Mavericks, Universal, and Common Owners - The Largest Shareholders of US Public Firms*

Amir Amel-Zadeh† Fiona Kasperk‡ Martin Schmalz§

July 10, 2022

Abstract

We construct a novel data set to show that, between 2003-2020, up to one-fifth of America's largest firms had a non-financial blockholder or insider as their largest shareholder. 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. Consolidation in the asset management industry increases universal ownership and common ownership of industry rivals. Extant results claiming indexing alone explains the rise of universal ownership cannot be confirmed with the new, more comprehensive data.

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 Philip Bond, 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 for valuable feedback. Vladimir Fedoseev 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.Amelzadeh@sbs.ox.ac.uk. Park End Street, OX1 1HP, U.K.

‡Fiona.Kasperk@sbs.ox.ac.uk. Park End Street, OX1 1HP, U.K..

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

1 Introduction

The apparent rise of the size and concentration of the asset management sector has attracted much recent attention by politicians, policy makers, and academics alike. One side hopes that "universal investors," 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). Another view fears precisely that scenario and has sponsored actions to reduce the concentrated power in the hands of BlackRock and its ilk.[1] Yet others presume that potential anti-competitive effects of common ownership of industry rivals that have been shown in particular product markets in particular industries may extend across the economy, and thus made far-reaching policy proposals to reorganize the asset management industry. Some asset managers declared that these proposals, if enacted, would be invasive enough to threaten their business model.[2]

In stark contrast to the advanced stage of this theoretical debate, we know very little about how large or influential these "universal investors" or "common owners" – i.e. widely diversified institutional asset managers – are, compared to other large investors with interests in only one firm. While market-level intra-industry research has hand-collected ownership data for all types of owners (e.g., Azar et al., 2018; Backus et al., 2021a), broad cross-industry studies that attempt to assess drivers and effects of common ownership at the firm- or industry- level are often based on a particular subset of ownership records, namely 13-F filings (e.g., Backus et al., 2021b; Koch et al., 2020; Lewellen and Lowry, 2020; Gilje et al., 2020). These filings cover institutional investors that manage $100 million or more in equity securities, but do not typically cover corporate insiders or non-institutional blockholders who own more than 5% of the equity of a company.

Below we illustrate why this data limitation potentially not only constrains our understanding of the portfolio interests of the largest shareholders of America's largest firms, but also calls into question the validity of the conclusions reached by this literature to date. Consider calculating a measure of common ownership, which captures to which extent the

[1] https://www.wsj.com/articles/lawmakers-seek-to-curb-voting-power-of-blackrock-vanguard-and-other-big-asset-managers-11652875481

[2] See Posner et al. (2016) for a prominent policy proposal; see Schmalz (2018); Backus et al. (2019); Schmalz (2021, 2022) for reviews of the literature.

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	Form 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. 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 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 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.

This example illustrates 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.[3]

We address the question to which extent the intuition from the above example generalizes to America's largest publicly traded firms. In particular, we first ask: who are America's

[3]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.

largest firms' largest shareholders, and what are these shareholders' portfolio interests? Second, what are the implications for the ownership structures of America's largest firms? In particular, how much "universal ownership" across all firms, and how much "common ownership" across industry rivals is there? Third, which role do blockholders and insiders play in explaining the variation across firms and over time in universal and common ownership? Lastly, does the consolidation of asset managers increase measures of universal and common ownership, or does the increase of textbook "indexing" alone explain the changing patterns of corporate ownership in America?

To address these questions, 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 filer type is the largest shareholder of a firm and of how many firms. We also examine how diversified the different filer types' portfolios are. 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."

Finally, we regress the measures of universal and common ownership on holdings by the different filer types as well as measures of textbook indexing from the literature. We also 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 non-financial blockholder or insider at or near the top of the shareholder registry. Second, whereas exceptions exist, 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 but in-

vestor concentration is higher once we account for the holdings of blockholders and insiders. Furthermore, 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 firms rather than concentrating holdings in one industry). Lastly, we find that consolidation through mergers in the asset management industry does increase both common ownership and universal ownership. Moreover, the holdings of the "Big Three" institutional asset managers appear to increase universal and common ownership above the level explained by textbook indexing. This finding contrasts with an earlier finding in the literature based on 13-F records alone.

A policy-relevant take-away from our analysis is 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 blockholders, but we show that in practice blockholders tend to be "mavericks" instead. Conceptually, whether blockholders contribute to common ownership depends on the blockholders' diversification.

In sum, we find that 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 data set of all types of owners of the largest publicly traded firms in America. We thus enable better measurement of universal ownership, common ownership, and its drivers.

Our contribution to the literature on common ownership most closely relates to Backus et al. (2021b) (BCS), who offer the most recent evidence on common ownership among the largest U.S. publicly traded firms.[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

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

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 footnote 12 of 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, Schmalz, and Tecu (2018) do when they compute the profit weights for airlines and as Backus, Conlon, and Sinkinson (2021a) do when they compute the profit weights for cereal." 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.

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

2 Data set construction

Building on the method of Backus et al. (2021b) we scrape and parse the ownership of S&P500 firms by institutional investors, corporate insiders as well as active and passive blockholders. We thereby utilize all ownership reports required by the Securities Exchange Commission (SEC) for investors in US public corporations. The fraction of share ownership that remains not captured by any filing is attributed to retail investors.

[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 72% 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 detail these differences throughout the paper and the appendix.

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 an corporate insider. A Form 3 need not yet report any equity ownership. Subsequently, whenever the insider changes his holdings of company equity, he is 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. 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 2000 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. Until 2003, the number of corporate insiders reporting equity ownership and their proportion of stock owned is negligible. However, before 2004, the most significant corporate insiders (holding >5% outstanding equity) were required to file 13-D or 13-G

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

reports and we therefore capture them through the added blockholder filings. Due to the limited number of insiders reports filed before June 2003, we focus our analysis exclusively on the period from June 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.1 how our approach differs from the methodology in Backus et al.. Once completed, the parsed data provides the Central Index Key (CIK) identifying 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 is that one report can be filed by multiple reporting owners. For example when the Walton siblings report their holdings in Walmart, this 13G will be filed jointly by Robson, John and Helen Walton, as well as the Walton Family Trust. Each sibling owns and votes on 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 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 share ownership 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, equity securities and derivatives are reported in separate tables and in each the filing owner provides a non-standardized, textual description of the security, 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. This routine is described in more detail in Appendix A.3. Following BCS, we only keep common stock holdings. As our sample is currently restricted to firms with single class common stock, we the collect the unique common stock class for each firm. 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 Walton Family Trust and siblings jointly filing Form 13-G, 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 issue is to adjust 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 previous step, this interpolation will not create any new duplicates.

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. We then spot check whether the resulting ownership structures are economically sensible (for example, ensuring

that we do not capture more than 100% of common stock outstanding), and also cross-check our results against various commercial data bases for consistency.

The completed data sample contains ownership structures of 932 firms and slightly less than 500 firms each quarter.

3 Results

3.1 America's largest shareholders and their portfolio interests

The average S&P500 firm is mostly owned by 13-F filing shareholders. Institutional investors own about 83% of outstanding equity in the average firm in 2019, while insiders and blockholders hold on average about 5%, respectively (Figure A.2). The cross-sectional variation in ownership, however, demonstrates much larger ownership stakes of blockholders and insiders in some firms. For an accurate measurement of the variation in ownership structures these investors cannot be neglected. Figure 1 plots the cross-section of ownership stakes held by 13-F filers, 13-D/G filers and Form 3/4/5 filers in each of the S&P 500 firms. Whilst the proportion held by 13-F filers is large in every case, there is a non-negligible number of firms where insiders or blockholders hold between 30-50% of shares. If we used only 13-F institutional ownership data, as provided by Thomson Reuters for example, we would be missing this variation.

In some cases there are multiple blockholders or insiders who jointly hold a large stake in the firm (e.g. the Walton Family Trust and the individual Walton siblings jointly own a large stake in Walmart), in other cases it is a single shareholder who owns a controlling stake (e.g. Randall Rollins in Rollins Inc.). In order to establish what owner is *the* largest owner of a firm, we plot the share of S&P500 firms where a 13-F filer, a 13-D filer, a 13-G filer, or an insider is the largest shareholder (Figure 2). The general conclusion from the cross-sectional graph is reinforced: individual 13-F investors are the largest shareholder in 80% of S&P500 firms in 2003 and increase that share to 90% of firms in 2020. Only in a fifth all firms in 2003 are blockholders and insiders the largest shareholder and this proportion declines to a tenth of firms in 2020. Assessing the ownership structures of America's largest firms without insiders and blockholders in our data, would miss the largest owners of these 10-20% of firms.

Omitting blockholders and insiders is especially dangerous in an analysis of overlapping ownership across firms, because these shareholders hold systematically different portfolios from 13-F institutional investors. As depicted in Figure 3 blockholders and insiders are mainly undiversified, henceforth referred to as "maverick" owners, that hold shares in a single firm in a given quarter. Only about 10% of blockholders and insiders are diversified, meaning they own multiple firms in a given quarter. 13-F filing investors, in contrast, are mostly diversified (95% of filers) and only a small fraction of less than 5% are maverick owners. We further distinguish between "common owners", describing an investor who owns shares in multiple firms competing in the same industry at a given point in time, and "universal" investors, who own at one moment in time more than 95% of all firms in the sample. Neither type of owner is found among blockholders or insiders. But more than 80% of institutional filers are common owners and another 5% of them are universal owners every quarter. Thus by neglecting insiders and blockholders we miss most of the maverick owners in American corporations, biasing our assessment of overlapping, common and universal ownership in the economy.

