBIS records with German individual party contributions.

DM Full name	City	German	lat	lon	DM Last name	DM First name	DM Gender	frequent name
CHRISTOPH GRÖNER	BERLIN	Yes	52.52468	13.40535	GRÖNER	CHRISTOPH	M	False
DIETMAR BÜCHER	IDSTEIN	Yes	50.2333	8.26667	BÜCHER	DIETMAR	M	False
FRANK HANSEN	SCHWABISCH HALL	Yes	49.1992	9.79056	HANSEN	FRANK	M	True
HANS PETER STIHL	REMSECK AM NECKAR	Yes	48.8667	9.28333	STIHL	HANS PETER	M	False
KLAUS ERICH GROTH	BERLIN	Yes	52.52468	13.40535	GROTH	KLAUS ERICH	M	True
CHRISTIAN KRAWINKEL	BERLIN	Yes	52.52468	13.40535	KRAWINKEL	CHRISTIAN	M	True
CHRISTIAN BURCHARD MAX OLDENDORFF	HAMBURG	Yes	53.6153	9.99269	OLDENDORFF	CHRISTIAN BURCHARD MAX	M	False
GEORG NEMETSCHEK	MUNCHEN	Yes	48.1374	11.5755	NEMETSCHEK	GEORG	M	False
RAINER OPOLKA	WENDISCH RIETZ	Yes	NA	NA	OPOLKA	RAINER	M	False
ANTONIS SCHWARZ	BERLIN	Yes	52.52468	13.40535	SCHWARZ	ANTONIS		True

Names and city stem from the original data. First and last names are extracted from the full provided name. German here is set to true for all records and indicates that a contribution was made in Germany. frequent_name again again indicates that the first or the last name is among the top 1% most frequent names among all names in ORBIS. Gender is predicted from the last name with the Python library gender guesser (Pérez 2016).

These two sources are matched using DeDupe (Dedupe.io 2022). DeDupe is a matching library exploiting supervised learning to learn which (weighted) attributes (and their interactions) predict correct matches. The library includes a workflow that suggests cases for labelling to make learning efficient. 1000 cases were labelled manually using DeDupe before the resulting learned weighted matching algorithm was applied to predict correct matches on the full datasets. As attributes, I used different forms of string comparison metrics of the different name parts, as well as some interactions between them. In addition, exact matches of gender and the city are taken into account. Finally, if coordinates are available of the city where a party donor is registered and a firm the individual holds a share in or is a manager of, the two geographic locations are compared using the Haversine Formula. The exact matching rules and metrics are presented in the following. For further details confer the documentation of Dedupe.io (2022).

```
fields = [
    {'field': 'DM Full name', "variable name": "fullname",
    'type': 'Text', 'has missing': True, 'corpus': fullnames()},
    {'field': 'DM Full name', 'type': 'String',
     'has missing': True, 'crf': True},
    {'field': 'DM Full name', 'type': 'Exact',
     'has missing': True},
    {'field': "DM Last name", 'type': 'String',
     'has missing': True, 'crf': True},
    {'field': "DM Last name", "variable name": "lastname",
    'type': 'Exact', 'has missing': True},
    {'field': "DM First name", "variable name": "firstname1",
    'type': 'String', 'has missing': True, 'crf': True},
    {'field': "DM First name", "variable name": "firstname2",
    'type': 'Text', 'has missing': True, 'corpus': firstnames()},
    {'field': "DM First name", 'type': 'Exact',
     'has missing': True},
    {'type': 'Interaction', 'interaction variables':
      ["firstname2", "firstname1", "lastname", "fullname"]},
    {'field': 'frequent_name', "variable name": "frequent_name",
    'type': 'Exact', 'has missing': True},
    {'type': 'Interaction', 'interaction variables':
      ["frequent_name", "lastname", "fullname"]},
```

```
    {'field': 'City', 'type': 'Exact', 'has missing': True},
    {'field': 'DM Gender', 'type': 'Exact',
     'has missing': True},
    {'field': 'latlong', "variable name": "latlong",
    'type': 'LatLong', 'has missing': True},
    {'type': 'Interaction', 'interaction variables':
      ["latlong", "lastname", "fullname"]}
]
```

From the 4,811,811 global director records, 937 could be matched to the 2,072 party donors with enough predicted certainty of a correct match of 50% or more. The same procedure was applied for the shareholder data. Here, 311 of the 188,074 individual shareholders could be matched to the 2,072 party donors. It needs to be stated here, that it is of course possible that matches are either false positives or false negatives. However, manual examination gave confidence at least in the precision of the result.

