european corporate governance institute

ecgi

Mavericks, Universal, and Common Owners - The Largest Shareholders of U.S. Public Firms

Finance Working Paper N° 838/2022
October 2022

Amir Amel-Zadeh
Saïd Business School, University of Oxford, CERF
Fiona Kasperk
Saïd Business School, University of Oxford
Martin C. Schmalz
University of Oxford, CEPR, CESIfo, C-SEB and ECGI

© Amir Amel-Zadeh, Fiona Kasperk and Martin C. Schmalz 2022. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

This paper can be downloaded without charge from:
https://doi.org/10.67929/ecgi-fin-838-2022

european corporate governance institute

ECGI Working Paper Series in Finance

Mavericks, Universal, and Common Owners - The Largest Shareholders of U.S. Public Firms

Working Paper N° 838/2022

October 2022

Amir Amel-Zadeh
Fiona Kasperk
Martin C. Schmalz

We thank Matt Backus, Chris Conlon, and Michael Sinkinson for making code available, and Rüdiger Fahlenbrach, Philip Bond, Kenneth Ahern, Steve Kaplan, Todd Gormley (discussant), Michael Sinkinson, seminar audiences at the Federal Reserve Bank of New York and the Universities of Bonn, Groningen, Oxford, and Rotterdam, and the audience at MaCCI, CRESSE, NBER Summer Institute (Corporate Finance), and EARIE for valuable feedback. Vladimir Fedoseev and Hady Elkady provided outstanding research assistance. Amel-Zadeh acknowledges support from the Saïd Business School Foundation and Schmalz acknowledges support from Deutsche Forschungsgemeinschaft under Germany's Excellence Strategy– EXC 2126/1–390838866.

© Amir Amel-Zadeh, Fiona Kasperk and Martin C. Schmalz 2022. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Abstract

Using a novel data set, we show that up to one-fifth of America's largest firms had an activist, non-financial blockholder, or insider as their largest shareholder in the past 20 years. Blockholders and insiders tend to be less diversified than institutional investors. Measures of ``universal'' and ``common'' ownership of firms are therefore lower than previously believed based on analyses of institutional investors' holdings alone, and the heterogeneity in ownership structures across firms is greater. Activism contributes positively to common ownership of industry rivals, as do the ``Big 3'' and consolidation in the asset management industry. We conclude that policy makers can reduce within-industry common ownership without sacrificing diversification or market indexing.

Keywords: universal ownership, common ownership, institutional ownership, blockholders, insiders, antitrust, governance

JEL Classifications: G23, G34, L21, L40

Amir Amel-Zadeh
Associate Professor of Accounting
University of Oxford
Park End Street
Oxford, OX1 1HP, UK
e-mail: amir.amelzadeh@sbs.ox.ac.uk

Fiona Kasperk
Researcher
University of Oxford
Park End Street
Oxford, OX1 1HP, UK
e-mail: Fiona.Kasperk.DPHIL@said.oxford.edu

Martin C. Schmalz*
Professor of Finance and Economics
University of Oxford, Saïd Business School
Park End Street
Oxford OX1 1 HP, UK
e-mail: Martin. Schmalz@sbs.ox.ac.uk

\

*Corresponding Author

Mavericks, Universal, and Common Owners - The Largest Shareholders of U.S. Public Firms*

Amir Amel-Zadeh† Fiona Kasperk‡ Martin Schmalz§

September 29, 2022

Abstract

Using a novel data set, we show that up to one-fifth of America's largest firms had an activist, non-financial blockholder, or insider as their largest shareholder in the past 20 years. Blockholders and insiders tend to be less diversified than institutional investors. Measures of "universal" and "common" ownership of firms are therefore lower than previously believed based on analyses of institutional investors' holdings alone, and the heterogeneity in ownership structures across firms is greater. Activism contributes positively to common ownership of industry rivals, as do the "Big 3" and consolidation in the asset management industry. We conclude that policy makers can reduce within-industry common ownership without sacrificing diversification or market indexing.

JEL Classification: G23, G34, L21, L40

Keywords: common ownership, institutional ownership, blockholders, insiders, antitrust, governance

*This working paper was previously circulated with the title "Measuring Common Ownership: the Role of Blockholders and Insiders". We thank Matt Backus, Chris Conlon, and Michael Sinkinson for making code available, and Rüdiger Fahlenbrach, Philip Bond, Kenneth Ahern, Steve Kaplan, Todd Gormley (discussant), Michael Sinkinson, seminar audiences at the Federal Reserve Bank of New York and the Universities of Bonn, Groningen, Oxford, and Rotterdam, and the audience at MaCCI, CRESSE, NBER Summer Institute (Corporate Finance), and EARIE for valuable feedback. Vladimir Fedoseev and Hady Elkady provided outstanding research assistance. Amel-Zadeh acknowledges support from the Saïd Business School Foundation and Schmalz acknowledges support from Deutsche Forschungsgemeinschaft under Germany's Excellence Strategy – EXC 2126/1–390838866.

†University of Oxford, Saïd Business School, Amir.Amelzadeh@sbs.ox.ac.uk.

‡University of Oxford, Saïd Business School, Fiona.Kasperk@sbs.ox.ac.uk.

§University of Oxford, Saïd Business School, Nuffield College, CEPR, ECGI, CESIfo, and C-SEB, martin.schmalz@sbs.ox.ac.uk

Electronic copy available at: https://ssrn.com/abstract=4059513

Disclosure statements

Amir Amel-Zadeh

I have no conflicts of interest with respect to this research. I pursue no outside activities that I believe materially influence my research activities. I do not hold paid or unpaid positions in organizations whose policy positions, goals, or financial interests are known to me to relate to this or any article I have written, refereed, or discussed. I do not run a research center that is funded by a particular interest group.

For full disclosure, I have held or may hold investment portfolios comprised of direct shareholdings, ETFs and debt products of or issued by organizations that may be affected by this research. I am non-executive director and hold a financial interest in a financial institution whose business is unrelated to this research. From time to time, I engage in compensated consulting activities with financial institutions and investment funds that may have a financial interest related to the article unknown to me. I am also board member of the UK Endorsement Board (UKEB) whose purpose is to influence, endorse and adopt new or amended international accounting standards issued by the International Accounting Standards Board (IASB) for use by UK companies.

I have in the past received financial support, in-kind support, e.g., access to data, or consultant fees from following organizations: ABN AMRO, Bank of New York Mellon, Cambridge Innovation Consulting, PanAgora Asset Management. This research does not require IRB approval and no party had the right to review the article prior to circulation.

Fiona Kasperk

I have no conflicts of interest with respect to this research. I pursue no outside activities that I believe materially influence my research activities. For full disclosure, I hold a portfolio comprised of ETFs and debt products issued by organizations that may be affected by this research.

Martin Schmalz

I hold a portfolio of ETFs and ETNs, which may bias my opinions in favor of these particular investment products or their issuers. I own a silent interest in a private manufacturing firm. I do not hold voting shares in individual public companies. To raise research funds (i.e. not

Electronic copy available at: https://ssrn.com/abstract=4059513

personal income), I have written one industry-sponsored paper that is unrelated to my other research (detailed below). I pursue no outside activities that I believe materially influence my research activities. I do not hold paid or unpaid positions in organizations whose policy positions, goals, or financial interests are known to me to relate to any article I've written, refereed, or discussed; I do not run a research center that would be funded by a particular interest group, etc. My CV lists grants and fellowships, and also discloses academic activities such as refereeing and short visits, for some of which I receive/d a nominal fee or honorarium.

Electronic copy available at: https://ssrn.com/abstract=4059513

1 Introduction

This paper documents who the largest shareholders of America's largest publicly traded firms are, and what characterizes these shareholders' portfolios. In particular, we distinguish between shareholders with concentrated stakes in one or a few firms, who we call Mavericks, and diversified shareholders. For the latter, we document to which extent their portfolios comprise other firms in the same or in different industries. By doing so, we analyze who and what drives common ownership of competitors as well as "universal ownership" of a large fraction of publicly traded U.S. equities.

Understanding corporate ownership structures and the drivers of common and universal ownership is important, because an advanced policy debate considers restrictions to shareholder rights so as to respond to concerns that common ownership increases product prices (Azar et al. (2018); Schmalz (2021)). Moreover, in considering measures to reduce the anti-competitive effects of common ownership, a debate has ensued about the causes of common ownership. While theoretical work suggests textbook indexing as one potential cause of the increase in common ownership of competitors (Rotemberg (1984)), the idea of restricting textbook indexing – perhaps the most successful financial innovation of the 20th century – seems unattractive to many policymakers and academics alike. At the same time, much hope rests on the idea that universal owners thanks to having 'a stake in the overall economy' and large stakes in individual firms, have both the incentive and the ability to induce portfolio firms to internalize ecological and social externalities (e.g., Hart and Zingales, 2017); or that common ownership across industries may actually have the effect of reducing product prices (Azar and Vives (2021)). It is thus an open empirical question whether policymakers can reduce common ownership without restricting textbook indexing and thereby reducing universal ownership with potentially unintended negative consequences.

To answer this question, we investigate the empirical drivers of common ownership and universal ownership among America's largest publicly-traded firms. In particular, we investigate to which extent common and universal ownership of S&P 500 single-class firms is driven by "passive" shareholders engaging in market-level indexing strategies and to which extent common ownership of competitors arises as a result of active strategies by mutual funds, activist hedge funds, and family offices of the world's richest individuals.

Electronic copy available at: https://ssrn.com/abstract=4059513

We first construct a uniquely comprehensive and freely accessible dataset of U.S. corporate ownership. To illustrate why answering our research question requires more comprehensive ownership records than typically used in the literature to date, consider calculating a measure of common ownership, which captures to which extent the most influential shareholders in one firm also have financial interests in the other, between three of America's largest firms by market capitalization, Amazon, Tesla, and Twitter (as of Q4 2021 / Q1 2022).

Amazon	%	Source		**Tesla**	%	Source		**Twitter**	%	Source
Jeff Bezos	9.8	F 4		Elon Musk	16.7	13-G		Elon Musk	9.7	13-D
Vanguard	6.6	13-F		Vanguard	6.1	13-F		Vanguard	8.8	13-F
BlackRock	5.7	13-F		BlackRock	5.1	13-G		Morgan Stan.	8.1	13-F
State Street	3.3	13-F		Capital Res.	3.7	13-F		BlackRock	6.5	13-F
T Rowe Price	3.2	13-F		State Street	3.1	13-F		State Street	14.6	13-F

Whereas Vanguard, BlackRock, State Street, and other asset managers indeed have holdings in all three firms, the largest and arguably most influential owners in Amazon and Tesla do not hold significant stakes in the other firm. The example illustrates the importance of considering individual insiders and blockholders when trying to assess how large or influential common owners – i.e. widely diversified institutional asset managers – are, compared to other large investors with interests in only one firm. Yet, prior research on the drivers and effects of common ownership at the firm or industry-level (e.g., Backus et al., 2021b; Koch et al., 2020; Lewellen and Lowry, 2020; Gilje et al., 2020) has mainly relied on ownership records in 13-F filings that cover institutional investors, but do not typically include the ownership records of corporate insiders and non-institutional blockholders. That is, in the above example, considering 13-F records alone would cause researchers to ignore Elon Musk and Jeff Bezos, and lead to the erroneous conclusion that there is a high and homogeneous level of overlapping ownership between Amazon and Tesla. However, in fact, the stark difference in ownership structures and the unusually *low* level of overlapping ownership between Amazon and Tesla is precisely due to the presence of Mavericks, i.e. relatively undiversified blockholders and insiders, at the top of the ownership roster.

At the same time, researchers using 13-F filings alone would fail to observe the overlapping

Electronic copy available at: https://ssrn.com/abstract=4059513

ownership holding by Musk between Tesla and Twitter. Hence, omitting ownership by non-financial blockholders and insiders can lead to erroneous conclusions in either direction. Which direction the bias goes on average is an empirical question.

Furthermore, not all institutional investors are passive owners. Perhaps surprisingly, even activist investors do not always hold only one position per industry. For example, well-known "activist" investor Bill Ackman's Pershing Square is among the largest shareholders of Burger King, Chipotle, and Domino's Pizza, contributing *positively* to the level of common ownership in that industry.[1]

These examples illustrate that relying on 13-F filings alone when calculating measures of overlapping ownership can bias the measured level and mask the true variation of overlapping ownership of firms, whether in the same industry, or across industries. This is important because using measures that mask the variation in overlapping ownership can lead to underestimating the effect of such ownership on a variety of outcomes; and to a failure to reject the null hypothesis of no overlapping-ownership effects, as well as other econometric biases.[2]

In this paper we identify the largest shareholders for all single-class S&P500 firms, examine their portfolio holdings and assess the effect of including blockholders and insiders in measures of common and universal ownership. For this we scrape and parse all ownership records from the SEC's EDGAR system for the years 2003-2020 for all single-class S&P 500 firms, and merge the ownership information thus obtained to construct ownership records at the firm level. In particular, we merge information from 13-F filings by institutional investors with 13-D and 13-G filings by blockholders, as well as Form 3, 4, and 5 filings by corporate insiders, and clean the data for duplicates and a large number of other errors. We first document which investor type, whether insider, activist, non-financial blockholder, financial institution or Big3, is the largest shareholder of a firm, of how many firms and whether their dominance varies over time. We identify activists following Brav et al. (2008) to capture the investor's likely engagement with their portfolio firms. We then examine how

[1]Pershing Square files 13-D in some and 13-F in other cases, however; management of the target firms may nevertheless perceive a 13-F position by Pershing Square differently than a similarly large 13-F position by a typical "passive" investor. ValueAct's insufficiently disclosed acquisition of an "activist" position in competitors Baker Hughes and Halliburton is another prominent example of activist common ownership (link).

[2]This concern has been shown to be relevant in practice. Antón et al. (2022) find that estimates of common ownership on managerial incentives roughly double once blockholders are included in the calculation of common ownership.

Electronic copy available at: https://ssrn.com/abstract=4059513

diversified the portfolios of the different investor types are, whether they are "Mavericks" who hold concentrated portfolios or diversified. Based on that, we calculate measures of universal ownership called "profit weights" or "kappas", which measure the extent to which the largest shareholders of one firm have a financial interest in other firms, compared to their financial interest in the base firm. We also calculate profit weights of within-industry firm pairs, thus measuring horizontal "common ownership." We are particularly interested in assessing whether and by how much profit weights are affected by the holdings of activists, non-financial blockholders and insiders.[3]

To more formally examine the determinants of universal and common ownership, we regress the measures of universal and common ownership on holdings by the different owner types as well as measures of textbook indexing from the literature. Finally, we assess how concentration in the asset management industry affects universal and common ownership measures when taking into account blockholders and insiders. For this, we calculate the increase of universal and common ownership implied by the BlackRock-Barclays Global Investors merger and test whether it predicts future levels of overlapping ownership.

Our key findings are as follows. First, ownership structures are more heterogeneous across firms than previously believed, based on studies examining only institutional investor portfolios. For example, between 10% and 20% of firms have a dominant activist, non-financial blockholder or insider among the largest shareholders. Second, whereas exceptions exist, non-financial blockholders' and insiders' portfolios tend to be much less diversified than institutional investors' portfolios. In fact, most of them are "mavericks" who hold only a single large stake in one firm. As a result, universal and common ownership levels are lower once we account for the holdings of blockholders and insiders. Third, in contrast to our expectations, we find that activists hold surprisingly diversified portfolios, frequently comprising multiple firms in the same industry. However, due to their selective diversification, activists contribute positively to common ownership. The holdings of the "Big Three" institutional asset managers also increase both universal and common ownership, over and above the level explained by textbook indexing. This finding contrasts with prior findings in the lit-

[3]Profit weights are a weighted sum of a given firm's shareholders' portfolio profits, where the weights are the shareholders' respective control shares in the firm. Profit weights capture the extent to which the most influential shareholders should like the firm to behave as if it put weight on other firms' profits in addition to its own profits.

Electronic copy available at: https://ssrn.com/abstract=4059513

erature based on 13-F records alone. Lastly, we find that consolidation through mergers in the asset management industry increases both common ownership and universal ownership persistently.

Moreover, throughout our sample period from 2003 to 2020 we find that common ownership levels are higher than universal ownership levels, implying that policy makers could reduce common ownership of industry rivals without reducing textbook indexing, which involves holding a widely diversified portfolio of all firms rather than concentrated holdings in one industry.

A policy-relevant take-away from our analysis is therefore that there need not be a trade-off between good governance aided by the presence of large blockholders and a reduction of common ownership of industry competitors. The latter was previously believed to be driven by large financial and non-financial investors, but we show that in practice non-financial blockholders tend to be "mavericks" instead. Conceptually, whether blockholders contribute to common ownership depends on the blockholders' diversification.

The finding that some active shareholders choose to hold product market competitors carries another important policy implication: anti-competitive effects of common ownership can start to be addressed without touching textbook-indexing, but by preventing active monopolization of industries by family offices, hedge funds, and conglomerates. Increases in common ownership can further be prevented by policing mergers of asset management firms, which don't create more textbook index funds, but which increase concentration of corporate control by centralizing voting and engagement across funds.

A takeaway for academic researchers is that non-financial blockholders and insiders play an important role in driving the level and variation of universal and common ownership among America's largest firms, and that omitting these investors can lead to qualitatively wrong conclusions about the level, causes and consequences of both universal ownership and common ownership. Using only institutional ownership from 13-F filings as the basis for research on overlapping ownership is therefore not an innocuous shortcut. Consequently, past findings in the literature based on institutional ownership alone should be interpreted with caution as they likely suffer from bias.

This paper contributes to the literature a uniquely comprehensive and freely available data set of all types of owners of the largest publicly traded firms in America. We thus

Electronic copy available at: https://ssrn.com/abstract=4059513

enable better measurement of universal ownership, common ownership, and its drivers.

Our paper relates to Charoenwong et al. (2022), who document that active mutual funds creating common ownership positions have higher risk-adjusted returns, vote more actively against management and thus have an incentive to soften product market competition. Our contribution relates even more closely to Backus et al. (2021b) (BCS), who offer the most recent evidence on common ownership among the largest U.S. publicly traded firms before this paper.[4] BCS also provide a great service to the profession – including but not limited to the literature on common ownership – by making the 13-F ownership data they assemble freely available. They thus helped the field overcome challenges related to 13-F data provided by Thomson Reuters through WRDS that had been popular in the literature thus far. We complement their contribution by constructing a data set along dimensions they acknowledge to be missing in their paper: "Occasionally, these controlling shareholders are inside or retail investors (e.g., the Walton family) ... it is possible to use data from SEC Forms 4, 5, 6, and 144 ... to construct industry holdings where available. Similarly, there is additional information on firm cross-holdings in 13-D and 13-G reports, which are more difficult to incorporate because they are not filed on a quarterly basis. These data are impractical to clean for analysis at the aggregate level. However, it is feasible and important to do so for case studies of particular industries as, e.g., Azar et al. (2018) do when they compute the profit weights for airlines and as Backus et al. (2021a) do when they compute the profit weights for cereal." (Backus et al., 2021b, footnote 12).