Another group of investors that we do not want to miss are the Big Three institutional asset managers, Vanguard, BlackRock and State Street. They feature regularly in the news, holding on average more than 20% of outstanding equity in any S&P 500 firm in late 2019 (Figure A.4). All three are universal and common owners since 2010. Unlike other universal shareholders, however, the Big Three are also the largest block of owners in about 70% of S&P 500 firms in 2019 (Figure A.5). Comparing Figure 2 and Figure A.5, the dominance of 13-F filing shareholders seems only partly attributable to the Big Three. Other institutional investors must be contributing as well.

Size and portfolio diversification vary widely for different filing shareholders, but what are the implications for corporate ownership structures? Repeating the exercise from Figure 1, we plot the cross-section of ownership stakes based on degree of diversification instead of filer type in Figure A.7. The figure demonstrates that in 2019 about 45-60% of equity of every S&P500 firm is owned by universal investors. This is more than the average share owned by the Big Three, implying that they are not the only universal investors in the sample. Another 20% of equity is held by other diversified investors. The remainder of between 0-20% of equity is controlled by maverick owners. A few firms have maverick

ownership stakes between 30-50%, some do not have any visible maverick shareholders. Over time (comparing with Figure A.8) the share of equity held by universal investors has been growing, because of a lower retail share of ownership (shareholdings that are not reported to the Securities Exchange Commission), because of lower ownership by diversified, non-universal investors, and due to declining maverick ownership. The remaining cross-sectional variation in ownership structures seems to be driven by mavericks, and is highly correlated with the ownership stakes of insiders and blockholders.

The prevalence and trends of common ownership are actively discussed for the potential anti-competitive effects and we therefore analyze the cross-section of ownership structures again, differentiating diversified and common owners. Whilst universal owners by definition appear among the owners of every firm in the sample, common owners will never appear in the cross-sectional graph for firms that do not have an industry competitor among the S&P500 sample. Figure A.9 plots the 2019 ownership structures of America's largest firms, this time with stakes held by common owners, diversified owners and mavericks. The ownership of mavericks and diversified investors do not change compared to Figure A.7. However, the share of diversified owners, who are not common owners is surprisingly small. In most firms that have an industry competitor within the S&P500, common shareholders own the largest proportion of 60-80% of equity. Ownership by mavericks and diversified, non-common investors becomes the main driver of variation in ownership structures.

The relative importance of universal and common ownership over time is depicted in Figure A.6. We find that the share of S&P 500 firms where the largest shareholder is a common owner has been increasing gradually from 58% in 2003 to 65% in late 2019. The share of firms where a universal owner is dominant has risen more markedly from 40% to almost 80% over the same period, overtaking common owners around 2014. If the anti-competitive incentives of common and universal shareholders differ, this may have an important impact on firm conduct. Mavericks are much less frequently the dominant shareholder in S&P 500 firms, and over time that frequency is declining (from 15% to about 5% between 2003-2019). Diversified investors, are the largest shareholder in the remainder of firms.

Thus far we have demonstrated that blockholders and insiders are systematically less diversified than 13-F institutional investors, and represent an important fraction of share-

holdings in many S&P 500 firms. Omitting them will lead to mis-measurement of the extent of universal and common ownership among America's largest firms and may bias the variation in ownership structures and our analysis of its drivers.

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", measures to which extent the most influential shareholders in firm A are invested 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 shareholders.

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 codes).

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

To analyze the drivers of these profit weights, we follow Backus et al. and decompose κ into two components. the contribution of ownership structure similarity, and the contribution of relative investor concentration.

$$\kappa_{ab} = \frac{\sum_s \beta_{as}\beta_{bs}}{\sum_s \beta_{as}\beta_{as}} = \frac{\langle \beta_{as}\beta_{bs} \rangle}{\langle \beta_{as}\beta_{as} \rangle} \tag{2}$$

Conveniently, this allows us to also decompose the variance of profit weights into its

contributions by ownership similarity and relative investor concentration.

$$\kappa_{ab} = \underbrace{\cos(\beta_a, \beta_b)}_{\text{ownership structure similarity}} \cdot \underbrace{\sqrt{\frac{IHHI_b}{IHHI_a}}}_{\text{relative investor concentration}} \tag{3}$$

$$Var(ln\ \kappa_{ab}) = Var(ln\ cos(\beta_a, \beta_b)) + Var\left(ln\ \sqrt{\frac{IHHI_b}{IHHI_a}}\right) + 2Cov\left(ln\ cos(\beta_a, \beta_b),\ ln\ \sqrt{\frac{IHHI_b}{IHHI_a}}\right) \tag{4}$$

The similarity of firms' ownership structures is captured by the cosine angle between the vectors of their respective ownership structures. If two firms' ownership structures are very similar, they have many owners in common who have financial interests in both firms. If these owners' holdings in both firms are additionally of equal size, then the common owners will have an equal incentive to care about both firms' profits and should completely internalize the externalities of one firm to the other. More similar ownership structures therefore increase the cosine similarity statistic and thereby the profit weight κ between two firms.

The relative investor concentration component measures the extent to which the shareholders in one firm have the ability to influence corporate actions relative to shareholders in the other firm. For firms a and b, for example, if firm a has a concentrated ownership structure, while firm b's owners are highly dispersed, then firm a shareholders will have more control over corporate strategies. This component is the directed part of the profit weights that implies that the weight firm a attributes to firm b profits will not be the same as the weight attributed by firm b to firm a profits. More concentrated ownership in firm b and less concentrated ownership in firm a will ceteris paribus increase the profit weight firm a attributes to firm b.

Based on these two components we investigate the extent to which blockholders, corporate insiders and asset manager concentration change the contribution of ownership structure similarity and relative investor concentration to profit weights.

On average common ownership profit weights are higher than universal ownership profit weights and both have increased significantly over time. Figure 4 depicts an increase in universal ownership from an average weight of 0.42 on other firm profits in 2003 to an average weight of 0.67 in 2020. Limiting our measurement to common ownership only, we

find that the average profit weight on firms in the same industry is consistently about 0.05 higher than for universal ownership and increased from 0.49 to 0.74 in our sample period.

The inclusion of insiders and blockholders with systematically lower portfolio diversification reduces both our measures of common and universal ownership. To analyze the effect of completing the data, we calculate profit weights first based only on 13-F institutional ownership records and then with the comprehensive data comprising institutional investors, blockholders and insiders. In the time series, we find that using only 13-F institutional ownership data leads to a systematic over-estimation of universal and common ownership. The downward shift in average profit weights over time in Figures 5(A) and 5(B) is almost parallel, suggesting that the lower diversification of insiders and blockholders is persistent over time.

Plotting only the quarterly average profit weights veils the cross-sectional variation introduced by completing the ownership data. To analyze how profit weights change for individual firm-pairs we plot a histogram of the percentage changes in universal and common ownership kappas for all firm-pairs across all quarters in Figures 6(A) and 6(B). The overwhelming effect of the blockholder and insider addition is a decrease of most profit weights by between 5% and 25%. A small fraction of kappas is reduced much more substantially, suggesting we missed very important, undiversified owners in these firm pairs. Another small proportion of profit weights actually increased due to including diversified blockholders. Overall the changes in universal ownership kappas are more nuanced than the changes in common ownership kappas, as the tails of the distribution common ownership changes are more pronounced.

We further decompose profit weights into ownership structure similarity and relative investor concentration and find a decreased ownership similarity is the key driver of lower universal and common ownership. The less diversified insiders and blockholders, who in some cases hold significant ownership stakes, reduce the cosine similarity between ownership vectors of commonly and universally owned firms. In the cross-section there remain a few cases where ownership similarity increases, but the majority of firm-pairs has between 1% and 30% less similar ownership structures (Figures A.13, A.14). The time-series averages of ownership similarity in Figures A.11 and A.12 mirror the trend and changes of the average profit weight. Ownership similarity of universally and commonly owned firms is increasing

over time, but including blockholders and insiders in the data reduces the average similarity almost in parallel for all quarters. The parallel shift suggests that insiders and blockholders are not becoming more or less diversified over time, but sustain their portfolio characteristics, which are different from those of 13-F filers.

The change in average relative investor concentration due to the addition of insiders and blockholders instead is small. The cross-sectional histogram of changes depicts both positive and negative adjustments in similar proportions (Figure A.15) and the time-series is not much affected.

Adding blockholders' and insiders' shareholding to the previously used 13-F institutional ownership data, implies the inclusion of owners with systematically different portfolios, resulting in more heterogeneous ownership structures. Relative investor concentration measures are not affected by the addition, but the relationship between ownership concentration and similarity becomes significantly negative, because the most concentrated owners are now more frequently maverick owners. Overall this leads to a persistent reduction in our measures of universal and common ownership.

3.3 Drivers of universal and common ownership

Three candidate drivers of overlapping ownership are discussed and evaluated in the literature: investor indexing, Big Three holdings and centralization of voting across funds, and mergers in the asset management industry. Using linear regression analysis we investigate whether these candidate drivers positively contribute to our measures of universal and common ownership, even when based on the complete corporate ownership structures.

We regress the profit weight attributed by firm a to firm b profits on the aggregate holdings of the "Big 3" institutional investors (BlackRock, State Street, Vanguard) in firm a and a measure of investor indexing first devised in Backus et al. (2021b). The indexing variable measures the distance between the average investors' portfolio and the market portfolio weighting all firms by their market capitalization. We control for firm a retail share of ownership (the share of its equity ownership that we do not capture), its market capitalization and its operating margin. To account for the potential firm-specific persistence of ownership structures and period-specific economic conditions (state of financial markets) we also include firm and quarter-year fixed effects. Because our profit weights are firm-pair measures and

there may be firm-pair specific differences in ownership structures that are persistent over time, we include ordered firm-pair fixed effects (regression on κ_{ab} has a different fixed effect than regression on κ_{ba}) and quarter-year fixed effects in an alternative specification. The remaining variation in our sample should be driven only by selective changes in ownership stakes by shareholders, that are not driven by general economic conditions.