Firm party contributions

For the German firm contributions, I used the ORBIS batch search integrated in the ORBIS portal to match firms. A list of firm names, country, and cities were uploaded and the portal automatically generated the best match to a firm record based on string matching. As known from the literature (cf. the article "When should we believe results using ORBIS firm data?" in this dissertation), the ORBIS batch search is only as good as it can possibly be based on simple firm names. Especially issues of entity ambiguity arise and a subsidiary instead of the main firm of a corporate group may be suggested as the best match. However, since the sample used comprises a large sample of firms including subsidiaries this should hopefully not be too problematic here. Many of the entities making firm donations are actually not firms but industry associations or similar entities. Probably because of this, only 195 of the 303 corporate entities listed in the original list from Lobbycontrol (2022) could be matched. The following table shows the first matches from the resulting list:

Table 4: Examples of matched ORBIS records with German firm party contributions.

Unternehmensname	Stadt	Land	ID Nummer	Score	Matched BvD ID	Matched company name
Deutsche Vermögensberatung	Frankfurt am Main	NA	NA	A	DE6070009184	DEUTSCHE VERMOEGENSBERATUNG AKTIENGESELLSCHAFT DVAG
Verband der Chemischen Industrie	Frankfurt am Main	NA	NA	A	DE6070128333	VERBAND DER CHEMISCHEN INDUSTRIE
WI Bad Wörishofen GmbH	Bad Wörishofen	NA	NA	A	DE7330787422	WI BAD WOERISHOFEN GMBH
R&W Industriebeteiligungen	Köln	NA	NA	A	DE5190173496	R & W INDUSTRIEBETEILIGUNGEN GMBH
Sachsenmilch Anlagen Holding GmbH	Wachau	NA	NA	A	DE3070525774	SACHSENMILCH ANLAGEN HOLDING GMBH
Dr. August Oetker KG	Bielefeld	NA	NA	A	DE4010000707	DR. AUGUST OETKER KG
Evonik Industries AG	Essen	NA	NA	A	DE5110066894	EVONIK INDUSTRIES AG
Allfinanz Deutsche Vermögensberatung AG	Frankfurt am Main	NA	NA	A	DE6070504721	ALLFINANZ DEUTSCHE VERMOEGENSBERATUNG AG
Max Aicher GmbH & Co. KG	Freilassing	NA	NA	A	DE8250000394	MAX AICHER GMBH & CO. KG
METALL NRW – Verband der Metall und Elektro Industrie Nordrhein Westfalen e.V.	Düsseldorf	NA	NA	A	DE5050130535	METALL NRW VERBAND DER METALL UND ELEKTRO INDUSTRIE NORDRHEIN WESTF.
Verein der Bayerischen Chemischen Industrie e.V.	München	NA	NA	A	DE8170452949	VEREIN DER BAYERISCHEN CHEMISCHEN INDUSTRIE E.V.

Firm lobbying to Bundestag

Firm lobbying was also performed with the ORBIS patch search provided in the portal. However, the source data now stems from the German lobby register (Bundestag 2022). The following table shows some examples for matches for the data extracted from the Bundestag website:

Table 5: Examples of matched ORBIS records with firms lobbying to Bundestag.

Unternehmensname	Score	Matched BvD ID	Matched company name
LEW Wasserkraft GmbH	A	DE8030083242	LEW WASSERKRAFT GMBH
akf servicelease GmbH	A	DE5410109786	AKF SERVICELEASE GMBH
LEW Verteilnetz GmbH	A	DE8030244355	LEW VERTEILNETZ GMBH
Helsing	A	DE8171115646	HELSING GERMANY GMBH
Rheinmetall AG	A	DE5050000137	RHEINMETALL AG
VELUX Deutschland GmbH	A	DE2150041722	VELUX DEUTSCHLAND GMBH
KSB SE & Co. KGaA	A	DE7150246802	KSB SE & CO. KGAA
DocMorris Services B.V.	A	NL14105470	DOCMORRIS SERVICES B.V.
Unite Network SE	A	DE8170547447	UNITE NETWORK SE
eBay GmbH	A	DE2011143116	EBAY GMBH
KRAIBURG Relastec GmbH & Co.KG	A	DE3350002747	KRAIBURG RELASTEC GMBH & CO. KG

From the 1,114 entities registered as firms ("Unternehmen") with lobby register in 2021, 1,066 could be matched in this way and were considered for the analysis.

