Firm Complexity and Conglomerates Expected Returns

Alexander Barinov

School of Business
University of California Riverside

E-mail: **abarinov@ucr.edu**
http://faculty.ucr.edu/∼abarinov

This version: November 2021

Abstract

The paper discovers that firm complexity is negatively priced in cross-section. High/low-complexity conglomerates have 35/20 bp per month more negative five-factor Fama and French (2015) alphas than single-segment firms, and this effect is stronger in subsamples with low institutional ownership, higher idiosyncratic volatility, and around earnings announcements. The complexity effect is robust to controlling for a long list of pre-existing anomalies and seems to be generated by the interaction of higher disagreement about conglomerates and short-sale constraints.

JEL Classification: G12, G14, G34

Keywords: conglomerates, anomalies, mispricing, limits to arbitrage, diversification discount

1 Introduction

Organizational complexity was recently used in several asset-pricing studies. Cohen and Lou (2012) show that conglomerates take longer to process industry-level shocks and respond to those one month later than single-segment firms, which creates predictability of conglomerate returns at short horizons using returns to pseudo-conglomerates formed from single-segment firms. Barinov et al. (2020) use firm complexity as a limits to arbitrage variable and show that conglomerates have stronger post-earnings-announcement drift.

Cohen and Lou (2012) argue that investors have limited resources they can dedicate to information processing and that incorporating industry-wide shocks into the price of a conglomerate takes more effort. Compared to single-segment firms, incorporating the same industry-wide shock into the price of a conglomerate calls for consideration of potential spillover effects of the shock into the other industries the conglomerate operates in, requires determining the exact (time-varying) weight of the affected segment in the conglomerate, etc. Thus, Cohen and Lou predict and find that it takes longer for prices of conglomerates to reflect industry-wide shocks.

Barinov et al. (2020) show that investors face similar problem when receiving firm-specific information about the conglomerate. In the absence of detailed information of how, for example, variable and fixed (overhead) costs really split between the conglomerate divisions, investors have to spend more time and effort deciphering what an earnings surprise of a conglomerate means for each of the conglomerate division. Hence, Barinov et al. predict and find that earnings information is incorporated slower into prices of conglomerates and conglomerates have stronger post-earnings-announcement drift.

Barinov et al. (2020) also show that costlier processing of information about conglomerates leads to less information about conglomerates being produced, as analysts and informed investors tend to abandon conglomerates. The first part is consistent with earlier evidence in Gibson et al. (2001) that breaking up a conglomerate improves analyst coverage and diminishes analyst disagreement.

This paper looks at long-run impact of complexity on expected returns. Since, as Cohen and Lou (2012) and Barinov et al. (2020) show, information about conglomerates is harder to produce and less of it is therefore produced, I hypothesize that complexity creates disagreement: if investors start with diffuse priors, less information means that their posteriors will also be relatively diffuse. This higher disagreement about conglomerates, in combination with short-sale constraints, creates overpricing and subsequent underperformance, as in Miller (1977). If short-sale constraints are binding, pessimistic investors are kept out of the market, and the market price represents the average valuation of optimists, which increases with disagreement (optimists become more optimistic, pessimists become more pessimistic, but (some) pessimists do not trade).

I define firm complexity in two alternative ways: as a dummy variable (Conglo) that separates single-segment firms from conglomerates (firms with business segments in industries with different two-digit SIC codes) and as a measure of sales concentration (Comp) among the business segments. I find that conglomerates, and in particular high-complexity conglomerates, have lower institutional ownership, smaller analyst following, larger analyst forecast errors, and higher analyst disagreement than peer firms. My findings extend the results in Gilson et al. (2001), who focus on a small sample of spin-offs and carve-outs and find that once a conglomerate breaks up, its analyst coverage and the quality of analyst forecasts increases to what is normal for single-segment firms within one or two years.

In portfolio sorts on complexity, I find that irrespective of which complexity measure I use, high-complexity conglomerates have negative five-factor Fama and French (2015) model alphas (FF5 alphas), which are 35 bp per month lower than positive alphas of single-segment firms. Low complexity conglomerates also have FF5 alphas that are by 20 bp per month smaller than those of single-segment firms. Due to high persistence of the conglomerate status, the first alpha differential persists for at least five years, while the second lasts for two years. I also find that the complexity effect is robust to controlling for the long-run underperformance of bidders after the merger is completed (Agrawal, Jaffe, and Mandelker, 1992).

Consistent with the complexity effect being mispricing generated by the interaction of disagreement and short-sale constraints, I find that the complexity effect is significantly stronger for low institutional ownership firms (low supply of shares for shorting), firms with high probability of having high shorting fees, and firms with high idiosyncratic volatility (high limits to arbitrage). The complexity effect is also disproportionately concentrated around earnings announcements, when the investors are more likely to correct their valuation mistakes upon seeing the earnings: around 30% of the complexity effect is realized during less than 5% of trading days of the year that surround earnings announcements.