Our contribution is to undertake the endeavor to scrape, parse, and clean also the remaining SEC filings, and to understand the extent to which incorporating these filings changes the measurement of common ownership, for the same subset of firms BCS study. A limitation of our analysis relative to BCS is that our data starts only in 2000. Hence, our paper and data set does not substitute for theirs, but offers a complement.

In addition, as part of the data set, we contribute a newly parsed data set of 13-F records, which we believe to be superior to extant data sets.[5]

[4]Earlier papers documenting the secular increase in common ownership, also based solely on 13-F data, are reviewed in Schmalz (2018).

[5]Due to differences in our parsing methodology, in 55% of filings the Central Index Key (CIK) we extract is different from the CIK BCS extract. We end up with 55% more CIKs than BCS. For details, see Appendix A. Other significant differences between the 13-F-only analysis presented in this paper and BCS' analysis are due to differences in how we consolidate asset managers' holdings (e.g. we consolidate holdings of BlackRock and BGI only at the time when the merger was consummated, as opposed to throughout the sample). We

Electronic copy available at: https://ssrn.com/abstract=4059513

2 Data set construction

Building on the method of Backus et al. (2021b) we scrape and parse the ownership of S&P 500 firms as reported on filings required by the Securities Exchange Commission (SEC). Different from Backus et al. we incorporate not only filings by large institutional investors, but also filings by corporate insiders and blockholders, utilizing all ownership reports for investors in US public corporations and available on the SEC Edgar Archive. The fraction of stock ownership that remains not captured by any filing is attributed to retail investors.

Specifically, we parse six different SEC filings. Table 1 provides an overview of the key attributes of each filing type. More detail can be found in the data appendix. To summarize, 13-F filings are required on a quarterly basis from institutional investment managers with more than $100 million assets under management holding equity securities (and certain equity options and warrants) that trade on a U.S. exchange. A 13-F filing lists all securities owned by the institutional investor at the end of the quarter. Investors who acquire more than 5% of any equity security of a company are required to file a form 13-D within 10 days after the acquisition. If such blockholders acquired the equity securities *"not with the purpose nor the effect of changing or influencing the control of the issuer"* (17 CFR §240.13d-1b.1.i), they are allowed to report any ownership change using the shorter form 13-G, which, for example, does not require the investor to disclose their intentions of acquiring the beneficial ownership.

Lastly, the SEC requires corporate insiders, such as officers and directors as well as beneficial owners of more than 10% of outstanding equity and several other groups of insiders to report transactions in company equity securities using Form 4. A person who becomes a reporting person under the Form 4 requirements will initially file a Form 3 report, indicating the person now classifies as a corporate insider. A Form 3 need not yet report any equity ownership. Subsequently, whenever insiders change their holdings of company equity, they are required to file a Form 4 within 2 days of the transaction. Within 45 days of the end of a company's fiscal year corporate insiders must report all transactions they did not previously report, on a Form 5. Corporate insiders who failed to file a Form 3 or Form 4 report for transactions in the previous fiscal year, are required to report these on the Form 5, too.

detail these differences throughout the paper and the appendix.

Electronic copy available at: https://ssrn.com/abstract=4059513

Form type	Owner	Filed electronically	Frequency
13-F	Institutional investors with >$100 million in AUM	since 1999	quarterly, within 45 days of each quarter end
13-D	Investors acquiring >5% of an equity share class	since 1998	within 10 days of crossing the 5% threshold and for any acquisition/disposition of 1% or more thereafter
13-G	Investors acquiring >5% of an equity share class that are	since 1998	
	(a) exempt		(a) within 45 days of the end of the calendar year
	(b) qualified institutional		(b) as of the end of the calendar year or month (if >10%)
	(c) passive investors		(c) within 45 days of the end of the calendar year or month (if >10%)
3	insiders and shareholders owning >10% of an equity share class	since 2003	within 10 days of becoming insider
4	insiders and shareholders owning >10% of an equity share class	since 2003	within 2 days of transaction
5	insiders and shareholders owning >10% of an equity share class	since 2003	only if failed to file Form 4 within 45 days of fiscal year end

Table 1. Short summary of SEC filing requirements

Unlike 13-F, 13-D and 13-G filings, Forms 3, 4 and 5 not only report the overall number of shares owned by a corporate insider, but also other information related to an insider's transaction. Insider ownership is additionally classified into direct and indirect holdings, where the latter reflects that the insider holds shares through a trust or a family member.

Our data starts in 2003 Q3 as SEC regulations only required 13-F, 13-D and 13-G reports to be filed electronically from 1998/99 and only required Forms 3, 4 and 5 electronically from 2003. Due to the limited number of insider reports filed before June 2003, we focus our analysis exclusively on the period from July 2003 until December 2019.

One important intermediate step in collating the ownership from 13-F filings is to consolidate separately filed 13-F reports of incorporated funds of large institutional investment managers. We detail in Appendix A.2 how our approach differs from the methodology in Backus et al.. Once completed, the parsed data provides the Central Index Key (CIK) iden-

Electronic copy available at: https://ssrn.com/abstract=4059513

tifying the filing owner, the CUSIP number identifying each security owned, a description of the security (e.g. "call option", "common stock"), the quantity of shares owned and the market value of the owner's holdings.

For blockholder and corporate insider reports we develop two respective parsing methods coded in Python. The 13-D and 13-G filings report the CIK of the main filing owner, the CUSIP of the security owned, a description of the security, and the amount of shares owned and voted on directly or in a shared facility.

One difficulty arising in 13-D and 13-G parsing (see Appendix A.3 for more details) is that one report can be filed by multiple reporting owners. For example when the Walton siblings report their holdings in Walmart, this 13-G will be filed jointly by Robson, John and Helen Walton, as well as the Walton Family Trust. Each sibling owns and votes some stocks directly, but they share the voting power of stocks held in the Family Trust. Hence, the number of securities owned with shared voting rights is reported several times. We deal with this duplication by first parsing each reporting investors' individual direct and shared securities ownership. In the Walmart case, this will produce four observations, one for each sibling and one for the trust. Ultimately, we keep the largest aggregate amount of shares reported by any one reporting person and discard all other reported shares. This yields a conservative estimate of the stock ownership allocated to the main filing owner.

The Form 3, 4 and 5 filings report detailed information about each transaction made by corporate insiders in their company's equity (and derivative) securities. For each transaction the security, the number of shares transacted, the ex-post number of shares owned and the nature of ownership (direct or indirect via s trust or spouse) by the insider is reported as an individual entry. To capture the insider's aggregate ownership after the transaction, we extract the title of the security and the residual number of shares owned for each transaction and then discard all entries, except for the last one for each nature of ownership. For example if Bill Gates reports two transactions on a Form 4, one reporting on shares he owns directly, the other on shares he owns indirectly via the Bill & Melinda Gates Foundation, then we keep the ex-post number of shares owned for both entries and add them up, to reflect the aggregate number of shares owned by Gates after the transactions.

A further complication arises because corporate insider ownership reports differ from 13-F and 13-D/-G reports in that they do not identify each security by a CUSIP. Instead,

Electronic copy available at: https://ssrn.com/abstract=4059513

equity securities and derivatives are reported in separate tables and the filer provides a non-standardized, textual description of the security transacted, such as "call option", "common stock", "equity", featuring various typos and ambiguities. While the CUSIP and security title reported in 13-F, 13-D and 13-G reports allows us to differentiate between common stock and other securities relatively easily, this is not always the case for the free-text description in insider filings. To retain only holdings of common stock in our data set, we deploy an extensive text cleaning, filtering and interpretation routine to the security description.[6] This routine is described in more detail in Appendix A.4. Following BCS, we only keep common stock holdings. We then merge the security description with the company's trading CUSIP identifier for the given quarter. The resulting output includes the filing owner's CIK, the owned security's CUSIP, a cleaned security description and the number of shares owned.

After cleaning and identifying holdings from all filing types (large institutional, insider and blockholder ownership), we aggregate the holdings of each owner in each company, facing two more challenges. First, stock holdings for a given owner, security and point in time can be filed in multiple types of reports. For example, the 13-G filed jointly by the Walton Family Trust and the Walton siblings will duplicate shares that are reported by the Walton siblings in their individual Form 4 filings. To avoid double-counting, we first remove different types of ownership reports filed by the same owner CIK in the same quarter. Because in some cases the owner CIK will not be the same across different filing types (in the Walton example the 13-G is filed by the Walton Family Trust and the Form 4 is filed by the individual sibling, where each party has its own CIK) we then remove stock holdings of exactly equal size across different file types for the same firm in the same quarter.

The final step requires adjusting for the different filing frequencies. While 13-F reports are filed quarterly, blockholder and insider reports are required only when transactions in the underlying security are made. To fill the gaps in blockholder and insider ownership between transactions, we forward fill the holdings reported by a given owner in a given security until the next reported ownership for this owner-security combination. The rationale is that ownership stakes should not change in the meantime without a report being filed. Having removed any quarterly duplicates of filings for the same owner-security combination in the

[6] While Form 4 filings separately report transactions in derivative and non-derivative securities, the type of non-derivative security is generally only described in a text comment or footnote.

Electronic copy available at: https://ssrn.com/abstract=4059513

previous step, this interpolation will not create any new duplicates. Nevertheless we check for potential duplication of ownercik-permno holdings again and eliminate them if they arise.

If at the end of our sample (1 January 2020), the most recent filing reported a non-zero ownership stake, we forward-fill this ownership stake until the end, for a maximum of four quarters. We then spot check whether the resulting ownership structures are economically sensible and cross-check our results against various commercial data bases for consistency.

The completed data sample contains ownership structures of 742 firms and slightly less than 500 firms each quarter.

3 Results

3.1 America's largest shareholders and their portfolio interests

Across our subset of S&P 500 firms we capture the ownership of 70-90% of outstanding equity. To evaluate who the largest shareholders are and what their portfolio incentives look like, we create the following categories: the Big 3, comprising BlackRock, Vanguard and StateStreet; activist investors, identified using the list of investors with active engagements between 2000-2016 by Brav et al. (2008) and Brav et al. (2010); asset managers, comprising all 13-F filers excluding the Big Three and those classified as activists; corporate insiders, who file Form 3, 4 or 5 reports excluding those already identified in one of the previous categories; and the remaining unclassified owners who are predominantly non-financial blockholders. These categories allow us to identify individual investors and their investment motivations better than categorizing them by filing type (i.e., 13-F filers, 13-D/G filers and Form 3/4/5 filers). This is because some owners report their ownership holdings on multiple filing types. For example Carl Icahn pre-dominantly files 13-F reports but also 13-D activist and Form 3/4/5 insider reports. Hence it is difficult to classify him exclusively as one particular filer. Instead, we classify him as activist based on his prior active engagements.

Figure 1 plots the cross-section of firms' ownership structures on 1 January 2020, with each firm represented by a bar. The bar height reflects the share of equity for which we can identify the owner, and the colored parts of each bar depict the aggregate amount of equity owned by each type of investor. The largest aggregate stake in most firms is held by asset managers (excl. the Big 3), amounting to 30-50% of common stock across firms.

Electronic copy available at: https://ssrn.com/abstract=4059513

This share was even larger at the beginning of our sample (see Figure A.1), but has declined due to the growth of Big 3 ownership. In 2020 the Big 3 hold between 15-25% of equity of all S&P 500 firms. Together, all 13-F filing institutional investment managers (incl. the Big 3) constitute between 60-80% of the equity ownership of most S&P 500 firms. As some non-financial blockholders and activists also file 13-F reports, using Thomson Reuters or other 13-F-only ownership data is a good approach when we are interested in the investors owning the largest fraction of equities. However, this approach misses important insiders and blockholders, who are responsible for most variation in the cross-section of ownership structures (see Figure A.2).

Activists are present in almost all sample firms owning between 1-10% of equity. Since activists like Value Act or Pershing Square can utilize a small ownership stake with great effect on a firm's performance or value, this is an important presence for corporate governance analysis. Ownership stakes of corporate insiders and non-financial blockholders are very small or non-existent in most firms. In the few firms where they do exist, however, their ownership reaches 10-50% of common stock. Moreover individual blockholders are by definition large owners with stakes $> 5\%$ of equity suggesting that the number of investors who, in aggregate, own 10-50% is likely small (see Figure A.4). Thus, a large part of the cross-sectional variation in ownership structures comes from insiders and non-financial blockholders, who, together with their incentives and level of diversification, have been ignored in many previous studies of corporate owners. Furthermore, in the time series the average and aggregate ownership stakes of insiders and non-financial blockholders are relatively persistent, whilst the holdings of asset managers and activists increase over time (see Figure A.3).

Before we evaluate the portfolio interests of our investor types, we identify who is the individual, largest shareholder of our sample firms. The largest investors will be the most influential for corporate voting outcomes and, especially if they are insiders or activists, will be able to influence corporate strategies materially to align with their own interests. Analogous to the aggregate ownership structures, asset managers, including the Big 3 are the dominant owners of almost 90% of S&P 500 firms in our sample (see Figure 2). Over time the balance between dominant asset managers other than the Big 3, and the Big 3 shifts. After the BlackRock-Barclays Global Investors merger in 2009, BlackRock, Vanguard

Electronic copy available at: https://ssrn.com/abstract=4059513

and State Street grow markedly and become the dominant shareholders of most firms. Each of the Big 3 owns on average 4-9% of the common stock of the largest U.S. firms in 2020 (Figure A.5), which makes one them the largest owner in almost 75% of our sample firms, and Vanguard on its own the dominant shareholder of more than 60% of firms (Figure A.6). About 10% of sample firms, however, are dominated by an insider, activist or non-financial blockholder. Insiders dominate more firms than blockholders or activists, but the joint dominance in a tenth of the sample is persistent between 2003-2020. We therefore have a non-negligible number of firms, where an engaged owner steers corporate activities according to his personal financial interests. The question is, what are these financial interests?

To identify the likely financial interests of our investor types, we differentiate them by how diversified their portfolio holdings are.[7] The first diversification class are "maverick" investors who hold the equity of only one firm at a given point in time. Mavericks have no obvious financial interest in other firms. All other investors who hold equity in more than one firm at a given point in time and therefore have a financial interest in the performance of multiple firms are referred to as "diversified" investors. Among diversified investors we distinguish between two special types of diversification. "Universal" investors own shares in almost all firms in our sample at a given point in time (specifically more than 95% of firms). With a financial interest in all firms, these universal owners could have the desired beneficial effect of reducing negative ecological and social externalities (Hart and Zingales (2017)). In contrast, "common" investors are diversified across multiple, but not all firms, and importantly own shares in multiple firms within the same industry (primarily identified by SIC-4 digit codes). [8] Common but non-universal ownership may give rise to the potential anti-competitive effects discussed in the literature on common ownership (Schmalz (2021)).

Averaging across all investors of a certain type (activists, insiders etc.) and across all quarters in our data, we measure what proportion of each investor type has maverick, diversified, universal and common portfolios. Figure 3(B) depicts two dominant combinations. Asset managers, the Big 3 and activists are mostly diversified, while corporate insiders and non-financial blockholders are mostly mavericks. Unsurprisingly all asset managers, including the Big 3, are diversified and almost all own multiple firms in one industry (common

[7]This differentiation is not perfect, as we cannot measure all portfolio holdings of individual investors, such as personal investments in bonds, ETFs, pension funds or in equities of non-employer firms

[8]for robustness we also use Hoberg and Phillips (2016) TNIC-3 product market competition indicators.

Electronic copy available at: https://ssrn.com/abstract=4059513

owners). However, whereas the Big 3 are also universal owners in every quarter, only half of the remaining asset managers hold stakes in all sample firms. Based on this assessment, the Big 3 contribute to overlapping ownership across all firms, while other asset managers contribute to general overlapping ownership, but potentially more to within-industry overlapping ownership. In contrast, 85% of insiders and 45% of non-financial blockholders are undiversified mavericks, who reduce overlapping ownership. The remaining 15% of insiders are diversified, but not within industries, while the remaining half of non-financial blockholders is split between diversified non-common owners (diversified only across industries) and common owners (diversified across and within industry). Hence, while insiders do not contribute to within-industry common ownership at all, some non-financial blockholders do contribute to it. Most surprising is the diversification of activists. Less than 10% of them hold shares in only one firm or are purely diversified across industries in a given quarter. All other activists are diversified also within industries or are universal investors. This finding contrasts with our ex-ante belief that activists are highly concentrated investors with a few stakes spread across industries. They turn out to own all S&P 500 firms in 10% of our sample and otherwise own multiple firms in a given industry, contributing to overlapping ownership within industries and across all firms. Over time the share of activists that are diversified and common owners is decreasing as more activists become universal owners (Figure A.7). The diversification of all other types of owners does not vary over time.

So far we have established that asset managers and the Big 3 are the largest individual and aggregate owners of most S&P 500 firms, while corporate insiders, activists and non-financial blockholders dominate about 10% of firms. In most cases, the largest ownership stake is therefore owned by highly diversified, universal and common owners who contribute to overlapping ownership across firms, within and across industries. Only 10% of firms are dominated by a maverick or diversified non-common investors who either reduce overlapping ownership or contribute primarily to cross-industry overlapping ownership. In contrast to our ex-ante beliefs activists are present with relatively small stakes in all firms and mostly increase overlapping ownership, across *and* within industries.

Electronic copy available at: https://ssrn.com/abstract=4059513

3.2 Universal and common ownership

To investigate the level and variation of universal and common ownership across firms we calculate the Edgeworth coefficient of effective sympathy between two firm's ownership structures. This coefficient, also referred to as "kappa" or "profit weight", measures to which extent the most influential shareholders in firm A are invested and therefore have a financial interest in firm B.

Following Grossman and Hart (1979) and Rotemberg (1984) we derive the Edgeworth coefficient of sympathy from a firm objective function maximizing shareholder value in the presence of overlapping ownership. Shareholders' cash flow rights are identified by the fraction of outstanding equity they own (β). Their voting rights (γ) are a function of the respective ownership stake in a given firm. Backus et al. (2021b) clarify the assumptions under which the objective function Q_a of firm a as a function of its shareholders' portfolios x can be re-written as the sum of own firm profits and weighted other firms' profits. The latter weights are the Edgeworth sympathy coefficients, or profit weights, labeled κ_{ab}. These profit weights are calculated for each firm-pair and serve as our measure of universal ownership. We study common ownership restricting our measurement to the profit weights of firm-pairs consisting of industry competitors (identified by SIC 4-digit codes or Hoberg and Phillips (2016) TNIC 3 product market competitor codes).

$$Q_a(x_a, x_{-a}) \approx \pi_a + \sum_{b \neq a} \underbrace{\left(\frac{\sum_s \gamma_{as}\beta_{bs}}{\sum_s \gamma_{as}\beta_{as}} \right)}_{\kappa_{ab}} \pi_b \tag{1}$$

Following most of the literature we assume proportional control, hence one share owned conveys one vote and $\beta_{bs} = \gamma_{bs}$ in the above equation. For robustness we also analyze common and universal ownership with alternative control assumptions. This is discussed in the robustness section 3.3.