USA

Individual PAC contributions

Matching was performed, analogous to the matching for Germany, for names of individual shareholders and names of individual managers separately. This is because in ORBIS names of managers and names of shareholders are distinct variables from distinct modules of the ORBIS database (for examples confer the section on German individual party contributions above).

Names, city, gender, and postcode stem from ORBIS. German and US are dummies whether the firm is a German or a US firm respectively. frequent_name is a dummy indicating whether the first or the last name is among the top 1% most frequent names among all names in ORBIS. This indicator is included as a mediator for liklihood of a correct match because more frequent names are more likely to be false positive matches of different persons with the same name. If a name is very frequent, more criteria had to be matched to be classified as a correct match (esp. same city).

The data for US PAC contributions stem from OpenSecrets (2022). The following data give an example of the data:

Table 6: Examples of US individual PAC contribution records.

id	name	city	state	gender	status	firm	firstname	lastname	frequent name
k0001516259	MARIO LOVO	CORAL GABLES	FL	M	LAWYER		MARIO	LOVO	False
p0004711763	CHRISTOPHER LORENZINI	LEVITTOWN	NY	M	NOT EMPLOYED		CHRISTOPHER	LORENZINI	True
p0003908650	JAMES LYTLE	GLENDALE	CA	M	NONE		JAMES	LYTLE	True
r0001004521	ANDREW MARTIN	SPRINGFIELD	IL	M	NOT EMPLOYED		ANDREW	MARTIN	True
r0000453298	CARLOS MARTIN	TACOMA	WA	M	NOT EMPLOYED		CARLOS	MARTIN	True
r0001421567	GEOFFREY MALECHA	EUGENE	OR	M	NOT EMPLOYED		GEOFFREY	MALECHA	False
j1002298274	ROSE MALY	LOS ANGELES	CA	F	PHYSICIAN	UCLA	ROSE	MALY	True
p00049032261	RICHARD L MARTIN SHORTER	EUGENE	OR	M	NONE		RICHARD L	MARTIN SHORTER	False
p0003531523	CAROL MARTIN DAVIS	SANTA FE	NM	F	ARTIST		CAROL	MARTIN DAVIS	False
q0000314634	MILLICENT MARTIN SHULTZ	BROOKLYN HEIGHTS	OH	F	RETAIL SHIFT SUPERVISOR	CVS	MILLICENT	MARTIN SHULTZ	False

Names, city, gender, employer, and status stem from the original data. US here is set to true for all records and indicates that a contribution was made in the US. frequent_name again again indicates that the first or the last name is among the top 1% most frequent names among all names in ORBIS.

These two sources are matched using DeDupe (Dedupe.io 2022) using different forms of string comparison metrics of the different name parts, as well as some interactions between them. In addition, exact matches of gender and the city are taken into account. 1000 cases were labelled manually using DeDupe before the resulting learned weighted matching algorithm was applied to predict correct matches on the full datasets. The exact matching rules and metrics are presented in the following.

```
fields = [
    {'field': 'name', "variable name":"fullname",
    'type': 'Text', 'has missing': True, 'corpus': fullnames()},
    {'field': 'name', "variable name":"fullname",
    'type': 'Exact',
     'has missing': True},
        {'field': "lastname", "variable name":"lastname",
        'type': 'Exact', 'has missing': True},
        {'field': "firm", "variable name":"firmname3",
        'type': 'String', 'has missing': True, 'crf': True},
        {'field': "firm", "variable name":"firmname2",
        'type': 'Exact', 'has missing': True},
        {'field': 'name', "variable name":"firmname1",
        'type': 'Text', 'has missing': True, 'corpus': firmnames()},
        {'field': "firstname", "variable name":"firstname",
        'type': 'String', 'has missing': True, 'crf': True},
        {'field': "firstname", 'type': 'Exact',
     'has missing': True},
    {'type': 'Interaction', 'interaction variables':
      ["firstname", "lastname", "firmname1"]},
    {'type': 'Interaction', 'interaction variables':
      ["firmname1", "firmname2", "firmname3"]},
    {'field': 'city', 'type': 'Exact',
     'has missing': True},
    {'field': 'US', 'type': 'Exact',
     'has missing': True},
    {'field': 'gender', 'type': 'Exact',
     'has missing': True}]
```