Our regression results in Table 2 yield positive and statistically significant coefficients on both investor indexing and Big Three holdings in every specification. This suggests that both indexing and the concentration of holdings by the Big Three contribute positively to universal and common ownership. The contribution of investor indexing is more economically significant than the contribution of the Big Three holdings. But in contrast to findings of Backus et al. (2021b), our coefficient on the Big Three remains positive and statistically significant despite controlling for indexing. This difference is not driven by the addition of insiders and blockholders in the data, but rather by our more accurate parsing of 13-F filings and a more comprehensive consolidation of fund families into one entity. Based on our results we *cannot* conclude that Vanguard, BlackRock and State Street *do not drive* universal and common ownership, over and above investor indexing.

We also find that larger firms and firms with a larger retail share of ownership attribute higher profit weights to other firms, both for universal and common ownership. The positive coefficient on market capitalization is likely driven by more institutional shareholdings in larger firms, leading to more diversified owners and higher profit weights. A larger share of retail ownership implies less shares are ultimately voted in our sample, implying the denominator of profit weights decreases and the profit weights increase.

Adding the aggregate holdings of insiders, blockholders and passive blockholders in firm a to our right hand-side explanatory variables, we validate the significant negative effect these filers have on our measures of universal and common ownership. The inclusion of these variables does not affect our findings on indexing and the Big Three holdings.

To analyze the role of asset management mergers in driving universal and common ownership we focus on the effects of the BlackRock - Barclays Global Investors merger in 2009.

The effect of the merger on profit weights is demonstrated by plotting the change in universal and common ownership measures between counterfactual "implied" profit weights against the actual profit weights pre-merger. Pre-merger profit weights are calculated using

the baseline ownership structures for Q1 2009, with the actual ownership stakes of BlackRock (henceforth BLK) and Barclays Global Investors (henceforth BGI). The "implied" profit weights are calculated on the ownership structures for Q1 2009, assuming that BLK and BGI are already merged into one entity (i.e. we sum up the individual BLK and BGI holdings in every firm).

The histograms of implied changes in profit weights demonstrate that universal ownership and common ownership would increase by 1-15% for most firm-pairs due to the merger (Figures 7(A), A.18(A)). For a sub-sample of firm-pairs the hypothetical merger implies a reduction in profit weights, suggesting that the relative investor concentration increases more for the base firm than for the other firm.

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 and thereby increase kappas. The hypothetically aggregated ownership stakes do imply an increase in investor concentration (Figure A.20, but this increase will be reflected twice in the profit weights (κ_{ab} increasing due to higher investor concentration in b coincides with κ_{ba} decreasing for the same reason) and therefore cannot drive a sustained increase in universal or common ownership, either. Instead, the merger increases the similarity of ownership structures across firm-pairs by 1-15% in most cases and thereby raises our measures of common and universal ownership (Figure A.19).

An open question is whether the implied profit weight increases by the merger correlate in any way with the actual changes in universal and common ownership post-merger. Plotting the changes in profit weights between the post-merger quarter (Q1 2010) and the pre merger quarter (Q1 2009) in Figure 7(A), the positive change is much less evident, and profit weights seem to increase and decrease arbitrarily.

To answer this question we regress the post-merger change in profit weights on the merger-implied profit weight changes. Our main specification of post-merger profit weight changes refers to the difference between kappas calculated for all firm-pairs in Q1 2010 and kappas calculated for the pre-merger quarter Q1 2009. We additionally analyze the immediate merger effects on profit weight changes in Q4 2009 and the more persistent effects on Q1 2011 and Q1 2012. We include an intercept to control for time fixed effects (as we are looking only at one quarter in each regression) and control for industry and firm fixed

effects, to eliminate any persistence in industry- or firm-specific ownership stakes of BLK and BGI. The remaining variation in post-merger profit weight changes is driven solely by the cross-industry variation in universal and common ownership.

Our results in Table 3 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 finding holds both for common ownership profit weights and universal ownership profit weights. Even when including additional controls for the pre-merger ownership similarity and investor concentration, market capitalization and aggregate ownership stakes of insiders, blockholders and retail investors, the coefficient on implied kappa changes remains unchanged in magnitude and statistical significance (Table **??**).

For robustness we also investigating 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.

The analysis suggests that investor indexing, the aggregate holdings of the Big Three and consolidation in the asset management industry all play a role in increasing common and universal ownership. The concentration of ownership by the Big Three contributes positively and significantly to our measures of common and universal ownership, even when controlling for investor indexing. The positive effect of asset management mergers is due to an increase in ownership structure similarity, rather than due to the creation of new common or universal ownership.

4 Conclusions

We demonstrate that the size and portfolio diversification of 13-F filing investors on the one hand and those of 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 is likely to lead to bias in analyses of the effect of corporate ownership structures on various outcomes. We show that adding these less diversified owners to the

regularly analyzed sample of 13-F institutional ownership reduces estimates of common ownership within industries and of universal ownership consistently. Based on a complete data set of corporate ownership we find more heterogeneous ownership structures and a more negative relationship between investor concentration and portfolio diversification. Furthermore, we find that common ownership within industries is higher than universal ownership, suggesting that it will be possible to reduce common ownership without reducing index investing, which is positively correlated with both. Over and above indexing, we find that the shareholdings of the "Big Three" institutional investors also contribute positively and statistically significantly to common and universal ownership, which is in contrast to previous findings in the literature. Lastly, our results suggest that mergers in the asset management industry predict future increases in both common and universal ownership, because they increase ownership structure similarity across firms.

The data set we construct 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 72% 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, the data we construct 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 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.

References

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

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

Backus, Matthew, Christopher Conlon, and Michael Sinkinson, "The common ownership hypothesis: Theory and evidence," *Economic Studies at Brookings*, 2019.

_ , _ , and _ , "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.

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.

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.

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.*

Posner, Eric A, Fiona M Scott Morgan, and E Glen Weyl, "A proposal to limit the anticompetitive power of institutional investors," *Antitrust LJ*, 2016, *81*, 669.

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).

_ , "Conceptual Breakthroughs on Common Ownership and Competition: A Framework For Evaluating Policy," in Marco Corradi and Julian Nowag, eds., *Intersections Between Antitrust and Corporate Law*, Cambridge University Press, 2022.

5 Figures & Tables

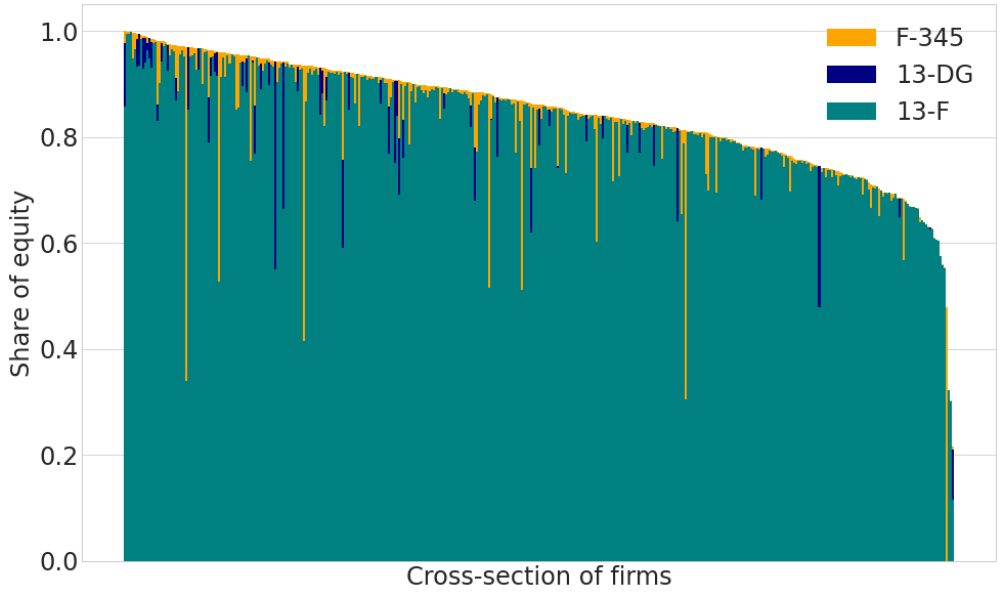


Figure 1. Ownership structures by filer type, 1 January 2020

Note: This figure depicts the captured ownership for each firm at the end of Q4 2019, differentiating ownership by 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 turquoise represents the share of equity ownership identified by parsing 13-F filings. The navy part of each bar measures 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.

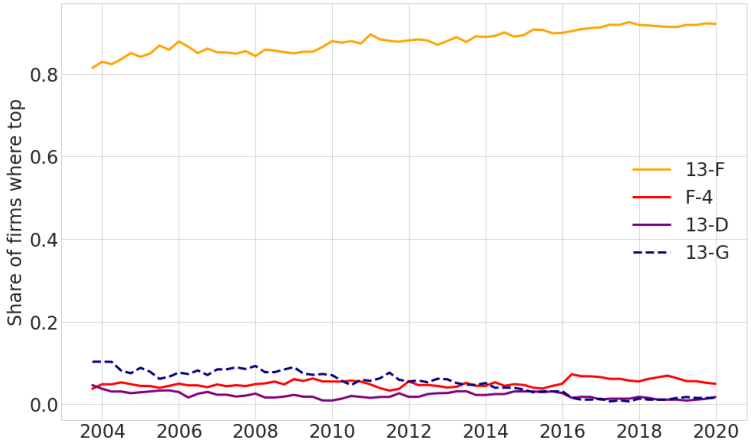


Figure 2. Top owner position by type of filer

Note: This figure depicts the quarterly share of firms in our sample for which a 13-F institutional filer, an insider, a 13-D filer or a 13-G filer reports the largest ownership stake.