Averaging across all firm-pairs (within industry for common ownership) for each quarter, we find that common ownership is about 10% higher than universal ownership (Figure 4). Described more intuitively, the average profit weight attributed to industry-competitors is 10% higher than the average profit-weight attributed to any S&P 500 firm. Both common and universal ownership measures have increased markedly over our sample time, from an average

Electronic copy available at: https://ssrn.com/abstract=4059513

weight of 0.42 to 0.69 for a generic firm and an average weight of 0.47 to 0.75 attributed to an industry-competitor. The difference between average universal and common ownership profit weights is almost constant between 2003 and 2020, suggesting that common ownership has not been increasing more, recently. However, plotting only the average profit weight for each quarter prevents an analysis of the underlying variation and distribution of profit weights across firm-pairs. The difference in average universal and common ownership profit weights could for example be driven by the tails of the distribution or by outliers. Looking at the distribution of firm-pair profit weights for all firms versus for industry competitors in Figure A.10(B) demonstrates that this is not the case. Instead, the distribution of common ownership profit weights is shifted to the right with respect to the universal ownership profit weight distribution.

Following Backus et al. (2021b) we also evaluate the prevalence of profit weights larger than one for universal and common ownership. A kappa larger than 1 means firm a shareholders care more about firm b profits than about firm a profits, implying an incentive to tunnel profits from one firm to the other. As evident in Figure A.11, across all firm-pairs such tunneling incentives exist in about 4% of cases in 2003, rising to 9% in 2020. Only considering within-industry firm-pairs, such incentives rise from 4% of cases to 13% of cases.

From our previous discussion of the largest shareholders and their portfolio diversification, we know that the Big 3, asset managers and activists will likely contribute positively to universal and common ownership, while corporate insiders and non-financial blockholders limit the extent of overlapping ownership. The previous literature, however has argued that the Big 3 do not contribute to overlapping ownership once we account for investor indexing (Backus et al. (2021b)). Activists on the other hand have rarely been associated with widespread overlapping ownership in the literature. In the following we therefore investigate to what extent the different types of investors contribute to universal and common ownership, when controlling for indexing and the overall shareholder structure.

This type of analysis necessitates complete ownership data, as the previously used Thomson Reuters 13-F data or the data set provided by Backus et al. (2021b) only comprise large financial institutions or asset managers, neglecting the systematically less diversified insiders and non-financial blockholders. Comparing our average and cross-sectional profit weights based on complete ownership data with those calculated using only 13-F institutional own-

Electronic copy available at: https://ssrn.com/abstract=4059513

ership reports, we find that such incomplete data over-estimates both universal and common ownership profit weights across the whole sample (Figures A.12(A), A.13(A)).

We analyze the contribution of the Big 3, activists, insiders and non-financial blockholders to universal and common ownership with linear regression analysis. The firm-pair profit weights are regressed on the aggregate holdings of each type of investor in the base firm (e.g. regress κ_{ab} on aggregate Big 3 + activist + insider + non-financial blockholdings in firm a), controlling for investor indexing (using the measure suggested by Backus et al. (2021b)), market capitalization, operating profits and retail share of ownership. To control for persistent differences in the ownership structures of firms a and b and for period-specific economic conditions that alter investing habits, we apply firm-pair and quarter-year fixed effects.

Table 3 shows the regression results for universal profit weights in the first three columns and for common ownership profit weights in columns four to six. Insiders and non-financial blockholders contribute significantly negatively to both universal and common ownership, as their maverick portfolio structure suggests. Activists instead contribute significantly positively to both universal and common, once we control for investor indexing (which negatively confounds the relationship between activists and overlapping ownership if not controlled for, see correlations in Table 2). Finally, the Big 3 contribute significantly positively to both universal and common ownership, too, despite controlling for investor indexing. This result contrasts with the results of Backus et al. (2021b) presented in column 3a of Table 3, who find a negative correlation between the Big 3 holdings and universal profit weights, once indexing is taken into account. Further investigation has showed that our result is not due to the added blockholder and insider filings, but instead is driven by our more accurate parsing of ownership reported on 13-F filings.

Comparing the coefficients on investor holdings in the common versus the universal ownership regression, insiders reduce common ownership more than universal ownership, while the coefficients are almost the same for non-financial blockholders. Activists contribute more positively to universal than to common ownership, while the opposite holds for the Big 3. This supports our previous analysis of the size and diversification of investor types: few insiders are diversified and none are common owners, suggesting some contribute to universal profit weights but none to common profit weights. An equal proportion of non-financial

Electronic copy available at: https://ssrn.com/abstract=4059513

blockholders is diversified across and within industries, yielding an equal effect on both profit weights. Activists are highly diversified and frequently common owners but less so than asset managers (omitted category here), hence they contribute less to common ownership profit weights than to universal profit weights. Only the effect of the Big 3, who are both more frequently common and universal owners than the remaining asset managers, is ambiguous ex-ante. Based on the regression results, the Big 3 are more diversified or hold larger positions among industry competitors, leading to a larger coefficient for common than for universal profit weights.

Our findings are relevant for academia and policy makers, as they inform the debate on steps to take regarding the growing and potentially harmful extent of common ownership. In contrast to Backus et al. (2021b) we find that the Big 3 asset managers contribute positively to common and universal ownership over and above what is implied by investor indexing. Hence, the centralization of voting power and financial interests across various index funds at the family level seems to enhance overlapping ownership beyond what an index portfolio requires. Second, we show that activists are, in contrast to our ex-ante beliefs, highly diversified, within and across industries, and contribute positively to both universal and common ownership. Hence, if one wanted to reduce the extent of common ownership without touching indexing and the Big 3, one may find a way to do so by limiting the within-industry diversification of hedge funds instead.

Another potential driver of universal and common ownership are mergers of asset managers, that consolidate holdings and voting power in one investor. In the literature such mergers are frequently used as instruments to investigate the effect of overlapping ownership on corporate outcomes (Lewellen and Lowry (2020)). To analyze the role of asset manager mergers in driving universal and common ownership we focus on the effects of the BlackRock - Barclays Global Investors (henceforth BLK and BGI, respectively) merger in 2009. Both BLK and BGI were universal owners of the S&P 500 firms before the merger, hence the joint entity does not create new common or universal owners. Instead it enhances overlapping ownership via larger holdings by the joint entity and thereby increases implied kappas (Figure A.14(A)).

We compare the merger-implied change in profit weights to the actual change in profit weights post merger to gauge if there is a persistent effect on overlapping ownership (Figure

Electronic copy available at: https://ssrn.com/abstract=4059513

5). Implied profit weights are calculated based on the ownership structures in 2009 Q1 by adding up the stakes of BlackRock and BGI to form the hypothetically merged portfolio. The implied change is the difference between implied and pre-merger (2009 Q1) profit weights, the actual change is the difference between post-merger (2010 Q1) and pre-merger profit weights. The scatter plot does not show a clear positive association between the implied and actual change in profit weights. Neither can we see a clear positive sign in the distribution of actual profit weight changes as we can for implied changes (Figure A.14(A)).

To establish whether the merger leads to a persistent increase in universal and common ownership we regress the post-merger actual change in profit weights on the merger-implied profit weight changes for each firm-pair. The main specification uses the actual profit weight change between 2010 Q1 and baseline 2009 Q1 as endogenous variable. The immediate merger effects on profit weight changes by 2009 Q4 and the more persistent effects by 2011 Q1 and 2012 Q1 are also evaluated. We include an intercept to control for time fixed effects (as we are looking only at one quarter in each regression) and control for firm fixed effects, to eliminate any persistence in firm-specific stakes of BLK and BGI. The remaining variation in post-merger profit weight changes is driven solely by the cross-firm variation in BlackRock and BGI ownership stakes that are combined.

Our results in Table 4 suggest that implied profit weight changes are a statistically significant predictor of the post-merger profit weight changes in the immediate Q4 2009 quarter, but also for longer term changes until Q1 2012. This qualitative finding holds both for common ownership profit weights and universal ownership profit weights. Quantitatively the relationship between implied and post-merger profit weights is larger for industry competitors, suggesting that in our specific example, where both parties were universal owners already before the merger, the joint entity amplified common ownership more than universal ownership.

For robustness we also investigate the discrete change in implied profit weights. Regressing the post-merger profit weight changes on an indicator variable that is 1 for the top third of changes in implied profit weights due to the merger. The treatment dummy is also significantly positively correlated with post-merger profit weight changes in all periods.

In addition to our previous findings, we show in this section that mergers among asset managers are another source of common and universal ownership growth. An alternative

Electronic copy available at: https://ssrn.com/abstract=4059513

method to reduce overlapping ownership would therefore require restrictions on asset manager mergers. In our example the positive effect of asset manager mergers is due to an increase in the holdings of one already large, universal and common owner. In other mergers another effect from the creation of new common or universal ownership may play an even more important role.

3.3 Robustness

In the preceding analysis we rely on the assumption of proportional control to calculate our measures of universal and common ownership. This assumption dominates the literature estimating profit weights and is founded theoretically by a the outcome of a standard voting model. To test whether our results are sensitive to this specific control assumption we recalculate our profit weights once attributing more control to larger owners ($\gamma_{bs} = \beta_{bs}^2$) and once reducing the influence of larger owners ($\gamma_{bs} = \beta_{bs}^{0.5}$). Untabulated results reveal that the trends in universal and common ownership, as well as the relatively higher level of common compared to universal ownership are the same for both assumptions. The results of our regression analysis also hold.

Another key component of the analysis of common ownership, is the classification of industry competitors based on SIC 4-digit industry codes. These industry assignments are frequently imperfect or missing, hence we also conduct the complete analysis using the Hoberg and Phillips (2016) TNIC-3 classification of product market competitors. The results on trends and levels of universal versus common ownership (Figure A.15) and the regression results for drivers of common ownership and merger effects (Table A.1, A.2, A.4) are unchanged by this choice.

4 Conclusions

We demonstrate that the size and portfolio diversification of 13-F filing institutional investment managers and activists on the one hand and those of non-financial blockholders and insiders on the other hand are systematically different. Because blockholders and insiders are the largest shareholders of 10-20% of S&P 500 firms, omitting such owners will lead researchers to mismeasure the variation in ownership structure and likely lead to bias

Electronic copy available at: https://ssrn.com/abstract=4059513

in analyses of the effect of corporate ownership structures on various corporate outcomes. Adding these less diversified owners to the regularly analyzed sample of 13-F institutional ownership reduces estimates of common, within-industries ownership and of universal ownership consistently. Based on a complete data set of corporate ownership we find more heterogeneous ownership structures and investigate the drivers of universal and common ownership. We find that common ownership within industries is higher than universal ownership throughout our sample, suggesting that it is possible to reduce common ownership without reducing index investing, which is positively correlated with both. Instead, family offices as well as active mutual funds and hedge funds, including those commonly understood to be "activists," induce more common ownership among competitors and may be the lever to limit such within industry overlapping ownership.

Moreover, we find that the shareholdings of the "Big Three" institutional investors also contribute positively and statistically significantly to common and universal ownership over and above the effect of market-indexing, which is in contrast to previous findings in the literature. This implies that the consolidation of portfolios managed by sub-funds of the Big Three to the family level gives rise to aggregate ownership and voting power that increases common and universal ownership beyond the level necessary for index diversification. Relatedly, mergers in the asset management industry predict future increases in both common and universal ownership. These findings indicate that regulators can reduce common ownership of competitors without either sacrificing diversification or "universal ownership," but by addressing actively chosen common ownership positions and by scrutinizing consolidation in the asset management industry.

The data set constructed should also be interesting for research that can rely solely on filings by institutional investors (13-F filings), including the finance literature, because the data set we provide is more comprehensive. In particular, we parse 55% more owner-CIKs from the SEC filings than prior work. Therefore, we propose that future research use the data set of ownership records used in the present paper, which we make freely available for academic use.

That said, this data set is unlikely to be free from errors, and new mistakes will be created as the data set gets updated. We invite all researchers to submit their proposed improvements and thus contribute to the continued maintenance of the first freely available

Electronic copy available at: https://ssrn.com/abstract=4059513

and comprehensive ownership data set for U.S. firms.

One limitation of the part of our analysis measuring "universal" and "common" ownership is that we use only one measure proposed in the literature. Our conclusions do depend on the particular measure we used. Many other measures of common ownership have been proposed and can also be calculated, sometimes more accurately, using the data we construct.

A remaining limitation concerns the scope of our analysis, which is limited to the subset of S&P 500 firms that do not have dual-class shares or controlling owners. This limitation is particularly constraining when calculating within-industry measures of "common ownership", because not all rivals are S&P 500 firms in many cases. Including the firms with dual-class stock structures is likely to further increase the variation in ownership structures. Future research could expand the data set accordingly. We also do not observe all holdings by activists or insiders, but only those that are required to be reported to the SEC. More comprehensive reporting requirements would allow researchers to lift this limitation. If reporting dates were harmonized across institutions and insiders, the accuracy of the data could be further improved.

Another takeaway for policymakers is the necessity of a *comprehensive* set of ownership records to accurately assess the extent of common ownership within industries and of universal ownership more generally in their jurisdictions. Furthermore, understanding the respective drivers is essential for policymakers deliberating on potential measures to limit common ownership without preventing investor diversification. Basing such understanding on institutional investors' filings alone can lead to wrong conclusions. For example, our research based on more accurate institutional ownership data and controlling for insider and blockholder ownership does not support the finding that the sizable holdings of the "Big Three" institutional investors do not contribute to universal and common ownership over and above investor indexing. Investigating question further appears a fruitful avenue for future research. Relatedly, assessing the likely effect of consolidation in the asset management industry on universal and common ownership cannot accurately be performed based on 13-F ownership alone. We thus hope that the provision of our expanded data set not only enables more high-quality research, but also enables competition authorities to measure the level of universal and common ownership more accurately, analyze the likely effect of proposed policies – or the likely effect of not enacting any.

Electronic copy available at: https://ssrn.com/abstract=4059513

References

Antón, Miguel, Florian Ederer, Mireia Giné, and Martin Schmalz, "Common Ownership, Competition, and Top Management Incentives," *SSRN Working Paper 2802332*, 2022.

Azar, José and Xavier Vives, "General Equilibrium Oligopoly and Ownership Structure," *Econometrica*, 2021.

_ , **Martin Schmalz, and Tecu Isabel**, "Anticompetitive Effects of Common Ownership," *Journal of Finance*, 2018, *74* (3).

Backus, Matthew, Christopher Conlon, and Michael Sinkinson, "Common Ownership and Competition in the Ready-To-Eat Cereal Industry," *NBER Working Paper*, 2021.

_ , _ , **and** _ , "Common ownership in America: 1980-2017," *American Economic Journal: Microeconomics*, 2021, *13* (3), 273–308.

Brav, Alon, Wei Jiang, and Hyunseob Kim, "Hedge fund activism: A review," *Foundation and Trends in Finance*, 2010, *4* (3), 1–66.

_ , _ , **Frank Partnoy, and Randall Thomas**, "Hedge fund activism, corporate governance, and firm performance," *Journal of Finance*, 2008, *63* (4), 1729–1775.

Charoenwong, Ben, Zhenghui Ni, and Qiaozhi Ye, "Active Mutual Fund Common Owners' Returns and Proxy Voting Behavior," *SSRN Working Paper*, 08 2022.

Gilje, Erik P., Todd A. Gormley, and Doron Levit, "Who's paying attention? Measuring common ownership and its impact on managerial incentives," *Journal of Financial Economics*, 2020, *137* (1), 152–178.

Gompers, Paul, Joy Ishii, and Andrew Metrick, "Extreme Governance: An Analysis of Dual-Class Firms in the US," *SSRN Electronic Journal*, 2008.

Grossman, Sanford J. and Oliver Hart, "A theory of competitive equilibrium in stock market economies," *Econometrica*, 1979, pp. 293–329.

Hart, Oliver and Luigi Zingales, "Companies Should Maximize Shareholder Welfare Not Market Value," *Journal of Law, Finance, and Accounting*, 2017, *2* (2), 247–274.

Hoberg, Gerard and Gordon Phillips, "Text-based network industries and endogenous product differentiation," *Journal of Political Economy*, 2016, *124* (5), 1423–1465.

Koch, Andrew, Marios Panayides, and Shawn Thomas, "Common ownership and competition in product markets," *Journal of Financial Economics*, 2020.

Lewellen, Katharina and Michelle Lowry, "Does common ownership really increase firm coordination?," *Journal of Financial Economics*, 2020, *forthcoming.*

Rotemberg, Julio, "Financial transaction costs and industrial performance," *MIT Sloan Working Paper*, 1984.

Schmalz, Martin, "Common Ownership Concentration and Corporate Conduct," *Annual Review of Financial Economics*, 2018, *10.*

_ , "Recent studies on common ownership, firm behavior, and market outcomes," *Antitrust Bulletin*, 2021, *66* (1).

Electronic copy available at: https://ssrn.com/abstract=4059513

5 Figures & Tables

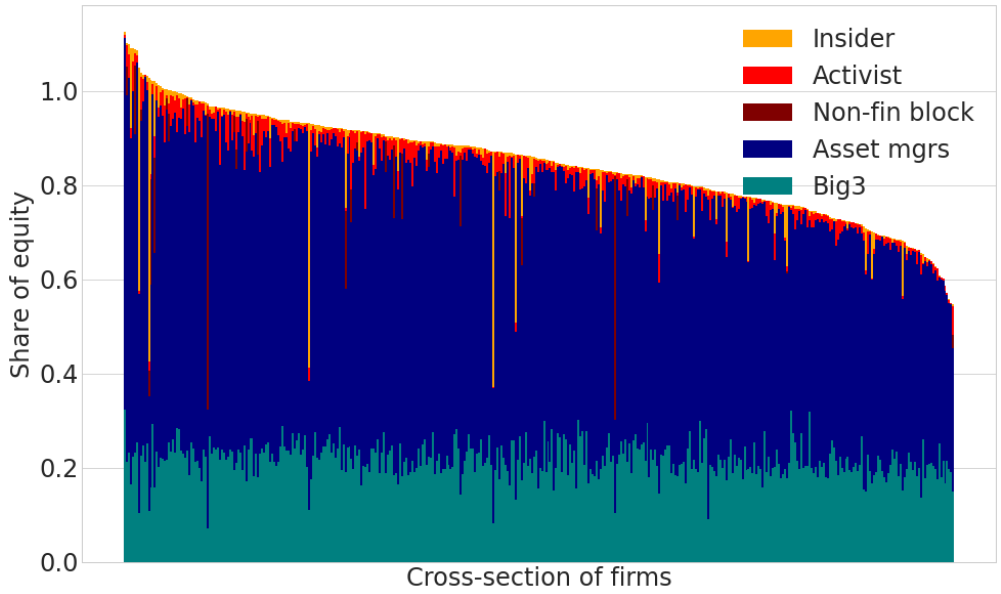


Figure 1. Ownership structures by filer type, 1 January 2020

Note: This figure depicts the captured ownership for each firm at the end of Q4 2019, differentiating ownership by how active the filer likely is. Each bar represents one firm. The bar height measures how much of the firm's stock ownership we capture. The red proportion of each bar represents the aggregate ownership of such investors identified by Brav et al. (2019) as "activists". The yellow proportion is the ownership of corporate insiders (filing Form 3/4/5 reports), who are not classified as "activists" by Brav et al. The navy part measures the share of equity owned by 13-F filings investors, who are neither activists, nor insiders, not BlackRock, Vanguard and State Street, as the "Big Three" ownership stakes in each firm are depicted in turquoise. The maroon part of each bar measures the remaining ownership captured by non-activist, non-insider, non-financial blockholders (13-G filers).