In this case, only the 3,518,326 director records who either sit on at least one US firm's board or are US citizens are matched. This is because the number of records for individuals who made a contribution in the US is very large with more than

3m and therefore there is a higher liklihood of false positive matches. Therefore the sample of individuals is reduced to those with any company affiliation in the US. Of these 3.5m directors, 51,122 could be matched to the 3,035,142 PAC donors with enough predicted certainty of a correct match of 50% or more. The same procedure was applied for the shareholder data. Here, 9,249 of the 233,803 individual shareholders could be matched to the 3,035,142 PAC donors. It needs to be stated here, that it is of course possible that matches are either false positives or false negatives. However, manual examination gave confidence at least in the precision of the result.

Firm PACs

For the US firm PACs, the ORBIS batch search integrated in the ORBIS portal to match firms was used again. A list of firm names, country, and cities were uploaded and the portal automatically generated the best match to a firm record based on string matching. As known from the literature (cf. the article "When should we believe results using ORBIS firm data?" in this dissertation), the ORBIS batch search is only as good as it can possibly be based on simple firm names. Especially issues of entity ambiguity arise and a subsidiary instead of the main firm of a corporate group may be suggested as the best match. However, since the sample used comprises a large sample of firms including subsidiaries this should hopefully not be too problematic here. 2,640 of the 2,694 corporate entities listed in the original list from OpenSecrets (2022) could be matched. The following table shows the first matches from the resulting list:

Table 7: Examples of matched ORBIS records with US firm PACs.

Unternehmensname	Stadt	Land	ID Nummer	Score	Matched BvD ID	Matched company name
American Dental Assn	NA	US	NA	A	US128799292L	AMERICAN DENTAL ASSOCIATION
National Assn of Home Builders	NA	US	NA	A	US297542549L	NATIONAL ASSN HOME BUILDERS
US Telecom Assn	NA	US	NA	A	US131289233L	US TELECOM ASSOCIATION
Michigan State Medical Society	NA	US	NA	A	US125986819L	MICHIGAN STATE MEDICAL SOCIETY
American Hotel & Lodging Assn	NA	US	NA	A	US263722139L	AMERICAN HOTEL & LODGING ASSOC
Texas Medical Assn	NA	US	NA	A	US128991475L	TEXAS MEDICAL ASSOCIATION
Dairy Farmers of America	NA	US	NA	A	US149094366L	DAIRY FARMERS OF AMERICA INC
Alaska Medical Assn	NA	US	NA	NA	NA	NA
Business Industry PAC	NA	US	NA	NA	NA	NA
National Council of Farmer Co Ops	NA	US	NA	NA	NA	NA
Federation of American Hospitals	NA	US	NA	A	US131235088L	FEDERATION OF AMERICAN HOSPITALS INC

Firm lobbying to Congress

US firm lobbying was also performed with the ORBIS patch search provided in the portal. The following table shows some examples for matches for the OpenSecrets data (OpenSecrets 2022):

Table 8: Examples of matched ORBIS records with firms lobbying to US Congress.

Unternehmensname	Score	Matched BvD ID	Matched company name
LafargeHolcim	A	BR60869336000117	LAFARGEHOLCIM (BRASIL) S.A. (Alias: LAFARGEHOLCIM)
Arkansas State University	A	US115127301GN	ARKANSAS STATE UNIVERSITY
Nevada Northern Railway Foundation	NA	NA	NA
HNTB Holdings	A	US271813253L	HNTB HOLDINGS
Trust for the National Mall	A	US249261635L	TRUST FOR THE NATIONAL MALL
EveryLife Foundation for Rare Diseases	A	US323965091L	EVERYLIFE FOUNDATION FOR RARE DISEASES
College Football Playoff	A	US275981756L	COLLEGE FOOTBALL PLAYOFF
Nordam Group	A	US149122831L	THE NORDAM GROUP LLC
Tenable Inc	A	US355135121L	TENABLE INC
1890 Universities Foundation	A	US305866940L	THE 1890 UNIVERSITIES FOUNDATION
Alzheimer's Assn	A	US114931976GN	ALZHEIMER'S ASSOCIATION

This way, 10,679 of the 13,000 listed firms and industry associations could be matched.