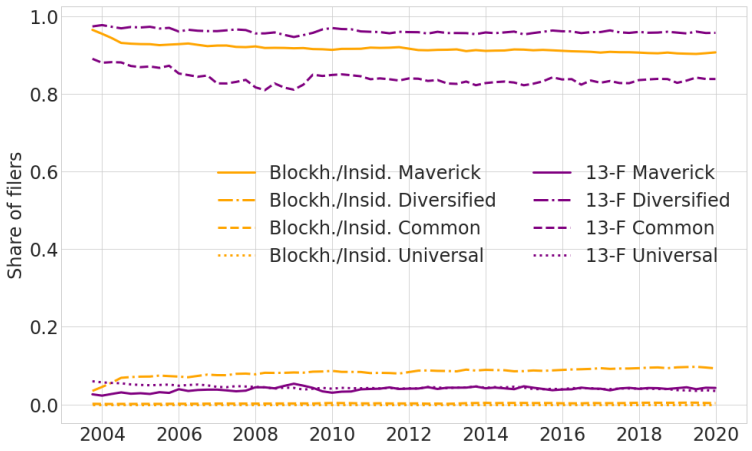


Figure 3. Diversification by filer type

Note: This figure depicts the quarterly share of 13-F institutional filers, and the quarterly share of all other filers that can be categorized as maverick owners (own 1 security in a quarter), as diversified owners (own multiple securities), as common owners (own multiple firms from the same industry) or as universal owners (own >95% of firms in the sample in a quarter).

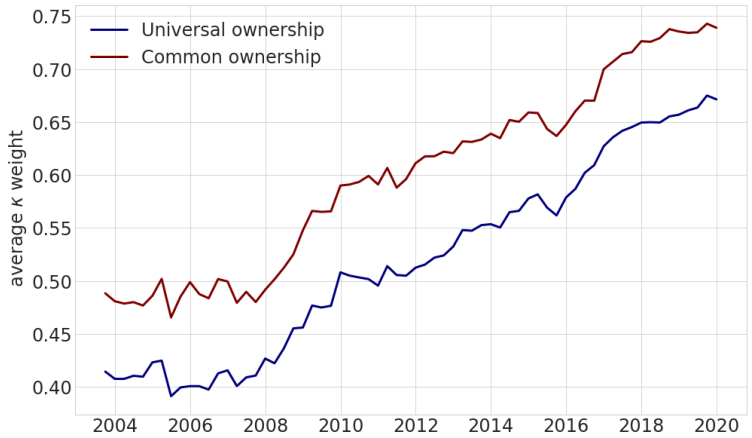
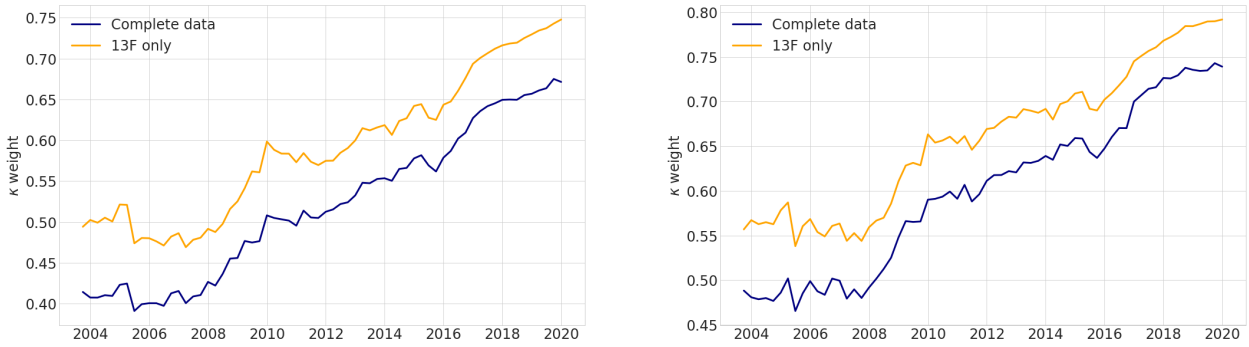


Figure 4. Average universal and common ownership profit weights

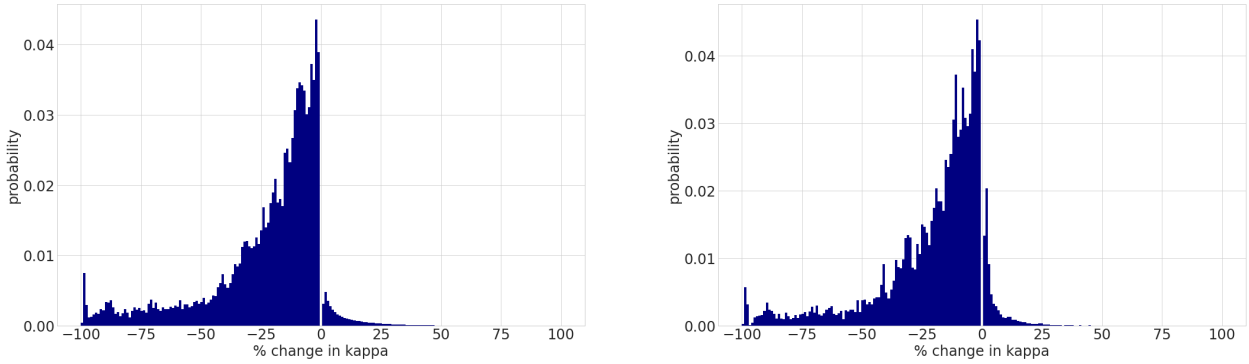
Note: This figure depicts the average universal ownership profit weights over time in blue and the average common ownership profit weights in dark red. The average common ownership profit weight is calculated using only profit weights of firm-pairs where both firms are in the same SIC industry. Profit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



(A) Universal ownership **(B)** Common ownership

Figure 5. Average profit weights with and without blockholders & insiders

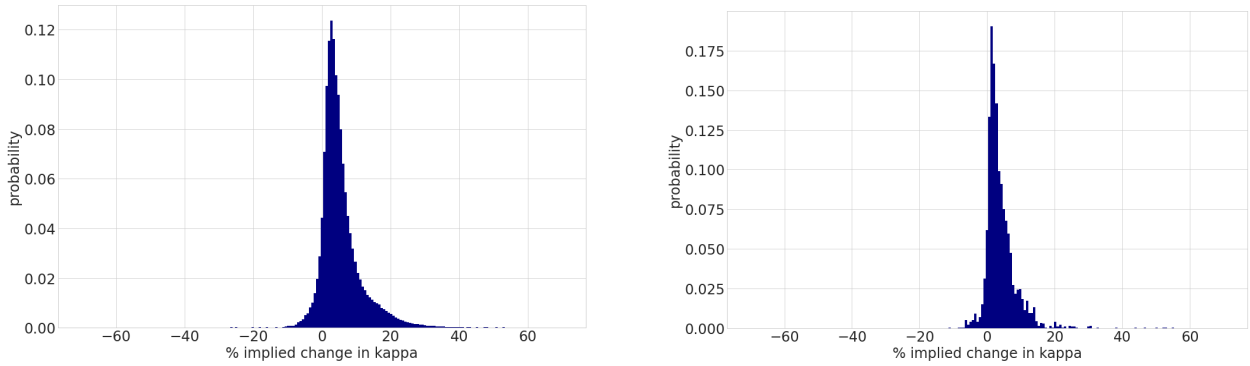
Note: This figure depicts the average universal and common ownership profit weights over time. The blue line represents average profit weights based on the complete data set, the yellow line represents averages based only on institutional ownership records. Profit weights are averaged each quarter across firms weighting all firms equally. Common ownership profit weights are averaged across firm-pairs with both firms in the same SIC 4 digit industry and weighting all firms equally.



(A) Universal ownership **(B)** Common ownership

Figure 6. 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. All kappa changes of less than 1% in either direction are ignored.



(A) Universal ownership

(B) Common ownership

Figure 7. Change in profit weights implied by merger

Note: This figure depicts the distribution of percentage changes in universal profit weights implied by the merger of BlackRock and BGI. The difference between each hypothetical "implied" profit weight, calculated based on 2009 Q1 ownership data but aggregating the BlackRock and BGI holdings, and its respective pre-merger profit weight using the baseline 2009 Q1 ownership data is considered. Universal ownership includes changes in all firm-pair kappas, common ownership includes changes in kappas for firm-pairs within the same SIC 4 digit industry.

	(1) κ_{univ}	(2) κ_{univ}	(3) κ_{univ}	(4) κ_{univ}	(5) κ_{univ}	(6) κ_{univ}
Insider Hold.	-0.249***		-0.338***	-0.225***		-0.312***
	(0.002)		(0.002)	(0.002)		(0.002)
13-D Holdings		-0.201***	-0.252***		-0.202***	-0.250***
		(0.002)	(0.002)		(0.001)	(0.001)
13-G Holdings		-0.129***	-0.207***		-0.121***	-0.194***
		(0.002)	(0.002)		(0.001)	(0.002)
Retail Share	0.590***	0.575***	0.553***	0.595***	0.581***	0.559***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Log(Market Cap)	0.058***	0.057***	0.057***	0.057***	0.056***	0.057***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Operating Margin	0.000**	0.001***	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Investor Indexing	1.081***	1.059***	0.962***	1.053***	1.026***	0.938***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Big 3 Holdings	0.050***	0.043***	0.087***	0.078***	0.073***	0.110***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
R-squared	0.622	0.622	0.623	0.796	0.796	0.797
R-squared Adj.	0.188	0.188	0.190	0.265	0.266	0.268
F-stat	493681	423307	374329	771036	661897	585180
p(F-stat)	0.000	0.000	0.000	0.000	0.000	0.000
Firm FE	✓	✓	✓			
Ordered Pair FE				✓	✓	✓
Quarter-Year FE	✓	✓	✓	✓	✓	✓
N	12799879	12799879	12799879	12799879	12799879	12799879

Table 2. Universal kappa regression on drivers

Note: This table presents results of a regression of universal ownership profit weights on aggregate holdings of corporate insiders (filing Form 3, 4, 5 reports), blockholders (filing 13-D reports) and passive blockholders (filing 13-G reports), the retail share of ownership (defined as 1 minus captured ownership fraction), log market capitalization, operating margin, aggregate holdings by the three largest institutional asset managers (BlackRock, Vanguard and State Street) and a measure of investor indexing (as suggested by Backus et al 2021). 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.