	Indexing	Big 3 holdings	Activist holdings	Non-fin blockh.	Insider holdings	Asset mgr holdings	Retail share
Investor indexing	1.0000	0.5797	-0.2339	-0.3361	-0.4482	-0.3189	0.3185
Big 3 holdings	0.5797	1.0000	0.1612	-0.1049	-0.1346	-0.1555	-0.2604
Activist holdings	-0.2339	0.1612	1.0000	0.0153	-0.0140	0.0390	-0.3412
Non-fin blockh.	-0.3361	-0.1049	0.0153	1.0000	0.0409	-0.0054	-0.2152
Insider holdings	-0.4482	-0.1346	-0.0140	0.0409	1.0000	-0.1470	-0.1359
Asset mgr holdings	-0.3189	-0.1555	0.0390	-0.0054	-0.1470	1.0000	-0.7913
Retail share	0.3185	-0.2604	-0.3412	-0.2152	-0.1359	-0.7913	1.0000

Table 2. Correlations between holdings

Electronic copy available at: https://ssrn.com/abstract=4059513

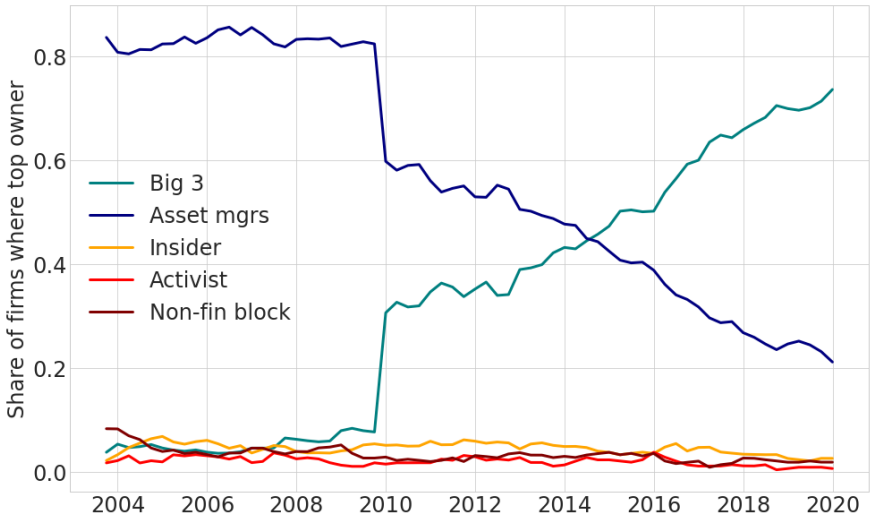
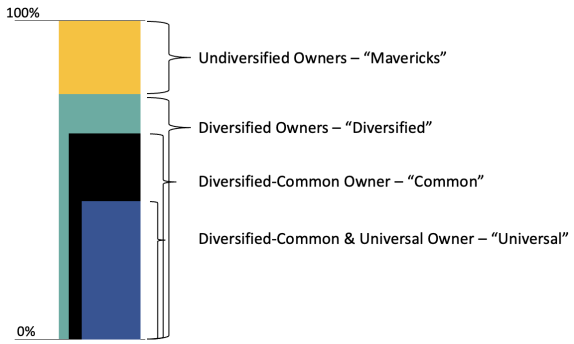
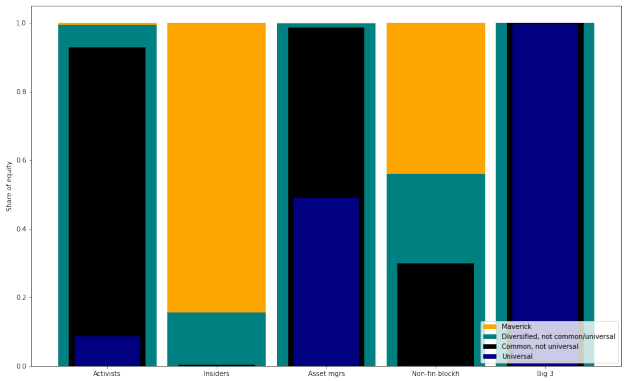


Figure 2. Top owner position by type of filer

Note: This figure depicts the quarterly share of firms in our sample for which an activist (as defined by Brav et al (2008, 2010), an insider (Form 3, 4, 5 filers), a non-financial blockholder (non-activist, non-insider, non-asset manager), an asset manager (13-F filers, excl Big 3) or one of the Big 3 (BlackRock, Vanguard, State Street) reports the largest ownership stake.



(A) Diversification types



(B) Diversification of certain investor groups

Note: This figure depicts which share of activist, insider, asset manager, Big 3 and non-financial blocks' holdings is held by an undiversified maverick owner, by a diversified owner (non-universal and non-common, i.e. only diversified across industries only), a common owner (holds multiple firms in one industry and across industries at the same time) and a universal owner (holds more than 95% of the firms in the sample at the same time, making him both a diversified and a common owner).

Electronic copy available at: https://ssrn.com/abstract=4059513

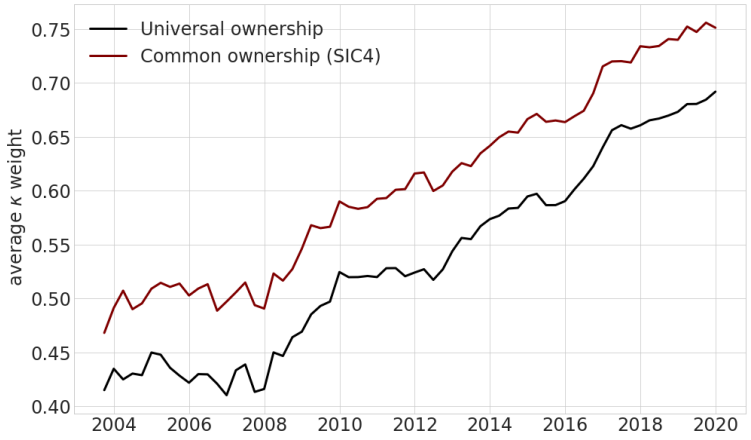


Figure 4. Average universal and common ownership profit weights

Note: This figure depicts the average universal ownership profit weights over time in blue and the average common ownership profit weights in dark red. The average common ownership profit weight is calculated using only profit weights of firm-pairs where both firms are in the same SIC industry. Profit weights are averaged each quarter across firms weighting all firms equally.

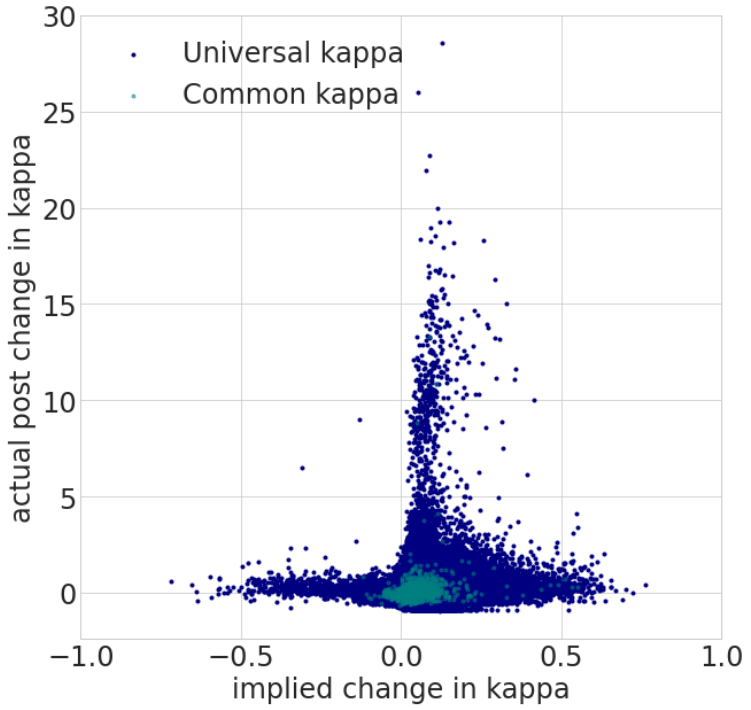


Figure 5. Change profit weights implied and actual post merger

Note: This figure depicts the changes in ownership structure similarity implied by the merger of BlackRock and BGI versus the changes in actual post merger profit weights in percentage point units. The "implied" firm-pair profit weight change is calculated as the difference between counterfactual profit weights with BlackRock & BGI holdings in 2009 Q1 added up vs the correct profit weights in 2009 Q1. Actual profit weight changes are the difference between profit weights in 2010 Q1 vs 2009 Q1.

Electronic copy available at: https://ssrn.com/abstract=4059513

	(1) κ_{univ}	(2) κ_{univ}	(3) κ_{univ}	(3a) κ_{BCS}	(4) κ_{com}	(5) κ_{com}	(6) κ_{com}
Insider holdings	-0.573***	-0.179***	**-0.181*** **		-0.685***	-0.283***	**-0.286*** **
	(0.002)	(0.002)	(0.002)		(0.019)	(0.020)	(0.020)
Activist holdings	-0.009***	0.450***	**0.446*** **		-0.132***	0.333***	**0.329*** **
	(0.002)	(0.002)	(0.002)		(0.022)	(0.023)	(0.023)
Non-fin blockholdings	-0.476***	-0.106***	**-0.110*** **		-0.520***	-0.126***	**-0.130*** **
	(0.002)	(0.002)	(0.002)		(0.021)	(0.022)	(0.022)
Investor Indexing		1.044***	**1.024*** **	1.110***		1.039***	**1.013*** **
		(0.001)	(0.002)	(0.001)		(0.017)	(0.019)
Big 3 holdings	0.954***		**0.095*** **	-0.236***	0.930***		**0.132*** **
	(0.003)		(0.003)	0.002	(0.034)		(0.037)
Retail Share	0.874***	0.698***	0.704***	0.670***	0.904***	0.708***	0.718***
Log(Market Cap)	0.071***	0.065***	0.066***	0.071	0.054***	0.046***	0.047***
R-squared	0.586	0.595	0.595	0.721	0.616	0.623	0.623
R-squared Adj.	0.233	0.262	0.262		0.221	0.245	0.245
p(F-stat)	0.000	0.000	0.000		0.000	0.000	0.000
Ordered Pair FE	✓	✓	✓	✓	✓	✓	✓
Quarter-Year FE	✓	✓	✓	✓	✓	✓	✓
N	10061188	10061188	10061188	13230003	95223	95223	95223

Table 3. Kappa regression on drivers

Note: This table presents results of a regression of profit weights on aggregate holdings of corporate insiders (filing Form 3, 4, 5 reports), of activists (identified in Alon Brav et al (2008, 2010)), of non-financial blockholders (non-insider, non-asset manager, non-activists), of the Big 3 asset managers (BlackRock, Vanguard and State Street), a measure of investor indexing (as suggested by Backus et al 2021), the retail share of ownership (defined as 1 minus captured ownership fraction), log market capitalization, operating margin. The first three columns consider all firm-pair profit weights and thus represent an analysis of our "universal" ownership measure. Column 3a provides results for an analogous regression performed in Backus et al 2021 on 13-F institutional ownership data only and serves to showcase how our improved parsing of Big 3 holdings changes the coefficient sign. Columns 4-6 present regression results when only considering firm-pairs in the same industry, interpreted as a measure of "common" ownership. We residualize and adjust the R-squared for quarter-year and firm-pair fixed effects (because κ_{ij} effect differs from κ_{ji}). One star denotes coefficients are significant at the 10% level, two stars significance at the 5% level, three stars significance at the 1% level.

Electronic copy available at: https://ssrn.com/abstract=4059513

	2009Q4	2010Q1	2011Q1	2012Q1
	$\Delta\kappa_{univ,post}$	(2) $\Delta\kappa_{univ,post}$	(3) $\Delta\kappa_{univ,post}$	(4) $\Delta\kappa_{univ,post}$
$\Delta\kappa_{univ,imp}$	**0.500*** **	**0.370*** **	**0.402*** **	**0.417*** **
	(0.008)	(0.009)	(0.010)	(0.011)
R-squared	0.519	0.642	0.845	0.904
N	147070	147070	147070	147070
	$\Delta\kappa_{com,post}$	(2) $\Delta\kappa_{com,post}$	(3) $\Delta\kappa_{com,post}$	(4) $\Delta\kappa_{com,post}$
$\Delta\kappa_{com,imp}$	**0.673*** **	**0.184***	**0.795*** **	**0.841*** **
	(0.090)	(0.102)	(0.117)	(0.137)
R-squared	0.633	0.759	0.923	0.956
N	1244	1244	1244	1244
Firm FE	✓	✓	✓	✓

Table 4. Regression on post-merger kappa change

Note: This table presents results for a firm-pair level regression of actual changes in profit weights due to the BlackRock-BGI merger on implied changes in profit weights. Actual changes are the difference between profit weights calculated for post-merger periods (2009Q4, 2010Q1, 2011Q1, 2012Q1) and profit weights calculated for the pre-merger period 2009Q1. Implied changes are the difference between counterfactual profit weights calculated based on 2009Q1 data when consolidating the BlackRock and BGI ownership stakes and the profit weights calculated on baseline 2009Q1 data. The upper half of the table considers all firm-pairs in our sample, and therefore refers to "universal" profit weights. The lower half is restricted to firm pairs that are part of the same industry and therefore refers to our measure of "common" ownership profit weights. We control for firm-fixed effects and a constant controls for quarter-specific fixed effects. One star denotes coefficients are significant at the 10% level, two stars significance at the 5% level, three stars significance at the 1% level.

Electronic copy available at: https://ssrn.com/abstract=4059513

6 Appendix

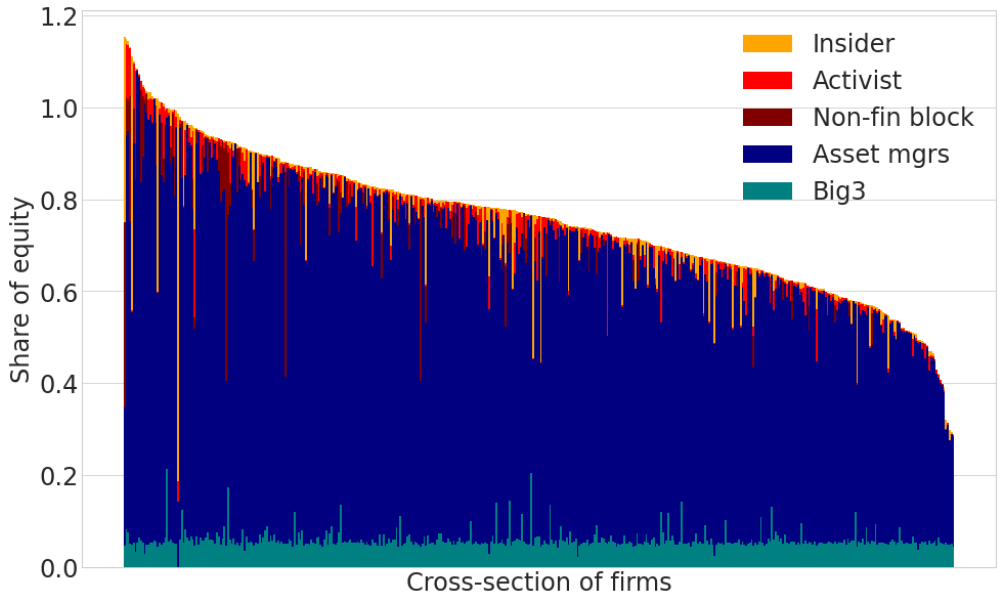


Figure A.1. Ownership structures by activism, 1 January 2005

Note: This figure depicts the captured ownership for each firm at the end of Q4 2004, differentiating ownership by how active the filer likely is. Each bar represents one firm. The bar height measures how much of the firm's stock ownership we capture. The red proportion of each bar represents the aggregate ownership of such investors identified by Brav et al. (2019) as "activists". The yellow proportion is the ownership of corporate insiders (filing Form 3/4/5 reports), who are not classified as "activists" by Brav et al. The navy part measures the share of equity owned by 13-F filings investors, who are neither activists, nor insiders, not BlackRock, Vanguard and State Street, as the "Big Three" ownership stakes in each firm are depicted in turquoise. The maroon part of each bar measures the remaining ownership captured by non-activist, non-insider, non-financial blockholders (13-G filers).

Electronic copy available at: https://ssrn.com/abstract=4059513

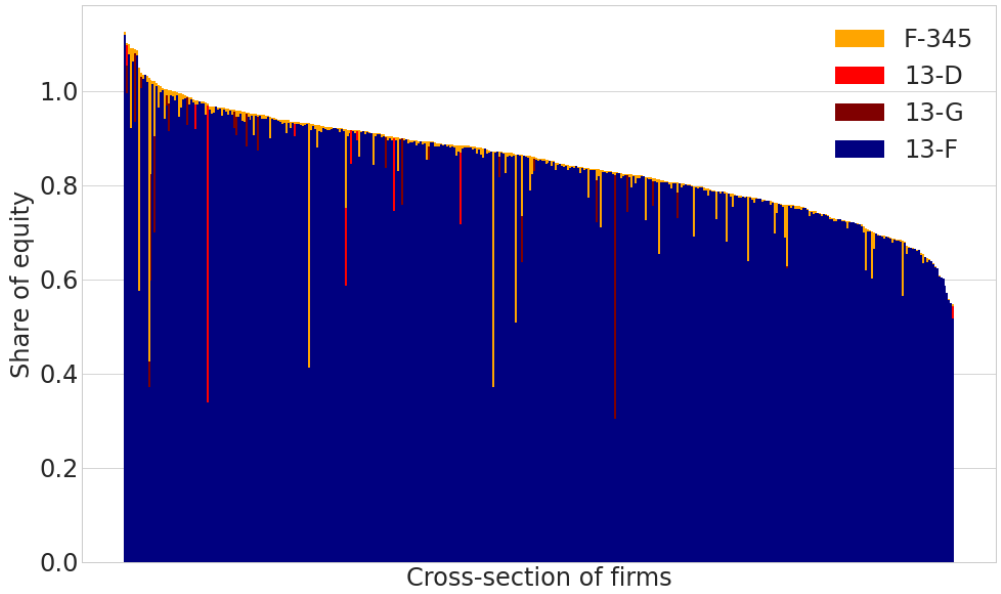


Figure A.2. Ownership structures by filer type, 1 January 2020

Note: This figure depicts the captured ownership for each firm at the end of Q4 2019, differentiating ownership by the filer reporting it. Each bar represents one firm. The bar height measures how much of the firm's stock ownership we capture. The proportion of each bar in blue represents the share of equity ownership identified by parsing 13-F filings. The red and maroon parts of each bar measure the additional ownership captured by parsing 13-D and 13-G filings, the orange part is captured by parsing Form 3, 4 and 5 filings.