4. Main and additional analyses

a. Full regression tables

Table 9: Full regression tables for table 5 presented in the main text.

	Model 1	Model 2	Model 3	Model 4
(Intercept)	334.03**	0.00	326.63**	0.00***
	(712.43)	(0.00)	(535.92)	(0.00)
de_coreness	0.77	0.70		
	(0.43)	(0.16)		
involvementSupervisory	8.38***	3.30***	1.16	1.03
	(4.19)	(0.76)	(0.18)	(0.13)
involvementExecutive	7.78***	1.84	2.17***	1.74***
	(3.02)	(0.43)	(0.30)	(0.18)
rich_controlTrue	1.02	1.73	1.08	0.72
	(0.26)	(0.17)	(0.16)	(0.13)
degree	1.07	1.14***	1.10***	1.09***
	(0.05)	(0.03)	(0.01)	(0.01)
founding_year	0.99***	0.99***	0.99***	1.00***
	(0.00)	(0.00)	(0.00)	(0.00)
revenue	1.21	1.77***	1.00**	1.00*
	(0.08)	(0.06)	(0.00)	(0.00)
listedUnlisted	0.31	0.15***	0.08***	0.12***
	(0.16)	(0.03)	(0.01)	(0.01)
fam_contTrue		0.97		0.93
		(0.33)		(0.20)
firm_cont		21.07***		80.64***
		(6.47)		(6.52)
NACE2C - Manufacturing		1298901.17		
		(4164427160.29)		
NACE2D - Electricity, Gas, etc.		2586252.82	6.97***	2.45***
		(8291825263.94)	(0.98)	(0.35)
NACE2E - Water Supply, Sewerage, etc.		324273.04	2.15*	6.48***
		(1039656838.52)	(0.70)	(0.88)
NACE2F - Construction		175985.43	0.83	0.53***
		(564229632.71)	(0.17)	(0.08)
NACE2G - Wholesale		617236.81	0.55***	0.67***
		(1978932515.19)	(0.08)	(0.06)
NACE2H - Transportation		1054872.07	2.98***	3.10***
		(3382041683.33)	(0.47)	(0.31)
NACE2I - Accommodation and Food		177329.45	0.77	0.70
		(568538702.46)	(0.20)	(0.12)
NACE2J - Information and Communication		2107143.07	2.26***	2.66***
		(6755744059.49)	(0.28)	(0.21)
NACE2K - Financial and Insurance Activities		638761.98	1.23	0.67***
		(2047944690.94)	(0.12)	(0.05)
NACE2L - Real Estate Activities		111174.15	0.72	1.21
		(356437164.55)	(0.28)	(0.22)
NACE2M - Prof., Scient. and Techn. Act.		635430.35	4.36***	2.55***
		(2037263080.20)	(0.44)	(0.19)
NACE2N - Administrative Activities		586814.65	1.79**	1.27
		(1881395556.10)	(0.32)	(0.16)
NACE2O - Public Administration		507173.89	0.30*	8.21***
		(1626058108.17)	(0.15)	(0.72)
NACE2P - Education		0.98	0.16*	4.92***
		(3145.88)	(0.09)	(0.40)
NACE2Q - Human Health and Social Work		529967.68	1.67***	3.49***
		(1699137608.27)	(0.24)	(0.23)
NACE2R - Arts, Entertainment and Recreation		2244759.93	0.75	2.39***
		(7196959638.70)	(0.27)	(0.33)

	Model 1	Model 2	Model 3	Model 4
NACE2S and T - Other Service Activities and Households as Employers		813254.00	2.82***	3.52***
		(2607386248.66)	(0.43)	(0.32)
NACE2U - Extraterritorial Organisations		0.94	0.00***	0.00***
		(3675.39)	(0.03)	(0.02)
us_coreness			1.82***	1.60***
			(0.11)	(0.10)
employees			1.00**	1.00***
			(0.00)	(0.00)
AIC	3006.87	12010.79	14610.32	34849.03
BIC	3103.67	12322.72	14920.96	35182.68
Log Likelihood	−1494.44	−5976.40	−7278.16	−17395.52
Deviance	2988.87	11952.79	14556.32	34791.03
Num. obs.	346655	346655	733175	733175