	2009Q4 $\Delta\kappa_{act,univ}$	2010Q1 $\Delta\kappa_{act,univ}$	2011Q1 $\Delta\kappa_{act,univ}$	2012Q1 $\Delta\kappa_{act,univ}$
$\Delta\kappa_{implied,univ}$	0.673***	0.476***	0.524***	0.552***
	(0.008)	(0.009)	(0.010)	(0.012)
R-squared	0.506	0.602	0.847	0.904
R-squared Adj.	0.049	0.020	0.017	0.014
F-stat	7542	3011	2518	2164
p(F-stat)	0.000	0.000	0.000	0.000
Firm FE	✓	✓	✓	✓
N	147840	147840	147840	147840

Table 3. Regression on post-merger universal kappa change

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 implied changes of 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.

6 Appendix

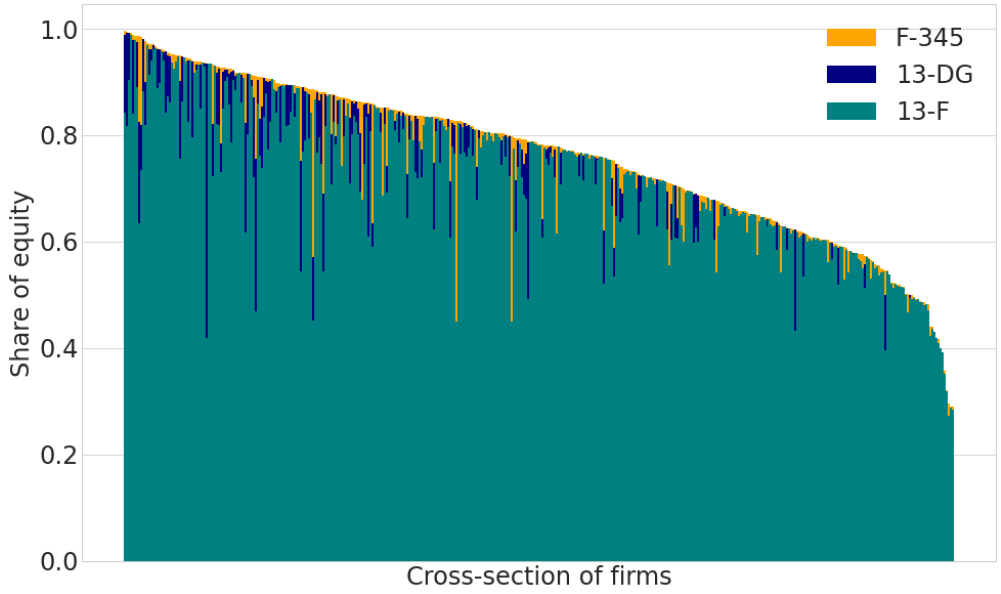


Figure A.1. Ownership structures by filer type, 1 January 2005

Note: This figure depicts the captured ownership for each firm at the end of Q4 2004, 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 turquoise represents the share of equity ownership identified by parsing 13-F filings. The navy part of each bar measures 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.

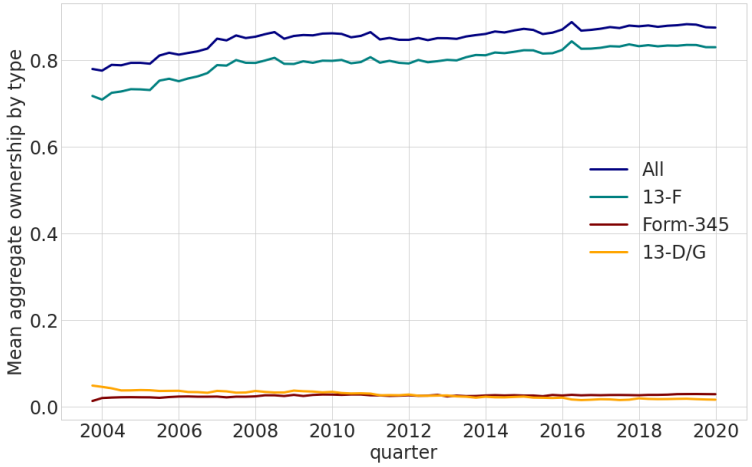


Figure A.2. Average aggregate ownership by filer type

Note: This figure depicts the aggregate share of outstanding equity owned by all 13-F institutional filers, all blockholders, all insiders and across all owners of the average S&P 500 firm for each quarter.

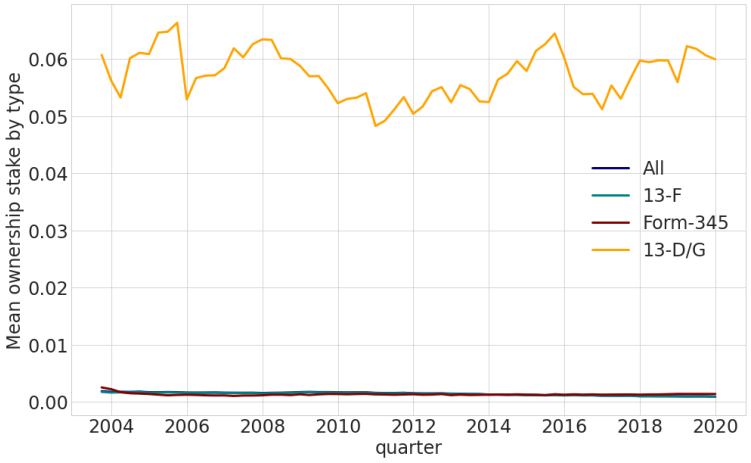


Figure A.3. Average ownership stake by filer type

Note: This figure depicts the average share of outstanding equity owned by an individual 13-F institutional filer, a blockholder, an insider and across all owners of an S&P 500 firm for each quarter.

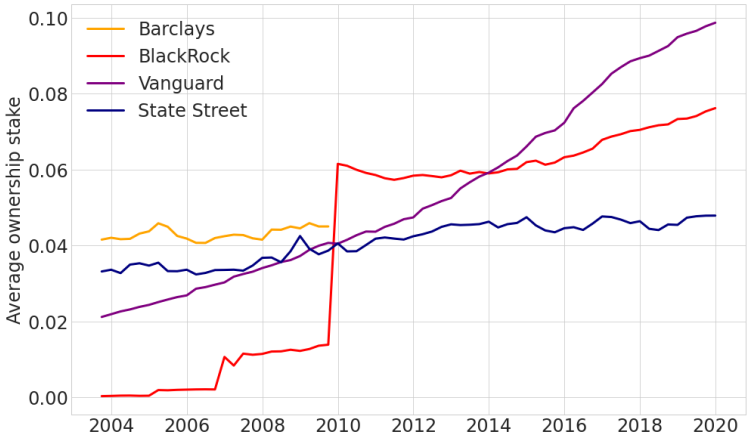


Figure A.4. Average ownership stake of Big 4

Note: This figure depicts the average share of outstanding equity of S&P500 firms owned by Vanguard, BlackRock, State Street, and Barclays Global Investors (until the merger with BlackRock) each quarter.

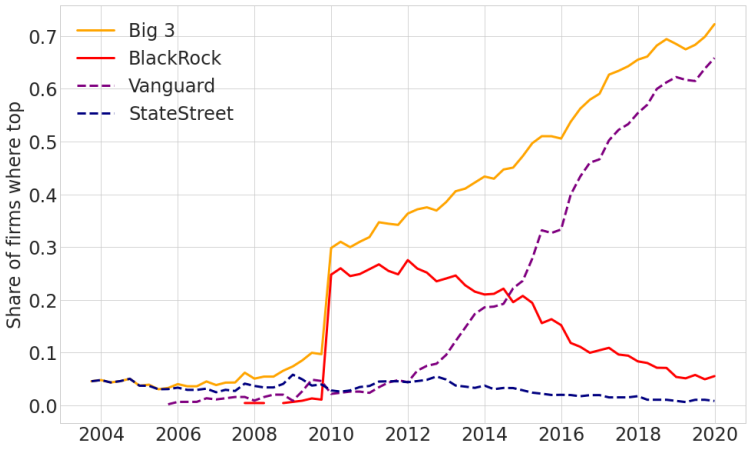


Figure A.5. Top owner position by Big 3

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

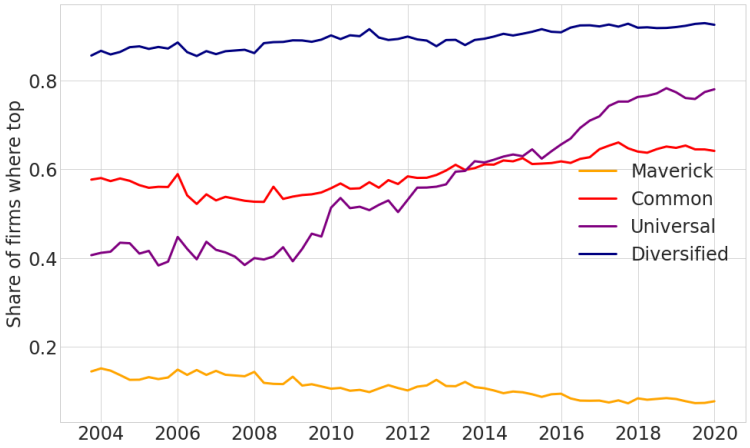


Figure A.6. Top owner position by diversification

Note: This figure depicts the quarterly share of firms in our sample for which a maverick owner, a diversified owner, a common owners or a universal owners reports the largest ownership stake. Maverick owners hold 1 security in a given quarter, diversified owners hold multiple securities per quarter, common owners hold multiple firms competing in the same industry and universal owners hold more than 95% of all firms in the sample.