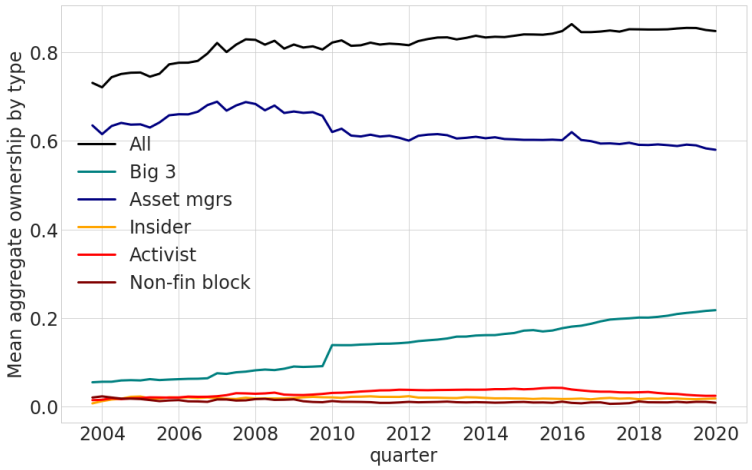


Figure A.3. Average aggregate ownership by investor type

Note: This figure depicts the aggregate share of outstanding equity owned by activists (as defined by Brav et al (2008, 2010), insiders (Form 3, 4, 5 filers), non-financial blockholders (non-activists, non-insiders, non-asset managers), asset managers (13-F filers, excl Big 3), the Big 3 (BlackRock, Vanguard, StateStreet) and across all owners of the average S&P 500 firm for each quarter.

Electronic copy available at: https://ssrn.com/abstract=4059513

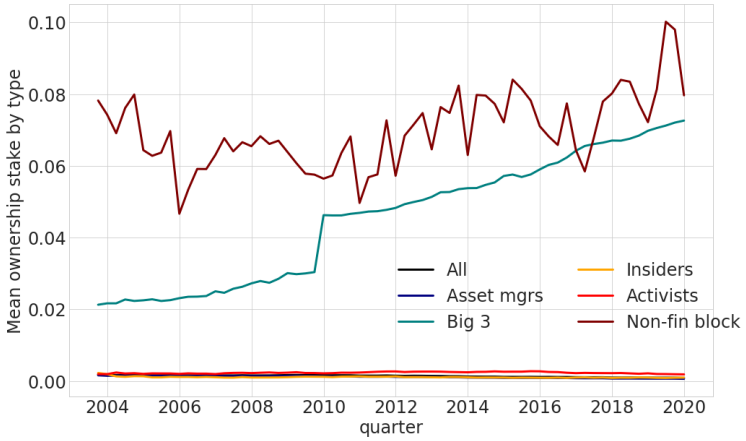


Figure A.4. Average ownership stake by filer type

Note: This figure depicts the average share of outstanding equity owned by an individual activist (as defined by Brav et al (2008, 2010), insider (Form 3, 4, 5 filers), non-financial blockholder (non-activists, non-insiders, non-asset managers), asset manager (13-F filers, excl Big 3), one of the Big 3 (BlackRock, Vanguard, StateStreet) and across all owners of an S&P 500 firm for each quarter.

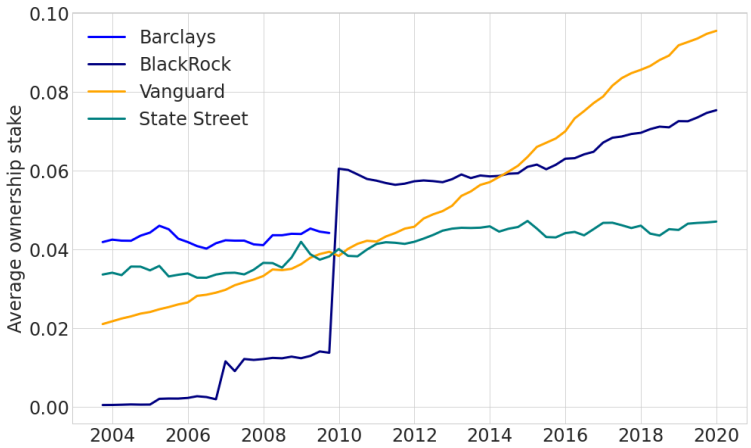


Figure A.5. Average ownership stake of Big 4

Note: This figure depicts the average share of outstanding equity of S&P 500 firms owned by Vanguard, BlackRock, State Street, and Barclays Global Investors (until the merger with BlackRock) each quarter.

Electronic copy available at: https://ssrn.com/abstract=4059513

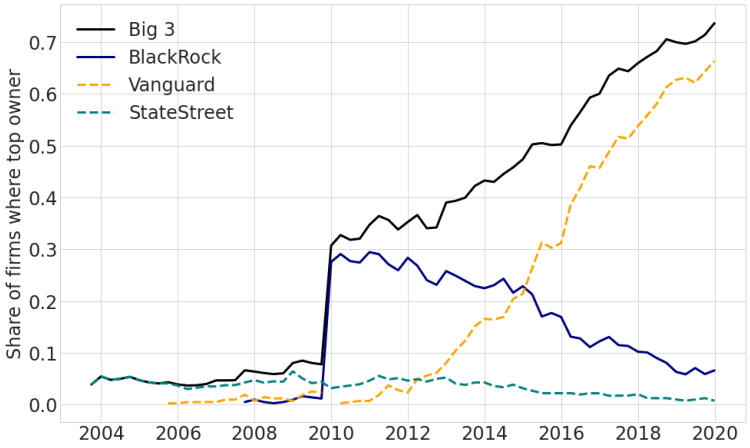


Figure A.6. Top owner position by Big 3

Note: This figure depicts the quarterly share of firms in our sample for which Vanguard, BlackRock, State Street, report the largest ownership stake, or one of the group does.

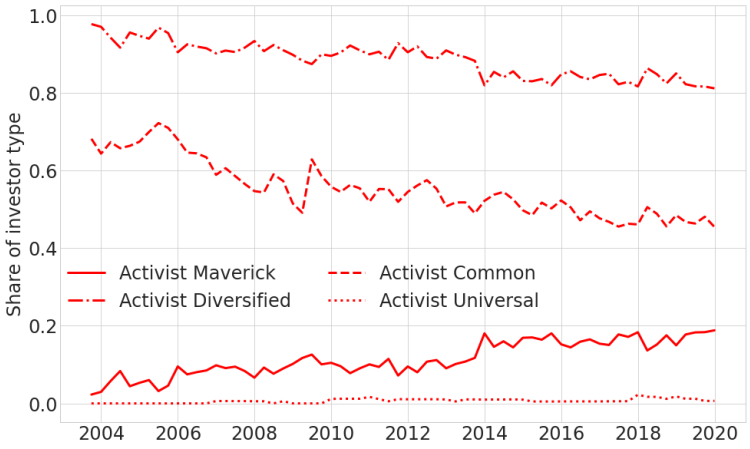


Figure A.7. Diversification by filer type

Note: This figure depicts the quarterly share of activist owners that can be categorized as maverick owners, as diversified owners, as common owners or as universal owners. Maverick owners hold 1 security in a given quarter, diversified owners hold multiple firms per quarter but not within the same industry, common owners hold multiple firms competing in the same industry and across industries but less than 95% of all firms in the sample and universal owners hold more than 95% of all firms in the sample making them both diversified and common owners, too.

Electronic copy available at: https://ssrn.com/abstract=4059513

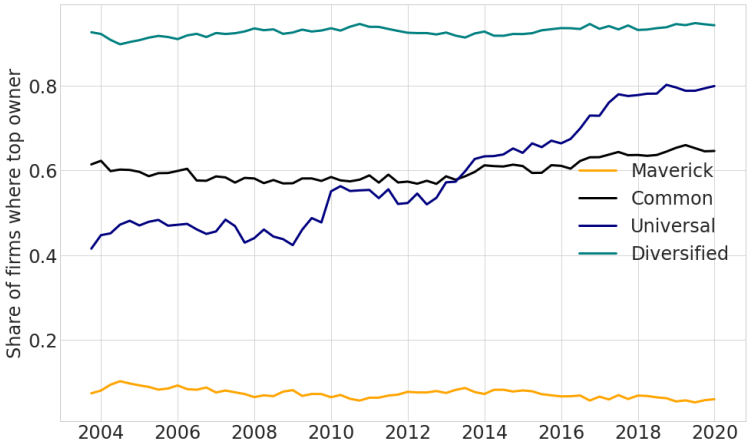
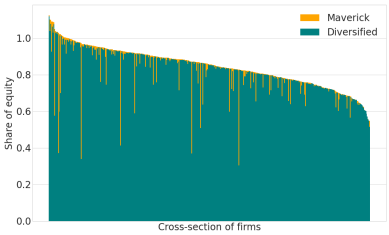
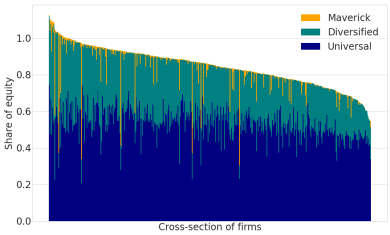


Figure A.8. Top owner position by diversification

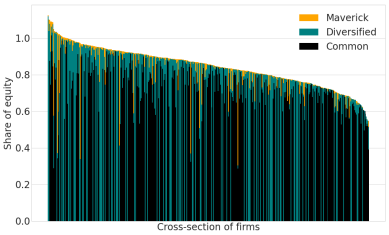
Note: This figure depicts the quarterly share of firms in our sample for which a maverick owner, a diversified owner, a common owner or a universal owner reports the largest ownership stake. Maverick owners hold 1 security in a given quarter, diversified owners hold multiple firms per quarter but not within the same industry, common owners hold multiple firms competing in the same industry and across industries but less than 95% of all firms in the sample and universal owners hold more than 95% of all firms in the sample making them both diversified and common owners, too.



(A) Maverick & diversified ownership



(B) Maverick, diversified & universal ownership

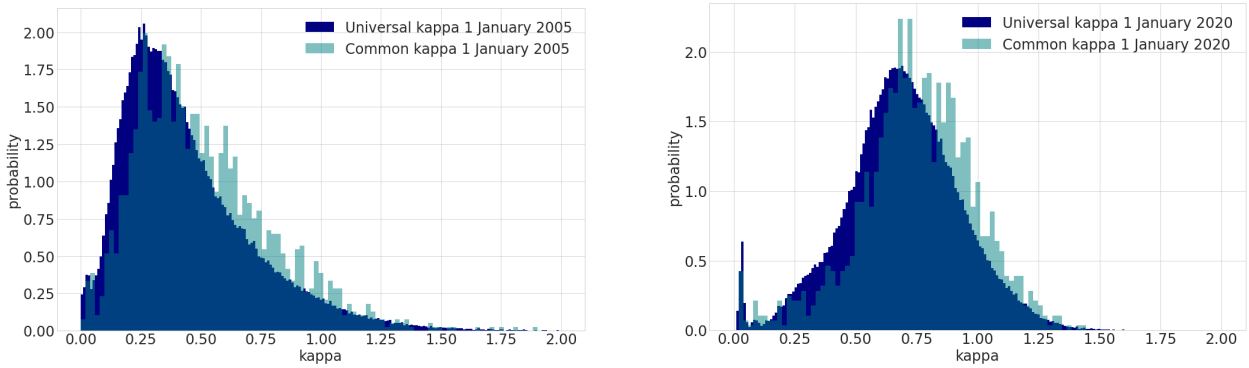


(C) Maverick, diversified & common ownership

Figure A.9. Ownership structures by diversification type, 1 January 2020

Note: These figures depict the captured ownership for each firm on 1 January 2020, differentiating ownership by the diversification of the investor. Each bar represents one firm. The bar height measures how much of the firm's stock ownership we capture. The proportion of each bar in turquoise represents the share of equity owned by diversified investors, who own shares in multiple companies but only across industries and in less than 95% of the sample firms. The blue proportion represents ownership by universal investors, where a universal investor owns more than 95% of the securities in our sample in the given period. The black part of each bar measures the holdings of common, but non-universal investors, who own shares in multiple firms in the same industry and across industries, but less than 95% of the sample. The orange part reports the holdings of undiversified shareholders, who own only shares in this company in the given quarter.

Electronic copy available at: https://ssrn.com/abstract=4059513



(A) 1 January 2005 **(B)** 1 January 2020

Figure A.10. Distribution of universal and common κ over time

Note: This figure depicts the distribution of universal and common ownership profit weights on 1 January 2005 and 1 January 2020. The distribution of universal kappas considers all firm-pairs. The distribution of common kappas considers only kappas for firm-pairs where both firms are in the same SIC 4 digit industry.

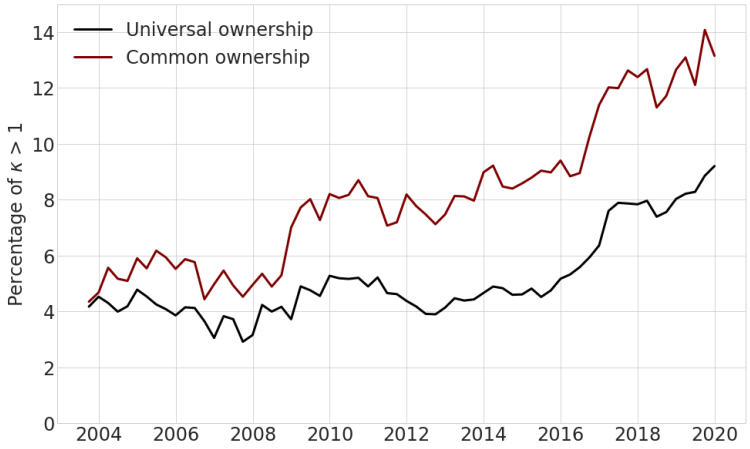
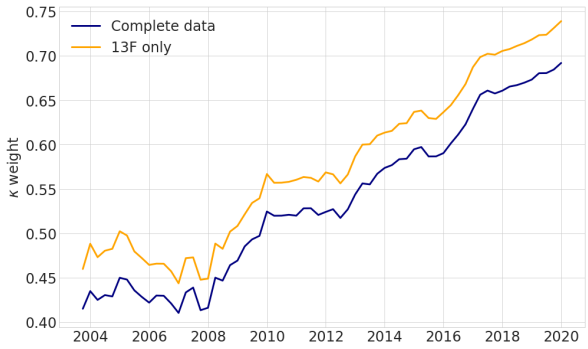


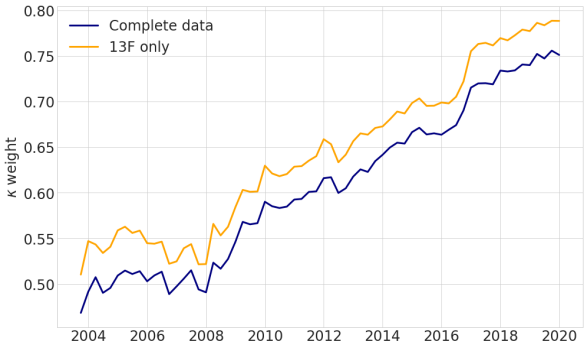
Figure A.11. Prevalence of tunneling incentives of universal and common owners

Note: This figure depicts the fraction of firm-pair profit weights that exceed 1 for universal ownership in black and for common ownership in dark red.

Electronic copy available at: https://ssrn.com/abstract=4059513



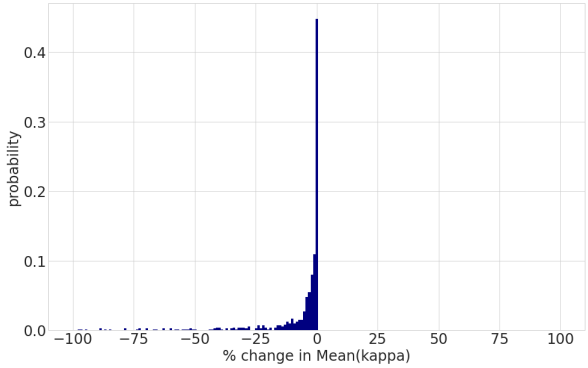
(A) Universal ownership



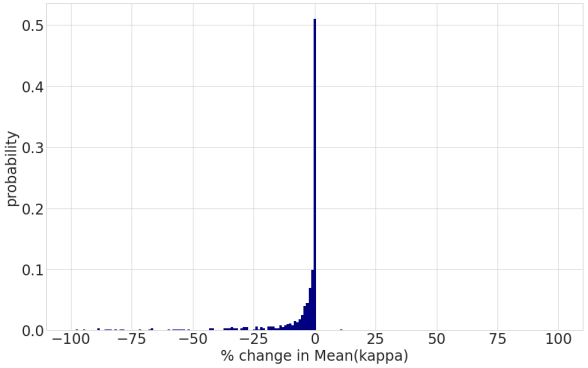
(B) Common ownership

Figure A.12. Distribution of change in κ with complete data

Note: This figure depicts the distribution of percentage changes in universal and common ownership profit weights when adding corporate insiders and blockholders to the 13-F institutional ownership records and recalculating kappas on the complete data set. The distribution of change in universal kappas considers all quarters and all firm-pairs. The distribution of common kappas considers all quarters but only kappas for firm-pairs where both firms are in the same SIC 4 digit industry.