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

b. Results without imputation

For the regression results presented in the main text, missing values were simply imputed by mean and mode values over all observations in the respective sample. I decided against using multiple imputation because there is reason to believe that values are not missing at random, and results using multiple regression deviated too much from those using mean/mode imputation and complete cases analyses. As Table 2 shows, complete case analysis shows overall very similar results to the mean/mode imputation based regressions presented in the main text. However, the sample is reduced to a small share of all firms. The only relevant differing results in terms of coefficient direction and significance are that super-rich involvement in management does not show significant associations with lobbying in Germany. Also closeness to the core of the inner circle now does have a significant but negative effect for firms to lobby. This seems counter intuitive. In summary, results are similar also when looking only at complete cases which in my view gives confidence in results presented in the main text.

Table 10: Full regression tables for table 5 presented in the main text.

	Model 1	Model 2	Model 3	Model 4
(Intercept)	0.05	0.00	180813392.93***	1.30
	(0.20)	(0.00)	(302799323.40)	(1.41)
de_coreness	0.72	0.61*		
	(0.38)	(0.12)		
involvementSupervisory	3.30*	1.49	1.85**	2.00***
	(1.65)	(0.32)	(0.31)	(0.27)
involvementExecutive	3.73**	0.75	2.82***	3.16***
	(1.82)	(0.19)	(0.46)	(0.39)
rich_controlTrue	0.91	1.46	1.04	0.75
	(0.33)	(0.18)	(0.18)	(0.14)
degree	1.02	1.06***	1.07***	1.06***

	Model 1	Model 2	Model 3	Model 4
	(0.05)	(0.02)	(0.01)	(0.01)
founding_year	1.00**	1.00*	0.99***	1.00**
	(0.00)	(0.00)	(0.00)	(0.00)
revenue	1.49***	1.88***	1.00*	1.00
	(0.08)	(0.04)	(0.00)	(0.00)
listedUnlisted	0.62	0.29***	0.11***	0.20***
	(0.32)	(0.05)	(0.01)	(0.01)
fam_contTrue		0.81		1.06
		(0.33)		(0.24)
firm_cont		5.59***		25.05***
		(2.04)		(2.17)
NACE2C - Manufacturing		72694.69		
		(287593283.67)		
NACE2D - Electricity, Gas, etc.		178366.79	5.09***	1.69**
		(705651133.07)	(0.82)	(0.26)
NACE2E - Water Supply, Sewerage, etc.		19826.76	1.26	2.98***
		(78438219.17)	(0.42)	(0.46)
NACE2F - Construction		19591.45	0.37***	0.24***
		(77507289.71)	(0.09)	(0.04)
NACE2G - Wholesale		42220.95	0.30***	0.29***
		(167033675.75)	(0.05)	(0.03)
NACE2H - Transportation		112472.31	1.78***	1.52***
		(444960707.41)	(0.31)	(0.17)
NACE2I - Accommodation and Food		43047.29	0.42**	0.24***
		(170302830.64)	(0.12)	(0.05)
NACE2J - Information and Communication		140527.25	1.26	1.18
		(555951096.46)	(0.17)	(0.10)
NACE2K - Financial and Insurance Activities		60179.96	1.04	0.45***
		(238082732.61)	(0.12)	(0.04)
NACE2L - Real Estate Activities		15698.93	0.32*	0.50***
		(62107803.12)	(0.14)	(0.10)
NACE2M - Prof., Scient. and Techn. Act.		75547.88	1.96***	0.99
		(298880998.08)	(0.23)	(0.08)
NACE2N - Administrative Activities		68652.14	1.06	0.58***
		(271600223.04)	(0.21)	(0.08)
NACE2O - Public Administration		86240.93	0.02***	1.40**
		(341184637.01)	(0.02)	(0.15)
NACE2P - Education		0.47	0.08***	1.35**
		(1856.69)	(0.04)	(0.13)
NACE2Q - Human Health and Social Work		103896.23	0.64**	1.40***
		(411032166.63)	(0.11)	(0.11)
NACE2R - Arts, Entertainment and Recreation		406095.17	0.31**	0.88
		(1606585582.30)	(0.12)	(0.14)
NACE2S and T - Other Service Activities and Households as Employers		75301.61	1.33	1.48***
		(297906732.55)	(0.23)	(0.16)
NACE2U - Extraterritorial Organisations		1.03	0.00***	0.00***
		(4639.62)	(0.01)	(0.02)
us_coreness			1.55***	1.31***
			(0.09)	(0.08)
employees			1.00*	1.00***
			(0.00)	(0.00)
AIC	1231.55	5960.04	8880.25	21898.79
BIC	1314.03	6225.24	9147.77	22186.12
Log Likelihood	−606.77	−2951.02	−4413.12	−10920.39
Deviance	1213.55	5902.04	8826.25	21840.79
Num. obs.	70574	69213	148449	148449