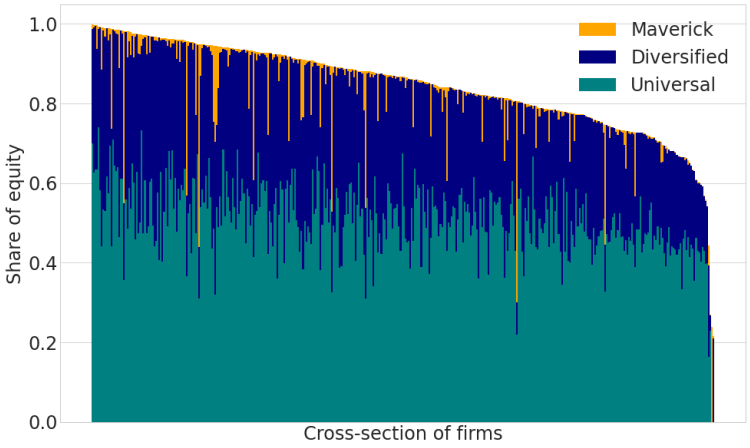


Figure A.7. Ownership structures by diversified-universal owners, 1 July 2019

Note: This figure depicts the captured ownership for each firm at the end of Q2 2019, 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 universal investors, where a universal investor own more than 95% of the securities in our sample in the given period. The navy part of each bar measures the holdings of diversified investors, who own shares in multiple firms, 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.

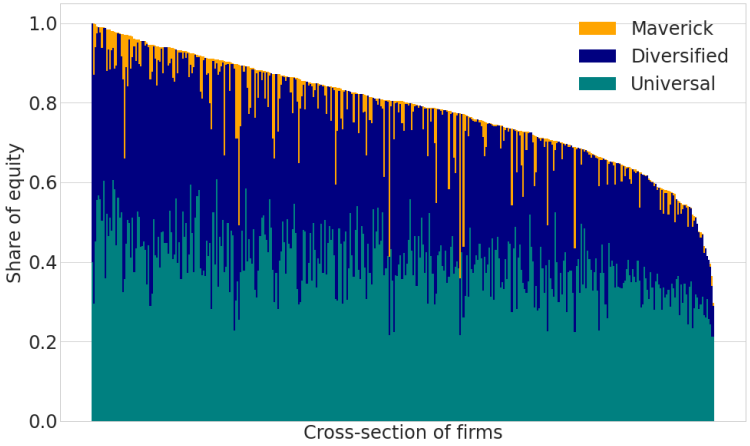


Figure A.8. Ownership structures by diversified-universal owners, 1 July 2006

Note: This figure depicts the captured ownership for each firm at the end of Q2 2006, 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 universal investors, where a universal investor own more than 95% of the securities in our sample in the given period. The navy part of each bar measures the holdings of diversified investors, who own shares in multiple firms, 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.

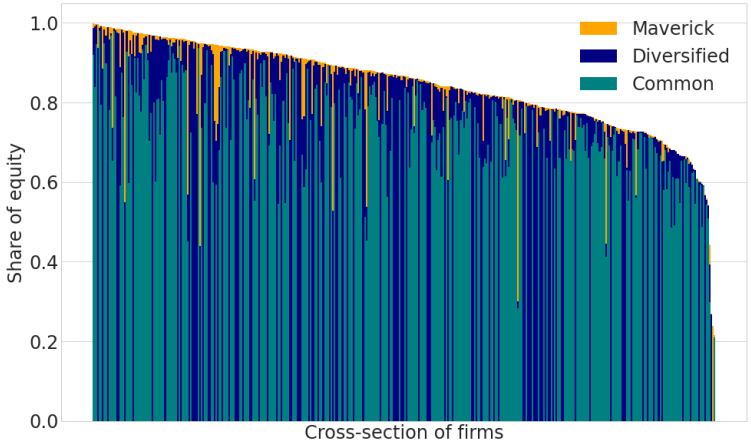


Figure A.9. Ownership structures by diversified-common owners, 1 July 2019

Note: This figure depicts the captured ownership for each firm at the end of Q2 2019, differentiating ownership by the diversification of the investor, considering the common ownership dimension. 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 common owners, that is all investors who coincidentally own shares in another firm in the same SIC 4 digit industry. The navy part of each bar measures the holdings of diversified investors, who own shares in multiple firms, but not in firms in the same SIC 4 digit industry. The orange part reports the holdings of undiversified shareholders, who own only shares in this company in the given quarter.

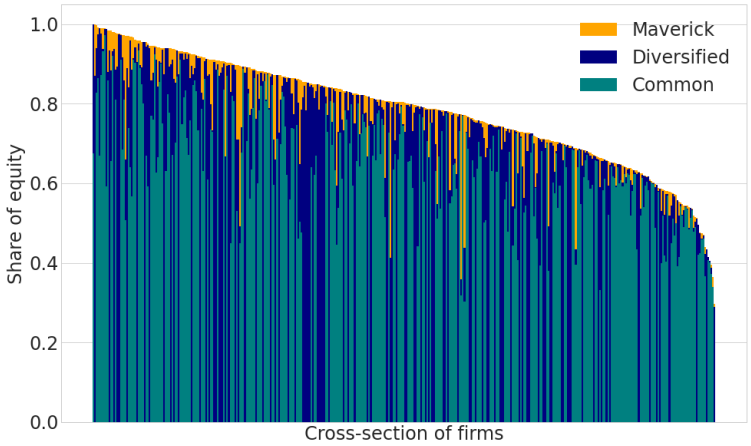


Figure A.10. Ownership structures by diversified-common owners, 1 July 2006

Note: This figure depicts the captured ownership for each firm at the end of Q2 2019, differentiating ownership by the diversification of the investor, considering the common ownership dimension. 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 common owners, that is all investors who coincidentally own shares in another firm in the same SIC 4 digit industry. The navy part of each bar measures the holdings of diversified investors, who own shares in multiple firms, but not in firms in the same SIC 4 digit industry. The orange part reports the holdings of undiversified shareholders, who own only shares in this company in the given quarter.

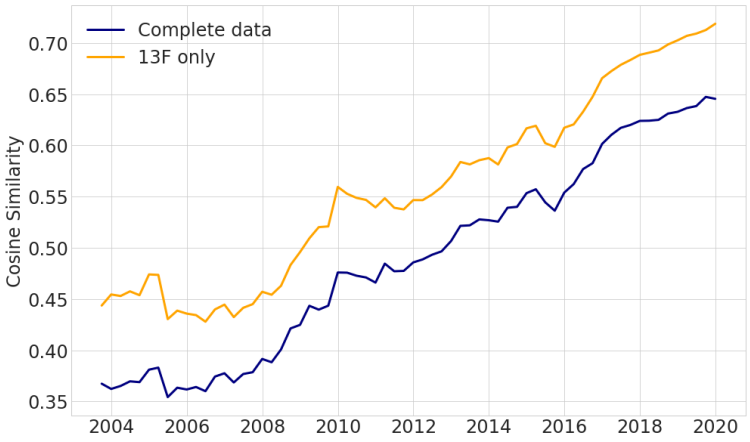


Figure A.11. Average universal cosine similarity with complete data

Note: This figure depicts the average universal ownership cosine similarity between ownership structures of firm pairs. The blue line represents averages based on the complete data set, the yellow line represents averages based only on 13-F institutional ownership records.

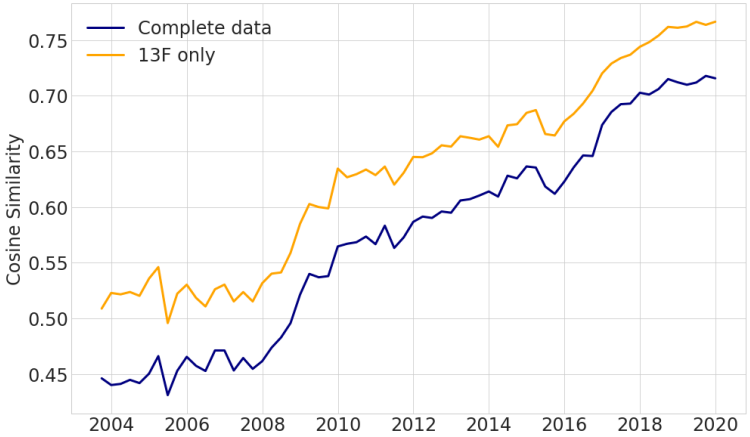


Figure A.12. Average common cosine similarity with complete data

Note: This figure depicts the average common ownership cosine similarity between ownership structures of firm pairs in the same SIC industry. The blue line represents averages based on the complete data set, the yellow line represents averages based only on 13-F institutional ownership records.

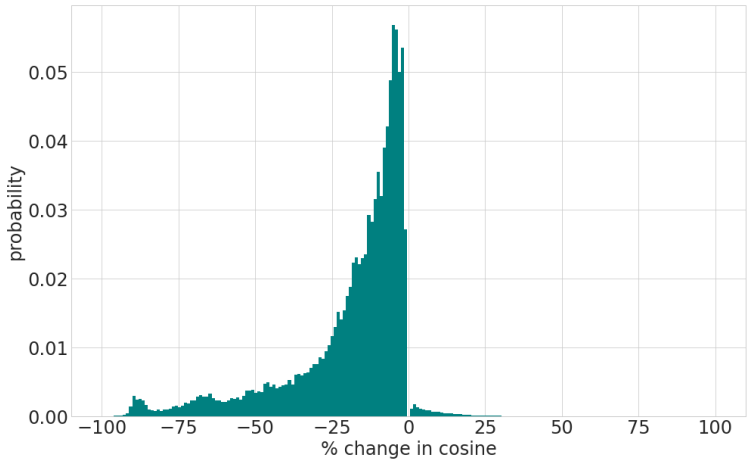


Figure A.13. Change in universal cosine similarity due to completing ownership data

Note: This figure depicts the distribution of percentage difference between ownership structure similarity of firm-pairs based on the complete data (including insiders and blockholders) and their ownership structure similarity based on the 13-F only data. Changes across all quarters and for all firm-pairs are plotted.