(A) Universal ownership

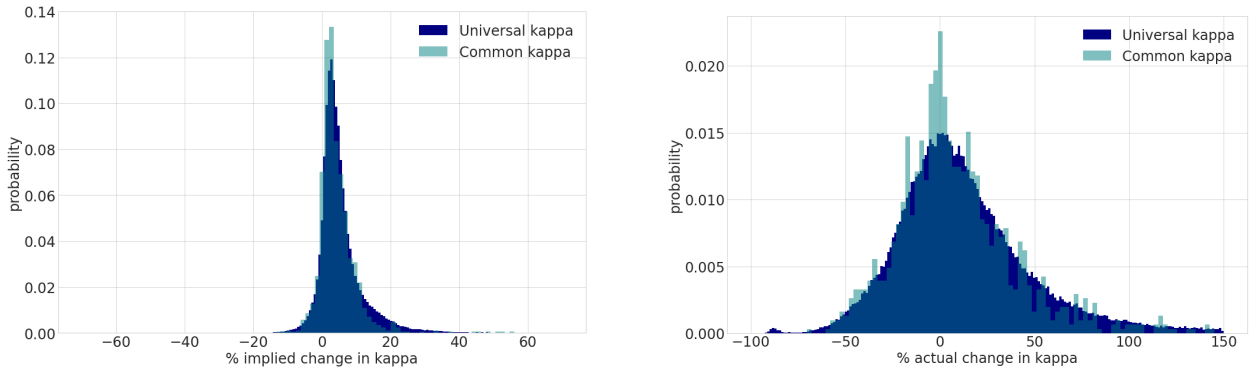


(B) Common ownership

Figure A.13. Distribution of change in κ_{univ} with complete data

Note: This figure depicts the distribution of percentage changes in universal and common ownership profit weights when adding corporate insiders and blockholders to the 13-F institutional ownership records and recalculating kappas on the complete data set. The distribution of change in universal kappas considers all quarters and all firm-pairs. The distribution of common kappas considers all quarters but only kappas for firm-pairs where both firms are in the same SIC 4 digit industry.

Electronic copy available at: https://ssrn.com/abstract=4059513



(A) Change in profit weights implied by merger **(B)** Actual change in profit weights post merger

Figure A.14. Distribution of changes in profit weights due to BLK-BGI merger

Note: This figure depicts the distribution of percentage changes in ownership structure similarity implied by the merger of BlackRock and BGI on the left and the changes in actual post merger profit weights on the right. The "implied" firm-pair profit weight change is calculated as the difference between counterfactual profit weights with BlackRock & BGI holdings in 2009 Q1 added up vs the correct profit weights in 2009 Q1. Actual profit weight changes are the difference between profit weights in 2010 Q1 vs 2009 Q1.

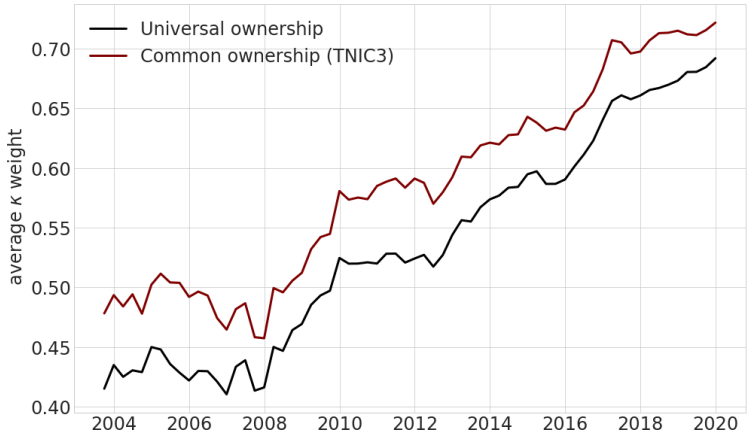


Figure A.15. Average universal and common ownership profit weights

Note: This figure depicts the average universal ownership profit weights over time in black and the average common ownership profit weights in dark red. The average common ownership profit weight is calculated using only profit weights of firm-pairs where both firms are in the same TNIC3 industry, as identified by Hoberg and Phillips (2016). Profit weights are averaged each quarter across firms weighting all firms equally.

Electronic copy available at: https://ssrn.com/abstract=4059513

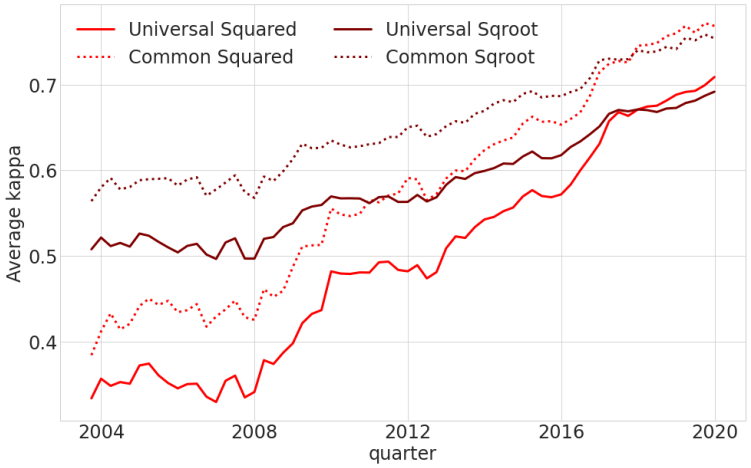


Figure A.16. Average universal and common profit weights varying control assumptions

Note: This figure depicts the profit weights under the assumption $\gamma_{bs} = \beta_{bs}^2$ in dark red and the profit weights under the assumption $\gamma_{bs} = \beta_{bs}^{0.5}$ in light red. Respectively, the unbroken line represents the average universal profit weights, while the dotted line represents average common ownership profit weights. The average common ownership profit weight is calculated using only profit weights of firm-pairs where both firms are in the same SIC-4 digit industry. Profit weights are averaged each quarter across firms weighting all firms equally.

Electronic copy available at: https://ssrn.com/abstract=4059513

	(1) κ_{com}	(2) κ_{com}	(3) κ_{com}
Insider Holdings	-0.756***	-0.291***	-0.295***
	(0.012)	(0.013)	(0.013)
Activist Holding	-0.069***	0.447***	0.442***
	(0.013)	(0.013)	(0.013)
Non fin Blockholdings	-0.709***	-0.260***	-0.265***
	(0.012)	(0.012)	(0.012)
Big 3 Holdings	1.262***		0.121***
	(0.020)		(0.022)
Investor Indexing		1.259***	1.233***
		(0.010)	(0.011)
Retail Share	0.864***	0.632***	0.641***
	(0.004)	(0.004)	(0.005)
Log(Market Cap)	0.077***	0.066***	0.067***
	(0.001)	(0.001)	(0.001)
R-squared	0.574	0.584	0.584
R-squared Adj.	0.230	0.268	0.268
F-stat	10948	13439	11764
p(F-stat)	0.000	0.000	0.000
Ordered Pair FE	✓	✓	✓
Quarter-Year FE	✓	✓	✓
N	256708	256708	256708

Table A.1. Common ownership kappa regression on filer types (TNIC3 industry)

Note: This table presents results of a regression of profit weights for firm pairs identified to be product market competitors by Hoberg & Philipps 2016 on aggregate holdings of corporate insiders (filing Form 3, 4, 5 reports), of activists (identified in Alon Brav et al ((2008, 2010))), of non-financial blockholders (non-insider, non-asset manager, non-activists), of the Big 3 asset managers (BlackRock, Vanguard and State Street), a measure of investor indexing (as suggested by Backus et al 2021), the retail share of ownership (defined as 1 minus captured ownership fraction), log market capitalization, operating margin. . We residualize and adjust the R-squared for quarter-year and firm fixed effects or for quarter-year and ordered firm-pair effects (κ_{ij} effect differs from κ_{ji}). One star denotes coefficients significant at the 10% level, two stars are significance at the 5% level, three stars are significance at the 1% level.

Electronic copy available at: https://ssrn.com/abstract=4059513

	2009Q4 $\Delta\kappa_{com,post}$	2010Q1 (2) $\Delta\kappa_{com,post}$	2011Q1 (3) $\Delta\kappa_{com,post}$	2012Q1 (4) $\Delta\kappa_{com,post}$
$\Delta\kappa_{com,imp}$	0.846***	0.478***	0.672***	0.774***
	(0.049)	(0.056)	(0.061)	(0.069)
R-squared	0.585	0.672	0.880	0.943
R-squared Adj.	0.074	0.019	0.031	0.032
F-stat	296	74	120	125
p(F-stat)	0.000	0.000	0.000	0.000
Firm FE	✓	✓	✓	✓
N	3716	3716	3716	3716

Table A.2. Regression on post-merger common kappa change, TNIC3 industry

Note: This table presents results for a firm-pair level regression of actual changes in common profit weights due to the BlackRock-BGI merger on implied changes of profit weights only consider firm-pairs consisting of product market competitors as identified by Hoberg & Philipps 2016. Actual changes are the difference between profit weights calculated for various post-merger periods (2009Q4, 2010Q1, 2011Q1, 2012Q1) and profit weights calculated for the pre-merger period 2009Q1. Implied changes are the difference between counterfactual profit weights calculated based on 2009Q1 data but when consolidating the BlackRock and BGI ownership stakes and the profit weights calculated on baseline 2009Q1 data. We control for firm and industry-fixed effects and a constant controls for quarter-specific fixed effects. One star denotes coefficients significant at the 10% level, two stars are significance at the 5% level, three stars are significance at the 1% level.

	2009Q4 $\Delta\kappa_{univ,post}$	2009Q4 - 13F (2) $\Delta\kappa_{univ,post}$	2010Q1 (3) $\Delta\kappa_{univ,post}$	2010Q1 - 13F (4) $\Delta\kappa_{univ,post}$
treat	0.046***	0.043***	0.056***	0.070***
	(0.001)	(0.001)	(0.001)	(0.001)
R-squared	0.502	0.637	0.848	0.919
R-squared Adj.	0.031	0.023	0.029	0.035
F-stat	3099	2316	2878	3583
p(F-stat)	0.000	0.000	0.000	0.000
Firm FE	✓	✓	✓	✓
N	98068	98068	98068	98068

Table A.3. Regression on post-merger universal kappa change, discrete

Note: This table presents results for a firm-pair level regression of actual changes in universal profit weights due to the BlackRock-BGI merger on a treatment dummy, that is equal to 1 for the third of firm-pairs where implied profit weights change the most compared to pre-merger profit weights. Actual changes are the difference between profit weights calculated for various post-merger periods (2009Q4, 2010Q1, 2011Q1, 2012Q1) and profit weights calculated for the pre-merger period 2009Q1. Implied changes are the difference between counterfactual profit weights calculated based on 2009Q1 data but when consolidating the BlackRock and BGI ownership stakes and the profit weights calculated on baseline 2009Q1 data. We control for firm and industry-fixed effects and a constant controls for quarter-specific fixed effects. One star denotes coefficients significant at the 10% level, two stars are significance at the 5% level, three stars are significance at the 1% level.

Electronic copy available at: https://ssrn.com/abstract=4059513

	2009Q4 $\Delta\kappa_{com,post}$	2009Q4 - 13F (2) $\Delta\kappa_{com,post}$	2010Q1 (3) $\Delta\kappa_{com,post}$	2010Q1 - 13F (4) $\Delta\kappa_{com,post}$
treat	0.039***	0.034***	0.064***	0.071***
	(0.005)	(0.006)	(0.006)	(0.007)
R-squared	0.568	0.692	0.900	0.954
R-squared Adj.	0.022	0.014	0.042	0.041
F-stat	57	37	108	107
p(F-stat)	0.000	0.000	0.000	0.000
Firm FE	✓	✓	✓	✓
N	2472	2472	2472	2472

Table A.4. Regression on post-merger common kappa change, discrete, TNIC3 industry

Note: This table presents results for a firm-pair level regression of actual changes in common profit weights due to the BlackRock-BGI merger on a treatment dummy, that is equal to 1 for the third of firm-pairs of product market competitors where implied profit weights change the most compared to pre-merger profit weights. Actual changes are the difference between profit weights calculated for various post-merger periods (2009Q4, 2010Q1, 2011Q1, 2012Q1) and profit weights calculated for the pre-merger period 2009Q1. Implied changes are the difference between counterfactual profit weights calculated based on 2009Q1 data but when consolidating the BlackRock and BGI ownership stakes and the profit weights calculated on baseline 2009Q1 data. We control for firm and industry-fixed effects and a constant controls for quarter-specific fixed effects. One star denotes coefficients significant at the 10% level, two stars are significance at the 5% level, three stars are significance at the 1% level.

Electronic copy available at: https://ssrn.com/abstract=4059513

Internet Appendix to *Mavericks, Universal, and Common Owners - The Largest Shareholders of U.S. Public Firms*

(Not for Publication)

Electronic copy available at: https://ssrn.com/abstract=4059513

A Data Appendix

A.1 Overview

This data set is compiled to serve as more complete source of corporate ownership data than what is currently available and used in the academic literature. The data construction exercise is a very tedious procedure which we will describe in detail in the following sections. We utilize all ownership reports required by the Securities Exchange Commission (SEC), comprising institutional investment managers' 13F-HR filings (henceforth *13F*), blockholders' SC 13-D and SC 13-G filings (collectively referred to as *13DG*), and corporate insiders' Form 3, 4 and 5 filings (collectively referred to as *F345*). For details on the respective SEC filing requirements, we refer the reader to the data chapter in the main paper.

Our sample is currently restricted to publicly traded U.S. firms, that are part of the S&P 500 index and have single class stock structures, however, the construction procedure will not change when we expand the sample to encompass more firms.

First we download all 13F, 13DG and F345 filings from EDGAR, the SEC's archive, and *parse* each report. Using file type-specific python code, we automatically extract the identifying information about the date, the owner, the owned security and the number of shares held from every report. Each file type is processed with a separate script, because their format varies significantly and present numerous challenges. Second we *identify* the owned securities using the unique *permno* identifier provided by CRSP[9]. Once uniquely identified, we reduce the data set to the desired sample dimensions, currently keeping only shareholdings of U.S. companies' common stock, if the company is part of the S&P 500 index and has a single class stock structure. In this reduced data set we *aggregate* the shareholdings for each owner-security-quarter combination. As some reports are filed less frequently than once a quarter, and others are revised multiple times within one quarter, we have to adjust the data by interpolating between individual filings or keeping only one holding for each owner-security-quarter.

After parsing, identifying securities and aggregating ownership for each file type separately, we *merge* these three source of ownership data in the last step. Doing so we have to account for duplicated and overlapping shareholding reports, in order to no overstate the holdings of an individual owner and the aggregate stocks owned for a subject company. Despite dropping duplicates we keep for each owner the information, which file types he reported ownership on, so as not to lose information about the intention and nature of each owner.

A.1.1 Variable description

The main variables in our data set are the *permno*, a five digit numeric code that uniquely identifies each security; the *ownercik*, a multi-digit numeric code (central index key assigned by SEC) that identifies each owner; the *rdate*, the end date of the quarter for which we capture the ownership stake; and the number of shares owned for the respective owner-security-quarter combination.

Additional variables occurring in our data construction are the *cik*, central index key of the subject company that is owned, and the historical *cusip*, an eight digit identifying code

[9] More detail of why we use this unique identifier is provided in section A.1.2

Electronic copy available at: https://ssrn.com/abstract=4059513

for each security, which we both do not use as they change over time and do not uniquely identify a security in our data set.

Every report states its type, such as"SC 13" for a 13G report, or "13F-HR" for a 13F, or "3/A" for an amendment to an earlier Form 3 report. We call this *filetype*, and use it to identify whether the owner is an institutional investment manager, a corporate insider, or another type institutional investor.

We also use four different dates when compiling the ownership data, to differentiate between different events and file types. We capture the date when an ownership report is actually filed with the SEC as "filing date", *fdate*, or the date of the transaction that necessitates a 13DG or F345 filing as "event date", *event_date*. Our data set identifies ownership for all owner-security combinations at the end of each quarter, providing the end of quarter "reporting date", *rdate*. For owner-security combinations where we have to interpolate holdings between filings that are multiple quarters apart, we highlight the quarter in which the ownership was originally reported as "filing-reporting date", *filing_rdate*.

A.1.2 Security identification data set

Ownership reports filed with the SEC use either a security's historical *cusip* or a combination of company *cik* and free text description to identify each security. Because *cusip* and *cik* codes for individual securities and firms change over time (e.g. due to security re-issuance and stock splits, or due to company name changes and acquisitions) we have to link each security to a more permanent and yet unique identifying key to use throughout our data set. The best available key, that is also accessible to researchers and link-able with other variables of interest, is a *permno*, the permanent security number assigned to each security by the Wharton Research Data Services platform CRSP.

We compile a data set, referred to as identification data set, that links a firm's *cik*, its securities' *cusip* and the respective security *permno* for each month and can be merged onto our ownership data set to yield the unique *permno* key. This data is a combination of the universe of monthly securities information available on CRSP and on COMPUSTAT, for the time period 1995Q4-2021Q4[10]. In addition, the identification data set contains further useful information for each security available on CRSP or COMPUSTAT, and compiled manually by us. We use information on the security's "share code", share price and number of shares outstanding provided by CRSP to drop non-common stock and non-traded securities. The company's incorporation country and SIC codes let us limit our data set to US firms and in some analysis to certain industries.

In a separate pipeline we compile a data set of dual class companies in the US between 2000-2020. The early years build on work by Gompers et al. (2008), but most of it is manually collected data from quarterly and annual reports and proxy reports filed with the SEC. This file provides a binary indicator for each quarter if a company has a dual class stock structure, and if so captures the number of shares outstanding and number of votes assigned to each class of stock. In combination with the link between security class title and security key (*cusip* and *permno*) parsed from the ownership reports, we can add to each security key-month the binary indicator of dual class stock structures and the number of votes or shares outstanding for each security (if the security is not traded it does not have a

[10]More details on this data combination exercise can be shared when in direct contact with Fiona Kasperk.

Electronic copy available at: https://ssrn.com/abstract=4059513

permno key, so we proceed to invent one by adding a letter, to the traded security's *permno*, hence keeping it unique).

A.2 13F reports

A 13F-HR report is required by institutional investment management firms with more than $100 million assets under management at the end of each quarter. It comprises a file header that is standard to all file types and then lists all securities owned by the institutional investor on the last day of each quarter in a table like fashion, see Figure A.18. These reports are filed in plain text format until 2013, with a variety of layout variation and mis-statements, that necessitate manual checks and a flexible parsing code. From 2013 on, many filings are reported in XML format, which is similar in structure to plain text, but the information contained in the table is linked to XML tags making it easier to identify individual items in the table.

A *cusip* can appear in several rows of the table, reporting chunks of the total shares of the security owned. Also, for the non-XML filings such chunks might be reported in sequential lines, omitting some information on the lines after the first one (see figure).