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

I. References

Amstadt, Allan. 2017. “German Zip Codes with Coordinates and Names.” Available at: https://gist.github.com/iteufel/af379872bbc3bf5261e2fd09b681ff7e. [Accessed: March 13, 2023].

Bachmann, Max. 2022. “Rapidfuzz.” Available at: https://pypi.org/project/rapidfuzz/.

Bornefeld, Andreas. 2019. “Vermögensforschung.” Available at: https://www.reichsten.de/en/index. [Accessed: April 26, 2019].

Bundestag, Deutscher. 2022. “Lobbyregister.” Available at: https://www.lobbyregister.bundestag.de/startseite.

Cohen, Adam. 2020. “Fuzzywuzzy Python Package.” Available at: https://pypi.org/project/fuzzywuzzy/. [Accessed: February 28, 2020].

Dedupe.io. 2022. “Dedupe.” Available at: https://github.com/dedupeio/dedupe.

Gulbranson, Derek. 2022. “Python-Nameparser.” Available at: https://github.com/derek73/python-nameparser.

Lobbycontrol. 2022. “Lobbypedia.” Available at: https://lobbypedia.de/wiki/Spezial:Abfrage_ausf%C3%BChren/Parteispenden. [Accessed: March 14, 2023].

OpenSecrets. 2022. “Bulk Data.” Available at: https://www.opensecrets.org/bulk-data.

Pérez, Israel Saeta. 2016. “Gender Guesser.” Available at: https://pypi.org/project/gender-guesser/. [Accessed: March 13, 2023].