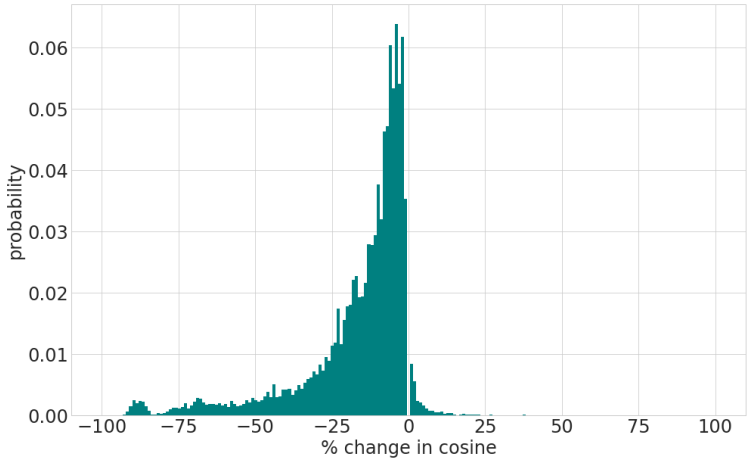


Figure A.14. Change in common ownership similarity due to completing ownership data

Note: This figure depicts the distribution of percentage difference between ownership structure similarity of firm-pairs based on the complete data (including insiders and blockholders) and their ownership structure similarity based on the 13-F only data. Changes across all quarters and for firm-pairs consisting of industry-competitors are plotted.

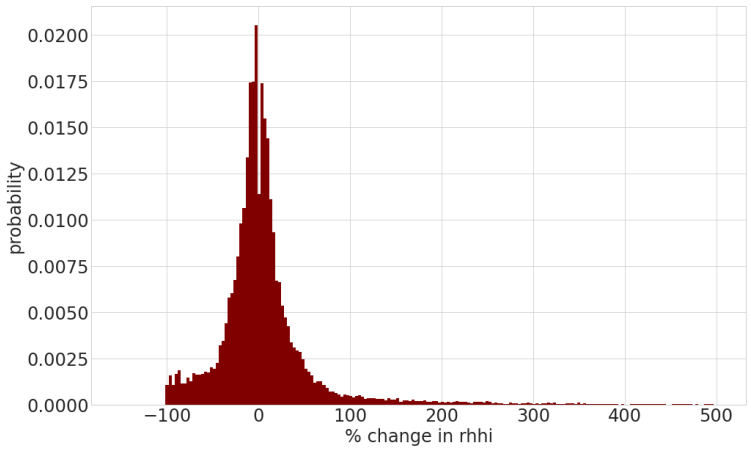


Figure A.15. Change in universal rel. investor concentration due to completing ownership data

Note: This figure depicts the distribution of percentage difference between relative investor concentration of firm-pairs based on the complete data (including insiders and blockholders) and the relative investor concentration based on the 13-F only data. Changes across all quarters and firm-pairs are plotted.

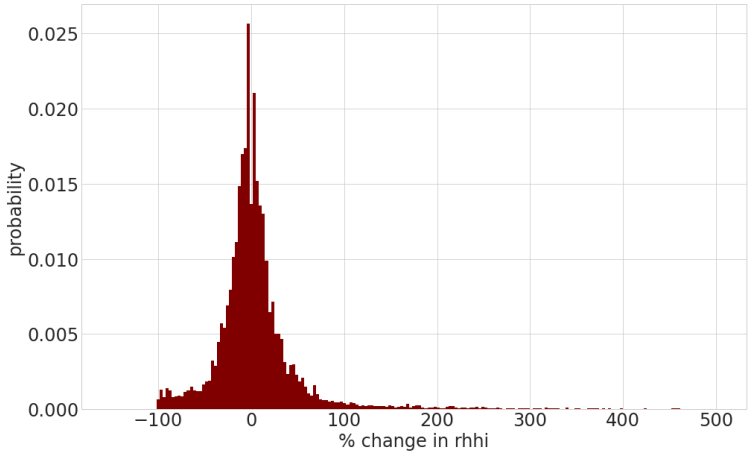


Figure A.16. Change in common rel. investor concentration due to completing ownership data

Note: This figure depicts the distribution of percentage difference between relative investor concentration of firm-pairs based on the complete data (including insiders and blockholders) and the relative investor concentration based on the 13-F only data. Changes across all quarters and for firm-pairs consisting of industry-competitors are plotted.

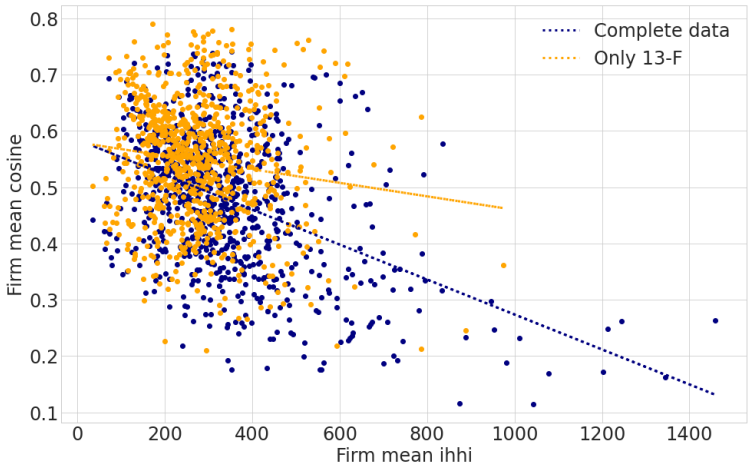
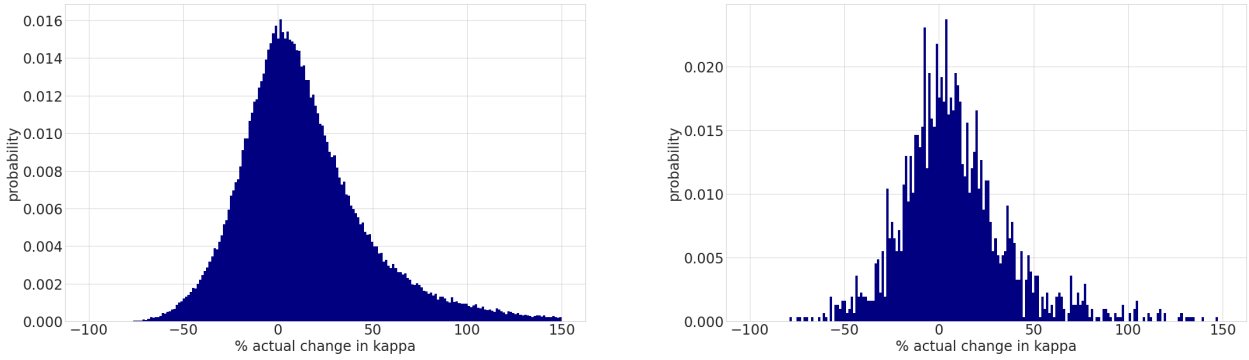


Figure A.17. Covariance similarity and concentration

Note: This figure plots the universal cosine similarity and investor concentration of each firm-pair (investor concentration is associated with both firms in the pair) and the linear relationship (correlation) between universal ownership structure cosine similarity and investor concentration. Positive covariance implies adding both types of investors simultaneously increases ownership similarity and investor concentration. Negative covariation implies the added investors lead to an increase in investor concentration while decreasing ownership structure similarity.



(A) Universal ownership

(B) Common ownership

Figure A.18. Change in actual profit weights post merger

Note: This figure depicts the distribution of percentage changes in universal profit weights after the merger of BlackRock and BGI. The change is the difference between each actual profit weight, calculated based on 2010 Q1 ownership data, and the respective pre-merger profit weight using the baseline 2009 Q1 ownership data. Universal ownership includes changes in all firm-pair kappas, common ownership includes changes in kappas for firm-pairs within the same SIC 4 digit industry.

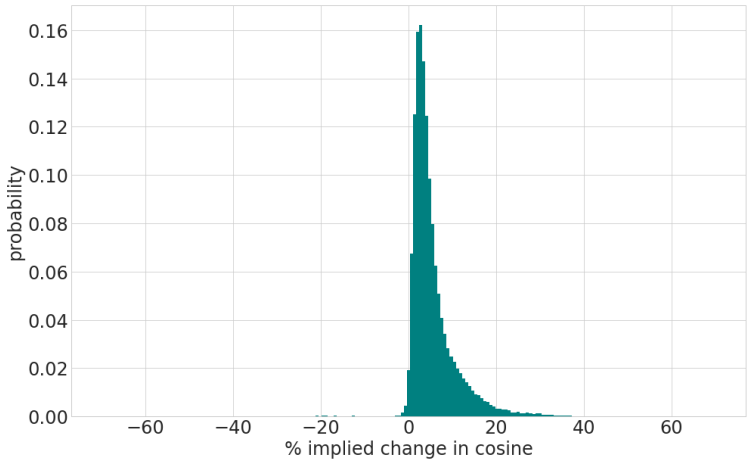


Figure A.19. Change in cosine similarity implied by merger

Note: This figure depicts the distribution of percentage changes in ownership structure similarity implied by the merger of BlackRock and BGI. The difference between each hypothetical "implied" firm-pair ownership similarity, calculated based on 2009 Q1 ownership data but aggregating the BlackRock and BGI holdings, and its respective pre-merger ownership similarity using the baseline 2009 Q1 ownership data is considered.