A.2.1 Parsing

From the filing header, which is mostly standardized across all file types (see Figure A.17) we parse the *ownercik*, *fdate* and *rdate* of each report. This information is common to and therefore duplicated for all shareholdings reported in the following holdings table. For a comprehensive ownership data set, we have to capture the *cusip* key of the respective security owned and the number of shares owned for each row of the holdings table. The total number of shares owned is usually a combination of the shares at "sole", "shared" and "none" free disposal, reported in the final three columns of the holdings table. To ascertain that the number we are parsing as number of shares owned is correct and not the market value of shares or a subset of the security *cusip*, or another erroneous value, our code checks if the number of shares captured is the result of some combination of the sole, shared and none numbers at the end of the row. Earlier 13F reports may also include the price of securities, in that case we also check if the number of shares captured is equal to the result dividing parsed the value by price.

In addition to mis-parsing the number of shares owned, several other parsing difficulties arise because the *ownercik*, filing and reporting date may be missing from the file header. Our code searches for date-like patterns and other mentions of "CIK" or "Central Index Key" in the report to fill these gaps. In the worst case we follow Backus et al. (2021b) and parse the *ownercik* from the file title.

A.2.2 Identification & Aggregation

Once all rows with information about each *cusip* and the respective number of shares owned is parsed and the header information on ownercik, rdate and fdate is added, we merge this data frame with our identification data set to link each *cusip* with the respective PERMNO key.

This step is often complicated because the *cusip* format can vary between 7, 8 or 9 digit codes, where in some cases the final digits are dropped and in other cases leading zeros are

Electronic copy available at: https://ssrn.com/abstract=4059513

dropped. We account for all these possibilities when identifying the securities, but any *cusip* that cannot be linked with a *permno* in the end is dropped from our sample[11].

Since the security holdings table can report ownership in the same security over multiple rows, we aggregate (sum) the number of shares owned by *permno* across multiple rows in one 13F report (filename).

At this point we can have multiple 13F reports for the same *ownercik-rdate* due to amendments. A 13F amendment (13F-HR/A) is filed when the investor restates some of his shareholdings because they were previously misreported. An amendment can, but does not have to include the entire ownership table again. It could also report only those ownership holdings that were previously erroneous. Therefore we cannot simply parse the last 13F filing in each quarter to capture the complete shareholdings of the investor. Instead we parse all 13F reports and amendments filed during the quarter and then keep the last entry for each individual security (*permno*). This procedure is based on the comprehensively validated assumption that if an amendment reports shareholdings in a *cusip*, and this cusip was also reported with a different number of shares on a previous filing in that quarter, then these are not newly added shares owned, but only a restated number of shares owned. Thus we can simply keep the latest reported number of shares owned for every *permno-rdate-ownercik*, sorting on filing date. In quarters where a security is subject to a stock split this procedure is amended, because restatements and backward looking adjustments for stock-splits cannot be differentiated. In these cases we keep the first reported ownership stake and discard the amendments, then manually adjusting the data for stock splits.

With that our 13F ownership data contains unique holdings for each *permno-rdate-ownercik* combination. Using the additional information from the identification data set we now drop ownership in securities that are not common stocks, part of a dual class stock structure, attributed to companies not incorporated in the US or companies in an industry we do not want to consider and companies that are not part of the S&P 500 index.

Finally, we compute the fraction of outstanding shares owned, referred to as *beta* for each *permno-rdate-ownercik* by dividing number of shares owned by the number of shares outstanding reported in CRSP.

A.3 13D & 13G reports

An investor whoe acquires shares in a company and subsequently owns more than 5% of outstanding equity of that company, or who despite selling some shares still owns more than 5% of outstanding equity has to disclose this ownership on a SC 13D or SC 13G report within 60 days of the transaction. 13D and 13G reports are very similar in structure and format. The only difference is that 13D reports contain an additional item where the investor files his intentions leading to the purchase or sale of shares. In the following we describe the general procedure to capture ownership reported in 13D and 13G filings, jointly referred to as 13DG filings.

Each 13DG filing consists of a header, a table that provides information about the owner, the security, the number of shares owned with sole and shared voting and dispositive power and the total number of shares reported ownership in, and finally a text section that repeats

[11]Securities of foreign issuers or such securities delisted at the end of a quarter or traded on very minor exchanges do not have a *permno* key and can therefore be dropped from our sample.

Electronic copy available at: https://ssrn.com/abstract=4059513

the main information summarized in the table with more detailed explanations in written form (see Figure A.20).

One special characteristic of 13DG filings is that they can be filed by groups of investors, where multiple investors share the voting and dispositive power of the transacted shares in some fashion. This happens for about 45% of all 13DG reports we download. In these cases the table reporting ownership is replicated for each co-filing investor, and the number of shares voted and disposed of solely and in shared fashion, as well as the total number of shares reported ownership in varies across the tables.

Two issues prevent an accurate identification of the individual interests of each co-filing investors. First, only the main reporting owner is identified by an ownercik in the file header, while all other investors are included by name only in the top of their respective ownership table. Second, the allocation of voting power among co-filing investors is often highly intransparent[12] and can only be understood by reading the accompanying footnotes and text items. Because the number of total shares owned can vary across reporting owners it is difficult to identify a single ownership stake to allocate to the filing *ownercik-permno-rdate* key.

A.3.1 Parsing

From the filing header we again parse the *ownercik* and *fdate* of the report, in addition to the event_date on which the transaction of shares occurred. We parse the *cusip* and share class from the following section and duplicated these common items for each reporting owner subsequently captured in the table. From the table we identify the reporting ownername, the number of shares voted solely, voted in shared fashion, disposed of solely and in shared fashion, and the total number of shares reported ownership in. Fortunately, in 76% of filings with multiple owners all investors report the same number of total shares owned. In these cases we capture the header ownercik, the cusip and the total number of shares owned.

In all other cases with multiple reporting owners we capture for each ownername the associated sole-voting, shared-voting, sole-dispositive, shared-dispositive and total number of shares reported. To report the most conservative estimate of aggregate ownership of the filing investors we use the single largest number of shares reported to be owned. We do not attempt to combine sole-voted and disposed of shares with some fraction of shared voting ownership, because the results are more error-prone than our current conservative method.

The main errors in parsing arise because of these co-filing investors and because the security-identifying *cusip* key is not reported in a standardized and well-identifiable location. We correct for the first error by manually checking any cases with excessive share ownership despite our conservative approach. To locate the cusip we search the whole filing for multiple custom patterns of places that can report a *cusip* and check that the parsed *cusip* does not violate the format requirements (as deduced from the CUSIP Global Services description).

A.3.2 Identification & Aggregation

For each filing investor we link the *cusip-rdate* pair using the identification data set to a unique *permno* key. This yields a unique *permno-rdate-ownercik* combination that is associated with the conservative number of shares owned estimate.

[12]Example is Walton family Trust and Walton siblings.

Electronic copy available at: https://ssrn.com/abstract=4059513

As there may have been multiple transactions in a given quarter we check for duplicates of the permno-rdate-ownercik combination and keep only the latest based on transaction (event) date. In case there is a stock split in that quarter for the respective security we adjust for it, depending on whether the event date is before or after the split. Lastly, because 13DG filings are mainly available in digitized form on Edgar since Q4 1999, we keep only the latest of all previous filings for each permno-ownercik and date it to the first quarter of 2000, as our starting point.

Using the additional information from the identification data set we now drop ownership in securities that are not common stocks, part of a dual class stock structure, attributed to companies not incorporated in the US or companies in an industry we do not want to consider and companies that are not part of the S&P 500 index.

Finally, we compute the fraction of outstanding shares owned, referred to as *beta* for each *permno-rdate-ownercik* by dividing number of shares owned by the number of shares outstanding reported in CRSP.

A.4 For 3, 4 & 5 reports

Company officers, directors, 10% owners and other insiders are required to file a Form 4 report within 48 hours of purchasing or selling shares in their company. Additionally, Form 3 reports are filed when a person becomes a corporate insider (e.g. due to election to the board) or on the first occasion an insider trades in his comany's stock. If an insider neglected his reporting requirements within the set timeline, he will file a Form 5 report instead of the Form 4 report.

The structure of Form 3, 4 and 5 filings is the same: each has a file header with information about the filing ownercik, the subject company cik and the filing date, followed by two tables. The first table lists all transactions in non-derivative securities of the company, where each transaction has is reported in a new row. The second table lists all transactions in derivative securities. The rows are in chronological order with the first transaction at the top and the last one at the bottom. For each row we are interest in the free text description of the security being transacted the number of securities owned post transaction and the nature of ownership (directly or indirectly, e.g. by trust or family) for each transaction (see Figure A.19.

A.4.1 Parsing

Parsing the downloaded Forms 3, 4 and 5 presents more challenges than the other filing types, because the security owned is not specified exactly by a *cusip*, but by a broader free text description and the company *cik*. Another challenge is that some insider reports are filed by several owners jointly again (as for 13DG filings). Unlike the 13DG multi-owner reports, however, in F345 filings each reporting owner is identified by name and *ownercik* and they all jointly own the reported shares and do not differentiate into sole or shared voting powers.

We parse the filing date and all *ownercik-ownername* combinations from the document header. From the non-derivative and derivative transactions table we parse for each row the free-text description of the security, the number of shares owned after the transaction and the nature of ownership.

Electronic copy available at: https://ssrn.com/abstract=4059513

Because the number of transactions across all these insider reports is so extensive, we reduce our data set before moving to the identification and aggregation step. As we are only interested in the number of shares owned in a certain fashion after the last transaction, we drop duplicates by free-text description-nature of ownership and keep only the last row. We have to differentiate by ownership type, because these will ultimately need to be summed together.

A.4.2 Identify & Aggregating

To identify the securities owned by corporate insiders we have to merge the combination of *cik* and free text description reported in F345 reports to our identification data set. The difficulty lies in the nature of the free-text description, which varies in accuracy, spelling and informativeness. If the security is described as "Common stock" or even more accurately as "Common stock Class A" it is easy to link the *cik*-free text to a *permno* in our identification data (especially since we are currently only consider single class companies). However, where the description is only a company name or general terms such as "equity" or a combination of these, it is much harder to identify the type and class of security being traded)it could be debt, options, preferred stock or any other security).

To filter through more than 40,000 different descriptions, we use spelling correction packages and manual rules. In a first step we label securities based on the most informative keywords found in the description. For example if the world "common" appears, the indicator for common stock is flagged, while when the word "option" appears, the option indicator is flagged. The Python library *difflib* allows to account for spelling mistakes by searching for words that are very similar to our pre-specified keywords. Similarity is measured by distance between two words and the acceptable distance can be varied manually. With this methodology we assign more than 30 binary labels for security types, such as common stock, preferred stock, restricted stock units, trusts, options, debt, depositary receipts, deferred compensation and many more (full list of most frequent labels and their categorization as either common stock or not in Table A.5). In addition, whenever we identify the term "class" or "series" in the free text field, combined with a single letter, we take note of this in another column, to use in future work where firms with dual class stocks structures are incorporated in the sample. The full list of labels and labeling code can be provided upon request. If the free text description offers too little information to infer the security type, for example when just the company name is reported or an unknown abbreviation is used, we make assumptions about the security type. One such assumption is that descriptions that only state the company name are likely to refer to the main common stock of the company. Another is that if the description is too ambiguous or incomprehensible, it is likely not common stock and we classify it as such and drop it consequently. This is a conservative assumption in order to avoid including false entries in our data set. All of our assumptions and rules are verified by manually accessing and reading the actual filing and comparing it to holding information accessible for example via Capital IQ.

Based on the resulting labels we determine which securities confer voting and cash flow rights and should be included in our ownership database. Obvious common stock or common equity entries are kept as such. Most labels such as debt, swaps, options, preferred equity, phantom stock and warrants, however, are dropped from our sample. Using only the remaining ownership data and security titles we merge in our identification data set (on cik-rdate, as unique securities) and thereby gain the unique PERMNO key.

Electronic copy available at: https://ssrn.com/abstract=4059513

Table A.5. Labels of Form 4-filed security types

label	classification
common stock	common stock like
class "." stock	common stock like
trusts	common stock like
preferred stock	not common stock
(restricted) stock unit	not common stock
stock options	not common stock
debt	not common stock
depositary receipts (ADR)	not common stock
depositary shares (ADS)	not common stock
deferred compensation	not common stock
non-qualified & qualified (...)	not common stock
"company name" trust	not common stock
closed & open-end fund	not common stock
restricted performance (derivative)	not common stock

Next we identify the main owner in a filing with several reporting owners (main_owner). Based on all pairs of co-filing owners and the number of co-filed reports we investigate graphs of the centrality of each co-filing owner in a network of co-filers (using *networkx* Python library, where graph has owners as vertices, and they are connected if they appear on filings together with the number of filings as edge weights).

For each *cik-rdate-ownercik-filetype* we keep the latest filing per quarter (by fdate and file number), and then keep only the ver latest insider report, preferring Form 5 to Form 4 to Form 3 entries. This yields a unique ownership stake for every *cik-rdate-ownercik*.

A.5 Merging

The final step of merging the three ownership data sets from 13F, 13DG and Form 345 filings together presents many more challenges due to duplicated ownership reports. There can be duplicated reports for the same owner in different filing types (e.g. BlackRock has to file a 13F report for its ownership stakes every quarter, but when it changes its position in a company where it owns more than 5% of outstanding equity it also has to file a 13G report). At the same time there can be multiple owners reporting the same shareholdings on the same file type (e.g. when both Bill and Melinda Gates report on a Form 4 their respective direct holdings, but also their indirect holdings in a company, because they are the joint owners of the Gates Foundation, which holds a stake in this company). Lastly there can also be multiple reports for the same shares by different owners and different file types (e.g. this is the case for Warren Buffett who files a Form 4 ownership report for firms where he is on the board and owns shares directly and indirectly through Berkshire Hathaway, and at the same time Berkshire Hathaway files its ownership on a 13F filing).

To remove such duplicates we first stack the three ownership data set constructed in the previous steps, and make sure the source file type for each row reporting ownercik-permno-rdate-shares is known. At this point the 13F data provides ownership information for each

Electronic copy available at: https://ssrn.com/abstract=4059513

quarter, while we have 13DG and F345 ownership reports only in quarters where a relevant transaction occurred. In addition, we still have multiple row for each F345 that is co-filed by multiple insiders and multiple rows for each individual insider if he reports ownership by different nature (e.g. direct and indirect via trust).

The task now is to remove any evident duplicates within and across file types, where the number of shares reported for a given *permno-rdate* is exactly (or almost exactly at 1% margin) the same for one *ownercik* across *file_types*, or for different *ownerciks* in the same *file_type* or for different *ownerciks* across *file_types*. Given duplicates across F345 and 13DG filings, we prefer F345 filings, which provide more accurate information whilst we prefer 13F reported ownership to any 13DG or F345 ownership, because it is easier to parse and reported more regularly.

The removal of duplicates is done first on the completely disaggregated data (capturing e.g. cases where indirect F345 ownership is the same as a 13DG holding of he related trust). Then we interpolate between 13DG and F345 filings and extrapolate both types of filings by four quarters (when the last filing is non-zero) and repeat the duplication check (capturing e.g. cases where a 13DG is filed one quarter before an associated F345 is filed). Lastly we aggregate multiple natures of ownership reported on one F345 report (e.g. add together direct and indirect ownership) to repeat the check for duplicates at the more aggregated level.

In a final step we keep only the main filing insider, and ascertain that extrapolated 13DG and F345 data does not create ownership reports in quarters where a company is dual class, not in the S&P 500 or not owned by a single 13F filer and therefore likely delisted by the end of the quarter.

Before we have an accurate and comprehensive data set of corporate ownership and "control" in U.S. public firms, we have to check for families of funds. For large investment management firms such as BlackRock or StateStreet we have multiple 13F filings by funds or regional offices that be belong to one family. However, all the shares owned by individual funds are ultimately voted and controlled by the family. Hence, we consolidate such separate 13-F filings into one *ownercik*, assuming that they do not double-file their ownership[13].

A.5.1 Spot checks

We spot check the captured "beta" (the fraction of outstanding equity held by an *ownercik* for a given *permno-rdate*) for obviously wrong, mis-parsed values (larger than 1) and for unlikely values (few institutional investment managers own more than 50% of a companies' equity). We identify several 13F filers (institutional investors) which correctly file ownership in excess of 50% of equity, but for most cases this constitutes a parsing error on our side or a reporting error on the investors' or CRSP's side (reported number of shares outstanding found to be wrong in several cases). There are 562 instances where we parse too high ownership stakes from 13F filings, which constitutes only 0.002% of all the 13F primary keys. We manually correct these instance in two ways. Where we capture a beta greater than 0.54 (empirically established threshold), we test whether dividing the number of shares owned by the price of securities at the quarter end (data from CRSP) yields a reasonable beta. This works in a number of cases and suggests our parsing code still captures market value of shares owned in some instances. Where the resulting "fixed beta" is still in excess

[13]This assumption was validated investigating the 13F filings in depth

Electronic copy available at: https://ssrn.com/abstract=4059513

of 0.5, we manually check and adjust the number of shares owned or number of shares outstanding.

Despite our extensive procedure for dropping duplicated entries within and across ownerciks and file types, some duplicates remain, which we check manually by reading the underlying 13DG and insider reports. We identify entities, such as Edward Lampert, ESL Investment, and RBC Investments, where Lampert is majority shareholder of ESL, which is majority shareholder of RBC Investments, and all three file ownership reports separately (but the number of shares reported varies). We implement manual fixes for such recurring joint ownership reports.

A.6 Data comparison

We have compared our S&P 500 ownership data set extensively to various other data sets, based on coverage (missing owners or firms), existence of clearly false stakes (in excess of 1 or 0.5 for institutions) and random examples of individual firm's ownership structures. Overall we find that our data is more comprehensive and accurate than Thomson Reuters' S34 data on institutional ownership, which captures on 13F ownership and has multiple parsing issues.

Our data is similarly comprehensive as Factset and Capital IQ. Both struggle with the duplicated ownership reports across owners and filetypes, as we do, implying that in each data set you will find that the aggregate stock holdings in a company can exceed the number of shares outstanding fr this company in the given quarter.

Most closely related in the approach is the data set by Backus et al. (2021b). Therefore we compare our results in more detail to theirs in the following.

A.6.1 BCS data

BCS parse the CIK of the filing entities from the filename under which a file is saved on the EDGAR database. In about 55% of the filings, however, the number combination in the filename is not the correct CIK of the filing entity, but instead a CIK of another entity assigned by SEC. Manually checking some of these CIKs reveals that they are mostly associated with entities that are consolidated by the SEC, hence not causing significant errors in the ownership measurement. In some cases, however, we find such CIKs are associated with companies unrelated to the filers. Hence, BCS attribute about 55% of the filings to the wrong institutional investor leading to significant errors. We fix this issue by parsing the CIK of the filing entity directly from the header text of the 13-F report.

Multiple other parsing issues of the BCS code arise because the format and structure of 13-F filings changes over time. Before the introduction of XML filing formats in 2013, the reports were produced and stored in plain text. Reading these reports with an automated approach is very difficult as each document format and structure may differ slightly from the other.