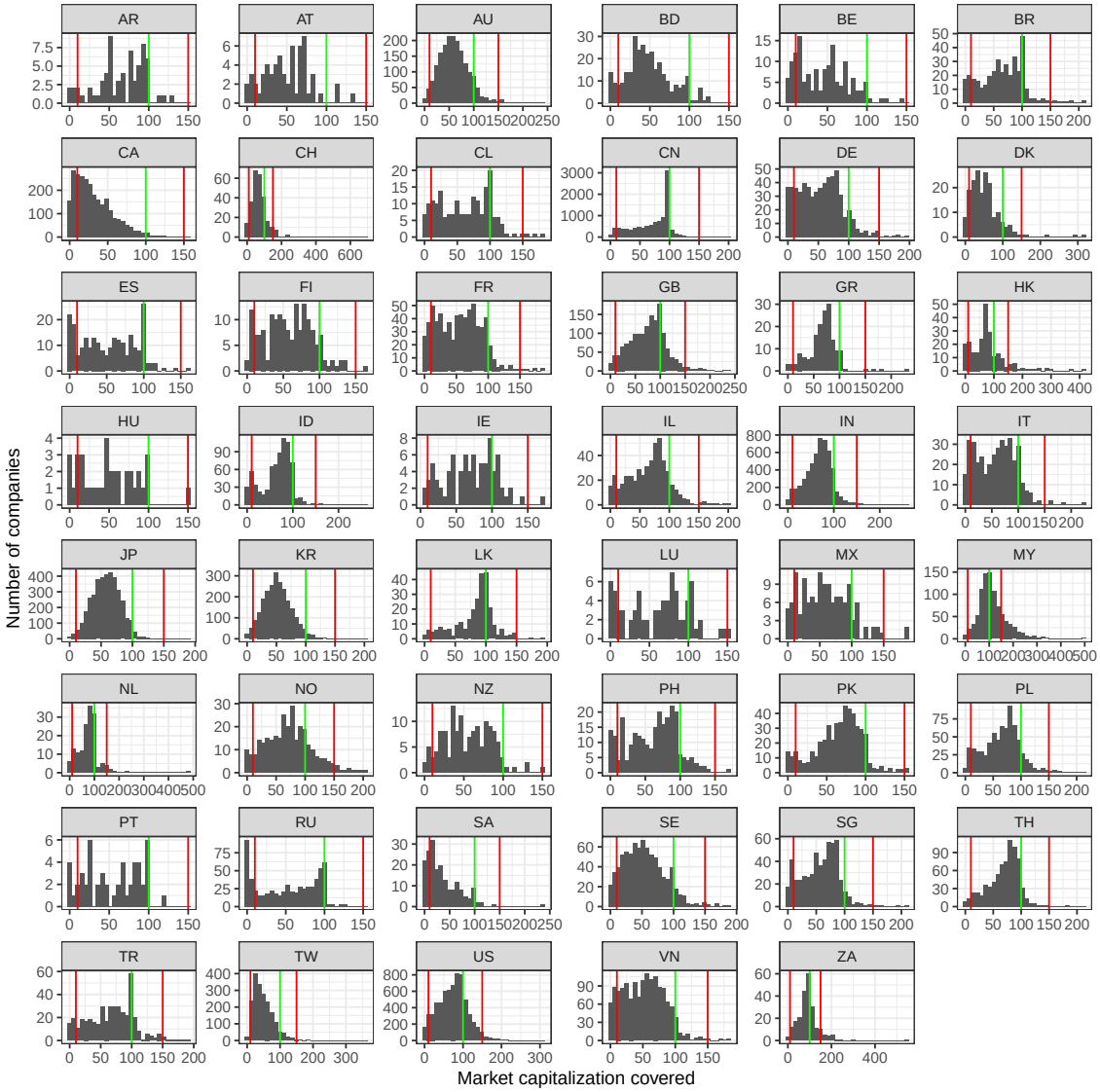
Rémy, Philippe. 2022. “Name-Dataset.” Available at: https://github.com/philipperemy/name-dataset/.

A Finance Capital 2.0 - Appendix

A.1 Coverage of shareholdings and sample cut-offs

Figure 1 shows a histogram of the distribution of the coverage of shareholdings, by firm and country. The x-axis displays the sum total of shareholder percentages reported per firm, excluding all free-float aggregates, the y-axis the number of companies for which Orbis reports X percent of shareholdings. In countries with high-quality shareholder data, the distribution is bell-shaped. In an ideal world, the density maximum would be in the region of 50-70% of shareholdings (see, e.g., Australia, Japan, or Sweden), and the bars would not exceed the green line, located at 100 percent. However, for a significant proportion of firms, Orbis over-reports, so that shareholdings sum to more than 100%. In order to mitigate this problem, we performed several algorithmic deduplication steps (see section **??**), after which we manually removed remaining duplicates for the largest shareholders by market capitalization. Unfortunately, these steps still leave 6481 companies with stakes summing up to more 100 percentage points.

Figure 1: Share ownership by main shareholder categories



Data: Orbis.

The International Max Planck Research School on the Social and Political Constitution of the Economy (IMPRS-SPCE) is a unique international three-and-a-half-year doctoral program offered jointly by the Max Planck Institute for the Study of Societies, the University of Cologne's Faculty of Management, Economics and Social Sciences, and the Faculty of Social Sciences at the University of Duisburg-Essen. Its research explores the relationship between the modern economy and its social and political foundations. Building on a long tradition in sociology and political science, the school aims to combine and develop the approaches of new economic sociology, comparative political economy, and organization studies. The Studies on the Social and Political Constitution of the Economy are a doctoral thesis series featuring dissertations by PhD students who have successfully completed the graduate training program of the IMPRS-SPCE.

MAX-PLANCK-INSTITUT FÜR GESELLSCHAFTSFORSCHUNG
MAX PLANCK INSTITUTE FOR THE STUDY OF SOCIETIES

Universität zu Köln

UNIVERSITÄT DUISBURG ESSEN
Open-Minded