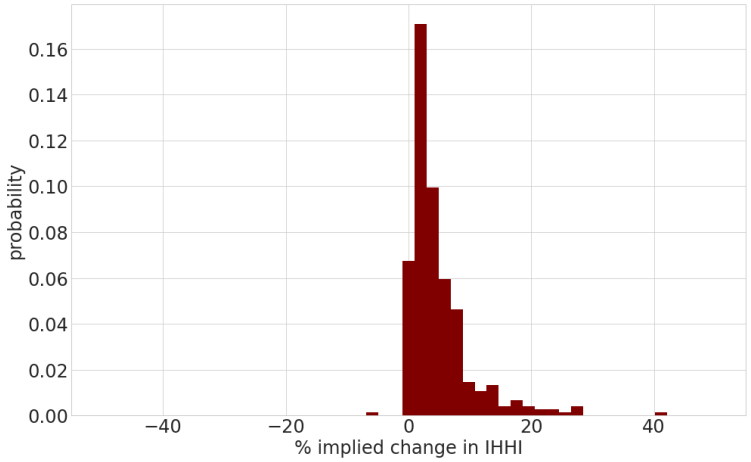


Figure A.20. Change in investor concentration implied by merger

Note: This figure depicts the distribution of percentage changes in investor concentration implied by the merger of BlackRock and BGI. The difference between the hypothetical "implied" investor concentration, calculated based on 2009 Q1 ownership data but aggregating the BlackRock and BGI holdings, and the respective pre-merger investor concentration using the baseline 2009 Q1 ownership data is considered.

	(1) κ_{univ}	(2) κ_{univ}	(3) κ_{univ}	(4) κ_{univ}	(5) κ_{univ}	(6) κ_{univ}
Insider Hold.	-0.249***		-0.338***	-0.225***		-0.312***
	(0.002)		(0.002)	(0.002)		(0.002)
13-D Holdings		-0.201***	-0.252***		-0.202***	-0.250***
		(0.002)	(0.002)		(0.001)	(0.001)
13-G Holdings		-0.129***	-0.207***		-0.121***	-0.194***
		(0.002)	(0.002)		(0.001)	(0.002)
Retail Share	0.590***	0.575***	0.553***	0.595***	0.581***	0.559***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Log(Market Cap)	0.058***	0.057***	0.057***	0.057***	0.056***	0.057***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Operating Margin	0.000**	0.001***	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Investor Indexing	1.081***	1.059***	0.962***	1.053***	1.026***	0.938***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Big 3 Holdings	0.050***	0.043***	0.087***	0.078***	0.073***	0.110***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
R-squared	0.622	0.622	0.623	0.796	0.796	0.797
R-squared Adj.	0.188	0.188	0.190	0.265	0.266	0.268
F-stat	493681	423307	374329	771036	661897	585180
p(F-stat)	0.000	0.000	0.000	0.000	0.000	0.000
Firm FE	✓	✓	✓			
Ordered Pair FE				✓	✓	✓
Quarter-Year FE	✓	✓	✓	✓	✓	✓
N	12799879	12799879	12799879	12799879	12799879	12799879

Table A.1. Common kappa regression on filer types

Note: This table presents results of a regression of common ownership profit weights (only considering firm-pairs that consist of industry competitors) on aggregate holdings of corporate insiders (filing Form 3, 4, 5 reports), blockholders (filing 13-D reports) and passive blockholders (filing 13-G reports), the retail share of ownership (defined as 1 minus captured ownership fraction), log market capitalization, operating margin, aggregate holdings by the three largest institutional asset managers (BlackRock, Vanguard and State Street) and a measure of investor indexing (as suggested by Backus et al 2021). 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.

	2009Q4 $\Delta\kappa_{act,com}$	2010Q1 $\Delta\kappa_{act,com}$	2011Q1 $\Delta\kappa_{act,com}$	2012Q1 $\Delta\kappa_{act,com}$
$\Delta\kappa_{implied}$	0.758***	0.309***	0.799***	0.875***
	(0.090)	(0.104)	(0.120)	(0.140)
R-squared	0.654	0.732	0.906	0.934
R-squared Adj.	0.051	0.006	0.032	0.028
F-stat	70	9	44	39
p(F-stat)	0.000	0.003	0.000	0.000
Firm FE	✓	✓	✓	✓
Industry FE	✓	✓	✓	✓
N	1302	1302	1302	1302

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

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 industry competitors. 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 $\Delta\kappa_{act,univ}$	2010Q1 $\Delta\kappa_{act,univ}$	2011Q1 $\Delta\kappa_{act,univ}$	2012Q1 $\Delta\kappa_{act,univ}$
treat	0.052***	0.048***	0.059***	0.071***
	(0.001)	(0.001)	(0.001)	(0.001)
R-squared	0.481	0.603	0.856	0.916
R-squared Adj.	0.049	0.035	0.037	0.041
F-stat	5067	3537	3774	4174
p(F-stat)	0.000	0.000	0.000	0.000
Firm FE	✓	✓	✓	✓
Industry FE	✓	✓	✓	✓
N	98560	98560	98560	98560

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.

	2009Q4 $\Delta\kappa_{act,com}$	2010Q1 $\Delta\kappa_{act,com}$	2011Q1 $\Delta\kappa_{act,com}$	2012Q1 $\Delta\kappa_{act,com}$
treat	0.048***	0.048***	0.100***	0.109***
	(0.008)	(0.009)	(0.010)	(0.011)
R-squared	0.653	0.774	0.968	1.007
R-squared Adj.	0.037	0.030	0.100	0.093
F-stat	34	28	98	90
p(F-stat)	0.000	0.000	0.000	0.000
Firm FE	✓	✓	✓	✓
Industry FE	✓	✓	✓	✓
N	868	868	868	868

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

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 industry 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.

Online Appendix to *Measuring Common Ownership: the Role of Blockholders and Insiders*

(Not for Publication)

A Data Appendix

A.1 Institutional investor ownership data

We parse the 13-F filings downloaded from the SEC EDGAR database with a modified version of the code used in Backus et al. (2021b) (BCS). The code modifications were necessary because we download a larger sample of filings and correct several errors related to filing attributes such as the filing entity's CIK and multi-row entries.

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 "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.

A.2 Blockholder ownership data

The 13-D and 13-G filings downloaded from EDGAR are very similar in structure and format, except for a paragraph where the filing entity states its intention leading to the purchase or sale of shares. The reporting date, the filing entity CIK, the subject entity CIK and the security's CUSIP and description are parsed with an adjusted version of the parsing code used in Schwartz-Ziv and Volkova (2021). All items can be identified from the html XML code of the first page of the filing.

We adjust for errors in the reported CUSIP identifier of securities by inferring the correct CUSIP from other filings with the same subject CIK and security description.

In addition to the identifying information on the filing and the subject entity and security, we parse the reported information on the intention of the purchaser and the resulting sole and shared voting and dispositive power of the filing entity. The intention may serve for further analysis of the effect of activist investing. In order to evaluate control right assumptions other than the most common proportionate control, sole and shared voting power information will be necessary.

A.3 Insider ownership data

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 3s are filed when an insider reports ownership for the first time and Form 5s are required if an insider neglects to report an acquisition or disposal of shares within 60 days of the transaction. Parsing the downloaded Forms 3, 4 and 5 from the EDGAR database presents more challenges than the other filing types, because the security owned is not specified exactly by a CUSIP. Instead, the insider reports the CIK of the company he is dealing in and a free text description of the security being acquired or disposed of. The number of shares owned after the transaction, the type of ownership (direct or indirect) and if indirectly owned, then the nature of such ownership (if by a trust or a family member) is reported. The most tedious construction step is the identification of the actual security being traded, as the free text description need not be accurate or informative. If the security is described as "Common

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

stock" or even more accurately as "Common stock Class A" it is relatively straightforward to link the CIK combined with the security description to a CUSIP identifier. 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. 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. Securities that are not held directly or in a trust by the reporting insider are also dropped, as the insider will not be able to make use of the residual rights associated with these securities. For now, we also drop all securities associated with companies having a dual class stock structure or non S&P 500 firms.

With the identifying information of filing date, CIK and common stock (class), we can match a CUSIP and PERMNO to most reported securities holdings. The match is based on a separately compiled, extensive database of all securities reported in COMPUSTAT and CRSP between 2000 and 2020. From the universe of all securities linked with a CRSP PERMNO or COMPUSTAT CUSIP, we drop such securities that are identified as non-equity either by their CUSIP (according to the rules of *Cusip Global Services*) or by the CRSP-provided identification. We then find more detailed information on the security type from 13-D, 13-G, 13-F filings, where in addition to the CUSIP of a security, a short free text description of the security is included. After cleaning these descriptions with a similar code as used for the Form 3,4,5 free text descriptions, we merge them with the time series data compiled from COMPUSTAT and CRSP. The resulting database contains the times when each security was traded on a public exchange and the type of the security, as far as it was ever reported in a SEC ownership report. We drop several securities that are identified as non-equity types and then merge this database with our Form 3,4,5 filings based on the associated CIK - security type - reporting date information.

A.4 Data aggregation

Joining the three ownership data sets from 13-F, 13-D & 13-G and Form 3, 4 & 5 filings creates many duplicates and several issues with the forward filling procedure of Form 3,4,5 data. The first inconsistency is cleaned out by searching all entries with the same reporting quarter, holding size, subject security and filing entity CIK. Entries that are clearly duplicated are removed. As filings dates may vary slightly leading to small variations in the holding size, we remove multiple entries where the holding size is not exactly the same but differs only by an insignificant fraction.

As the 13-F data provides a more regular indication of the existence of a security, we can check the accuracy of the forward fill method used in the Form 4 data construction with the data from 13-F. Specifically, we can check the last reporting date of a security in 13-F ownership data and if we forward filled an insider's holding of that security beyond this final reporting date, and do not have another insider holding report after that date, we know that the threshold cutoff of three years was too long. We can remove all insider holdings after the last reported 13-F holding of the security.