First, some filings have the order of the "value" and "shares" columns reversed. BCS manually compile a list of filings where they found this to be the case, swap the columns for such filings, and implement additional checks to see if the parsed number of shares is sensible. To avoid relying on such a manual approach and to account for other possible variations in the columns of ownership reports, we utilize the fact that the total shares reported in an entry of a 13-F filing is equal to the sum of three numbers placed at the end of each row: the

Electronic copy available at: https://ssrn.com/abstract=4059513

"sole", "shared" and "none" columns provide the number of shares for which voting power is sole, shared or not owned. Our approach is to find the number in each row that equals the sum of these three. If this approach fails, we use the BCS method.

Second, earlier 13-F filings report some security holdings over multiple rows. BCS only parse the first row of such entries and skip the subsequent rows attributed to the same security. The neglected rows often contain very large ownership stakes, leading to significant errors in the data. We make sure that our code reads every line of the 13-F reported holdings and adds up those reporting on the same security.

Third, we improve on the consolidation of institutional ownership reports filed by previously separate asset managers that merged at some point. BCS also consolidate such merging entities, however they already consolidate holdings before the merger actually happened. Thereby corporate ownership by BlackRock is massively overstated in the quarters where it had not yet acquired Barclays Global Investors. We correct for this by consolidating holdings only after two institutional investors actually merged.

Fourth, we perform a more comprehensive identification of dual class firms to construct a data set containing only single class companies. We identify the years and quarters in which a given corporation had a dual class stock structure and only drop it for these quarters. We also extend the list of dual class companies to account for younger and very short-lived dual class firms. BCS instead drop a firm from the sample entirely, if it has ever had a dual class stock structure.

Another small amendment of our code makes sure the reporting date of the filing is parsed in cases when it is missing from the header. While BCS assign the filing date to the reporting date in such cases, we check other parts in the filing text from which one can parse the reporting date.

Lastly, we amended the BCS code in order to create a complete database of corporate ownership for public US corporations. The original parsing code does not download and parse the 13-F filings completely, but instead parses only holdings for a pre-specified list of about 5000 securities. The list contains the unique CUSIP identifiers of the securities of S&P 500 firms. Securities of non-S&P 500 firms and even securities of S&P 500 firms that are not correctly listed by the CUSIP are therefore missed. In order to create the exhaustive database we aim for, our code downloads all 13-F filings and parses all security holding contained in the reports.

Electronic copy available at: https://ssrn.com/abstract=4059513

```
FILED AS OF DATE:               20210210
DATE AS OF CHANGE:              20210210

SUBJECT COMPANY:

        COMPANY DATA:
                COMPANY CONFORMED NAME:                 Tesla, Inc.
                CENTRAL INDEX KEY:                      0001318605
                STANDARD INDUSTRIAL CLASSIFICATION:     MOTOR VEHICLES & PASSENGER CAR BODIES [3711]
                IRS NUMBER:                             912197729
                STATE OF INCORPORATION:                 DE
                FISCAL YEAR END:                        1231

        FILING VALUES:
                FORM TYPE:              SC 13G
                SEC ACT:                1934 Act
                SEC FILE NUMBER:        005-85943
                FILM NUMBER:            21612104

        BUSINESS ADDRESS:
                STREET 1:               3500 DEER CREEK RD
                CITY:                   PALO ALTO
                STATE:                  CA
                ZIP:                    94304
                BUSINESS PHONE:         650-681-5000

        MAIL ADDRESS:
                STREET 1:               3500 DEER CREEK RD
                CITY:                   PALO ALTO
                STATE:                  CA
                ZIP:                    94304

        FORMER COMPANY:
                FORMER CONFORMED NAME:  TESLA MOTORS INC
                DATE OF NAME CHANGE:    20050222

FILED BY:

        COMPANY DATA:
                COMPANY CONFORMED NAME:                 VANGUARD GROUP INC
                CENTRAL INDEX KEY:                      0000102909
                IRS NUMBER:                             231945930
                STATE OF INCORPORATION:                 PA
                FISCAL YEAR END:                        1231
```

Figure A.17. Filing header example

Electronic copy available at: https://ssrn.com/abstract=4059513

UNITED STATES SECURITIES AND EXCHANGE COMMISSION
Washington, D.C. 20549
FORM 13F

FORM 13F INFORMATION TABLE

OMB APPROVAL	
OMB Number:	3235-0006
Expires:	Oct 31, 2018
Estimated average burden hours per response:	23.8

COLUMN 1	COLUMN 2	COLUMN 3	COLUMN 4	COLUMN 5			COLUMN 6	COLUMN 7	COLUMN 8		
			VALUE	SHRS OR	SH/	PUT/	INVESTMENT	OTHER	VOTING AUTHORITY		
NAME OF ISSUER	TITLE OF CLASS	CUSIP	(x$1000)	PRN AMT	PRN	CALL	DISCRETION	MANAGER	SOLE	SHARED	NONE
1 800 FLOWERS COM INC	CL A	68243Q106	2,756	86,485	SH		SOLE	2	34,255	0	52,230
1 800 FLOWERS COM INC	CL A	68243Q106	3	96	SH		SOLE	4	96	0	0
1 800 FLOWERS COM INC	CL A	68243Q106	15	469	SH		SOLE	7	469	0	0
1 800 FLOWERS COM INC	CL A	68243Q106	3,250	101,975	SH		SOLE	18	83,710	0	18,265
1 800 FLOWERS COM INC	CL A	68243Q106	16,417	515,114	SH		SOLE	19	515,114	0	0
1 800 FLOWERS COM INC	CL A	68243Q106	29,507	925,870	SH		SOLE	21	925,870	0	0
1 800 FLOWERS COM INC	CL A	68243Q106	34,790	1,091,611	SH		SOLE	22	1,076,379	0	15,232
1 800 FLOWERS COM INC	CL A	68243Q106	643	20,191	SH		SOLE	23	20,191	0	0
1 800 FLOWERS COM INC	CL A	68243Q106	54	1,681	SH		SOLE	25	1,681	0	0
1 800 FLOWERS COM INC	CL A	68243Q106	2,815	88,324	SH		SOLE	26	1,801	0	86,523
1 800 FLOWERS COM INC	CL A	68243Q106	113	3,548	SH		SOLE	28	3,548	0	0
1 800 FLOWERS COM INC	CL A	68243Q106	1,221	38,325	SH		SOLE	30	38,325	0	0

Figure A.18. 13F filing example

FORM 4

UNITED STATES SECURITIES AND EXCHANGE COMMISSION
Washington, D.C. 20549

OMB APPROVAL	
OMB Number:	3235-0287
Estimated average burden hours per response:	0.5

☐ Check this box if no longer subject to Section 16. Form 4 or Form 5 obligations may continue. *See* Instruction 1(b).

STATEMENT OF CHANGES IN BENEFICIAL OWNERSHIP

Filed pursuant to Section 16(a) of the Securities Exchange Act of 1934 or Section 30(h) of the Investment Company Act of 1940

1. Name and Address of Reporting Person*
WALTON ALICE L

(Last) (First) (Middle)
P.O. BOX 1860

(Street)
BENTONVILLE AR 72712

(City) (State) (Zip)

2. Issuer Name **and** Ticker or Trading Symbol
Walmart Inc. [WMT]

3. Date of Earliest Transaction (Month/Day/Year)
09/28/2020

4. If Amendment, Date of Original Filed (Month/Day/Year)

5. Relationship of Reporting Person(s) to Issuer (Check all applicable)
Director X 10% Owner
Officer (give title below) Other (specify below)

6. Individual or Joint/Group Filing (Check Applicable Line)
X Form filed by One Reporting Person
Form filed by More than One Reporting Person

Table I - Non-Derivative Securities Acquired, Disposed of, or Beneficially Owned

1. Title of Security (Instr. 3)	2. Transaction Date (Month/Day/Year)	2A. Deemed Execution Date, if any (Month/Day/Year)	3. Transaction Code (Instr. 8)		4. Securities Acquired (A) or Disposed Of (D) (Instr. 3, 4 and 5)			5. Amount of Securities Beneficially Owned Following Reported Transaction(s) (Instr. 3 and 4)	6. Ownership Form: Direct (D) or Indirect (I) (Instr. 4)	7. Nature of Indirect Beneficial Ownership (Instr. 4)
			Code	V	Amount	(A) or (D)	Price			
Common Stock	09/28/2020		G	V	1,525,000(1)	D	$0	11,323,580	D	
Common Stock	09/29/2020		S		606,875	D	$137.0491(2)	388,797,678	I	By Trust
Common Stock	09/29/2020		S		145,210	D	$137.8084(3)	388,652,468	I	By Trust(4)
Common Stock								1,000,891,131	I	By Limited Liability Company(5)

Figure A.19. F345 filing example

Electronic copy available at: https://ssrn.com/abstract=4059513

CUSIP No.: 88160R101

1. NAME OF REPORTING PERSON
I.R.S. IDENTIFICATION NO. OF ABOVE PERSON

The Vanguard Group - 23-1945930

2. CHECK THE APPROPRIATE [LINE] IF A MEMBER OF A GROUP

A.

3. SEC USE ONLY

4. CITIZENSHIP OF PLACE OF ORGANIZATION

Pennsylvania

(For questions 5-8, report the number of shares beneficially owned by each reporting person with:)

5. SOLE VOTING POWER

0

6. SHARED VOTING POWER

1,305,269

7. SOLE DISPOSITIVE POWER

54,445,577

8. SHARED DISPOSITIVE POWER

3,368,733

9. AGGREGATE AMOUNT BENEFICIALLY OWNED BY EACH REPORTING PERSON

57,814,310

10. CHECK BOX IF THE AGGREGATE AMOUNT IN ROW (9) EXCLUDES CERTAIN SHARES

N/A

11. PERCENT OF CLASS REPRESENTED BY AMOUNT IN ROW 9

6.10%

Figure A.20. 13DG filing example

Electronic copy available at: https://ssrn.com/abstract=4059513

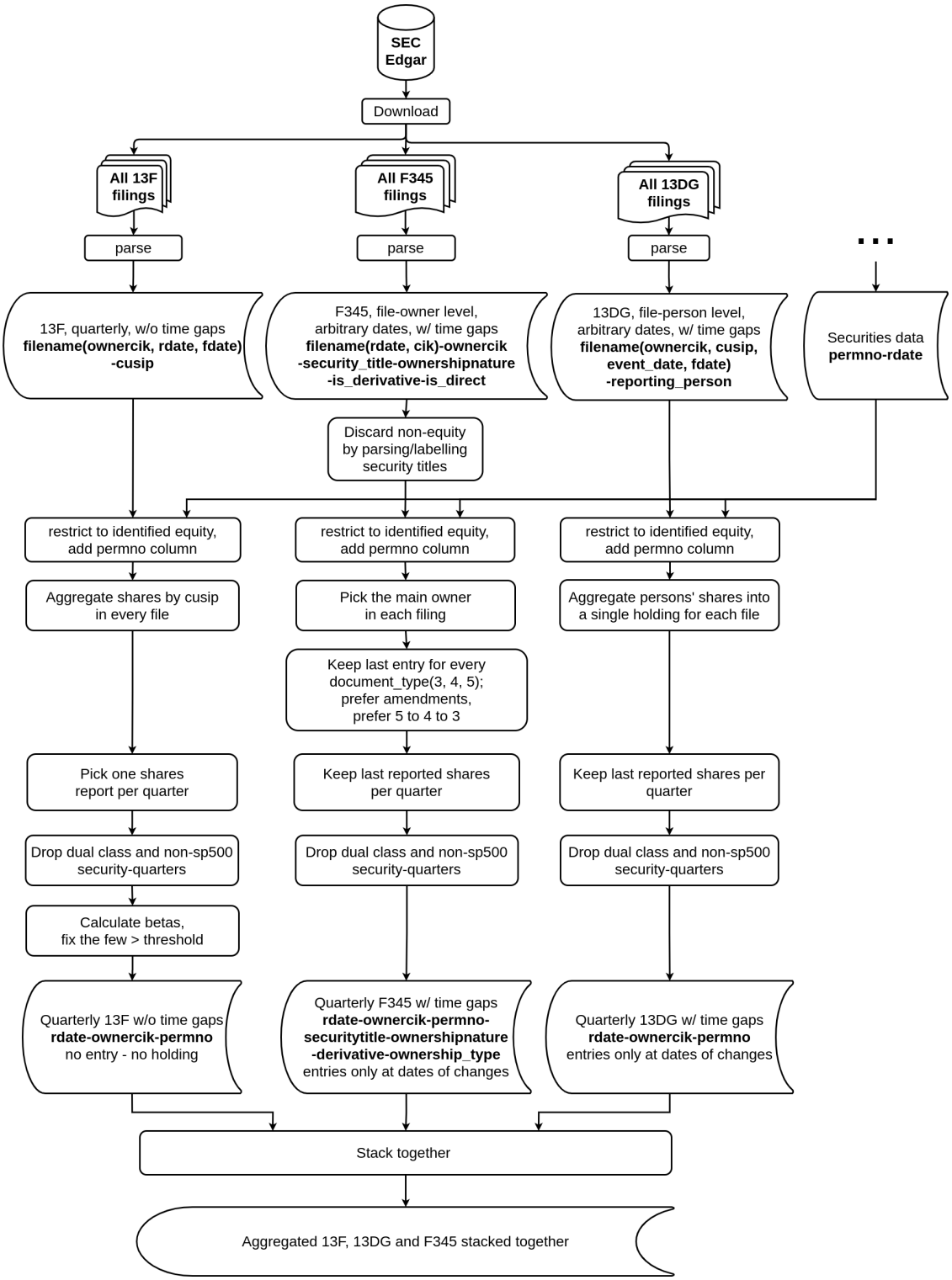


Figure A.21. Aggregation process

Electronic copy available at: https://ssrn.com/abstract=4059513

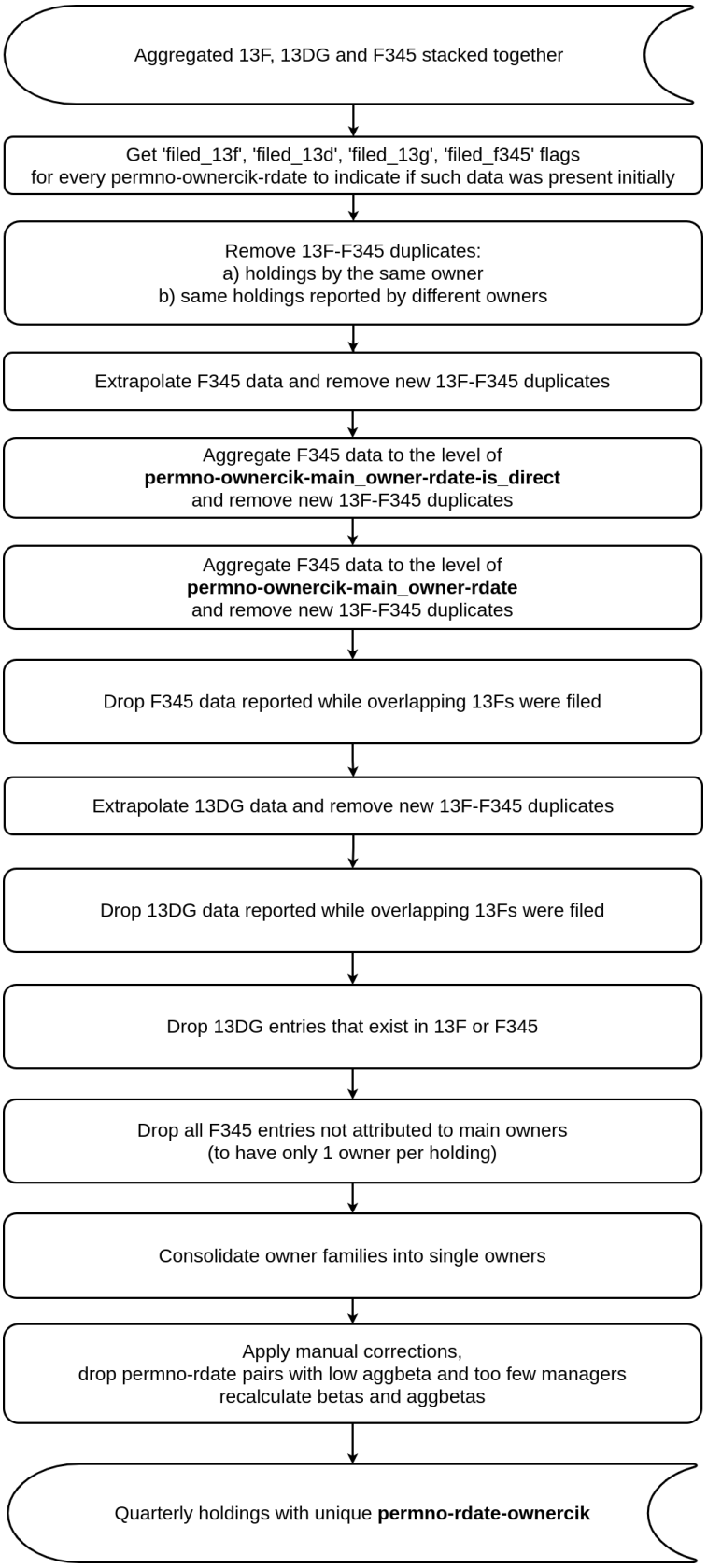


Figure A.22. Merging process

Electronic copy available at: https://ssrn.com/abstract=4059513

european corporate governance institute

about ECGI

The European Corporate Governance Institute has been established to improve *corporate governance through fostering independent scientific research and related activities.*

The ECGI will produce and disseminate high quality research while remaining close to the concerns and interests of corporate, financial and public policy makers. It will draw on the expertise of scholars from numerous countries and bring together a critical mass of expertise and interest to bear on this important subject.

The views expressed in this working paper are those of the authors, not those of the ECGI or its members.

www.ecgi.global

european corporate governance institute

ECGI Working Paper Series in Finance

Editorial Board

Editor	Mike Burkart, Professor of Finance, London School of Economics and Political Science
Consulting Editors	Renée Adams, Professor of Finance, University of Oxford Franklin Allen, Nippon Life Professor of Finance, Professor of Economics, The Wharton School of the University of Pennsylvania Julian Franks, Professor of Finance, London Business School Mireia Giné, Associate Professor, IESE Business School Marco Pagano, Professor of Economics, Facoltà di Economia Università di Napoli Federico II
Editorial Assistant	Asif Malik, Working Paper Series Manager

www.ecgi.global/content/working-papers

european corporate governance institute

Electronic Access to the Working Paper Series

The full set of ECGI working papers can be accessed through the Institute's Web-site (www.ecgi.global/content/working-papers) or SSRN:

Finance Paper Series	http://www.ssrn.com/link/ECGI-Fin.html
Law Paper Series	http://www.ssrn.com/link/ECGI-Law.html

www.ecgi.global/content/working-papers