The complexity effect turns out robust to changes in research design and controlling for other potentially related effects. In particular, the complexity effect survives in cross-sectional regressions even after I control for other anomalies often attributed to the interaction of disagreement and short-sale constraints, such as the analyst disagreement effect of Diether, Malloy, and Scherbina (2002), the idiosyncratic volatility effect of Ang, Hodrick, Xing, and Zhang (2006), and the short interest effect of Asquith, Pathak, and Ritter (2005). Since firms with high idiosyncratic volatility/analyst disagreement/short interest are, on average, quite small, while complex firms are normally relatively large, there is little overlap between the latter anomalies and the complexity effect even though the economic mechanism at work is likely to be the same.

The last observation is important: one can trade on the complexity effect without having to short small, illiquid, distressed, and extremely volatile companies the other disagreement-based strategies depend on. The complexity effect also seems to persist for several years, in contrast to the idiosyncratic volatility effect, the short interest effect, and other related anomalies, which last for at most a year, and thus trading on them entails high turnover and high trading costs.

While the focus of the paper is the link between firm complexity and future returns, as well as the trading strategy this link implies, the existence of the complexity effect can affect our interpretation of the diversification discount of Berger and Ofek (1995) and Lang and Stulz (1994). The negative alpha of (complex) conglomerates suggest that conglomerates are overpriced and are traded at too high valuation multiples. One way to interpret this evidence is as supportive of studies suggesting that the diversification discount is a measurement issue and if conglomerates are matched to single-segment peers in a different way, conglomerates no longer seem to trade at lower multiples (see, e.g., Villalonga, 2004a, 2004b, Hoberg and Philips, 2014, and Hund, Monk, and Tice, 2021).

Another way to reconcile the complexity effect in expected returns and the diversification discount in valuation multiples is to assume that even if conglomerates trade at lower multiples, those multiples are still not low enough and conglomerates are still overpriced. Conglomerates' complexity leads to increased disagreement between investors and, coupled with short-sale constraints, makes the market to underestimate the cash flow losses from inefficiencies within a conglomerate. As the information about these inefficiencies comes out, the returns to conglomerates suffer and the diversification discount slowly builds up. This new interpretation of the diversification discount is consistent with my finding that newer conglomerates underperform more.

While it is beyond the scope of the paper to fully judge whether the relatively low multiples of conglomerates arise because of measurement issues or accumulate gradually as a result of the continued underperformance I document, my paper supplies two pieces of evidence consistent with the diversification discount arising at least partly due to ´´slow bleeding”, i.e., prolonged underperformance of conglomerates.

First, I show that in the years preceding conglomerate formation future conglomerates have higher valuation multiples than their peers. The valuation multiples sharply decline in the first five years after the conglomerate is formed and keep declining at least for another decade. Second and most importantly, I find that conglomerate complexity is positively related to the conglomerate's valuation multiples in the first decade of its life (suggesting that high complexity conglomerates are initially more overpriced), but the sign of the relation flips to significantly negative in the second decade of the conglomerate's life (consistent with stronger underperformance of complex conglomerates, which is part of the complexity effect in returns).

In Robustness Appendix, I address several studies suggesting that conglomerates have lower cost of capital and potentially low risk. Hann et al. (2013) find that conglomerates have lower implied cost of capital and attribute this fact to coinsurance between business segments. Consistent with that, Hann et al. find that conglomerates' cost of capital is smaller if the segments' cash flows are less correlated, and the effect of coinsurance is stronger if conglomerates are also financially constrained. In their paper, Hann et al. look at weighted average cost of capital, which includes cost of debt in addition to cost of equity, and they define cost of equity as implied cost of capital that sets the discounted value of analyst earnings forecasts (adjusted for payments to debtholders) equal to the stock price.

Looking at average future returns rather than implied cost of capital, I do not find that the strength of the complexity effect depends on cash flow correlation between segments, and the relation between the complexity effect and financial constraints is mixed: in fact, controlling for the five Fama-French factors, conglomerates load more positively rather than negatively on high-minus-low financial constraints factors. I also find no relation between cash flow correlation between segments and my measures of firm complexity, supporting my conclusion that the complexity effect is distinct from the cost of capital effect Hann et al. (2013) document.

Several papers consider performance of conglomerates during downturns, looking primarily at cash flow effects (Gopalan and Xie, 2011, Matvos and Seru, 2014) or the effect of the Great Recession (Kuppuswamy and Villalonga, 2010). In Section 6.1, I consider the effect of several business cycle variables that are popular in the asset pricing literature (default spread, VIX, TED spread, etc.) on the risk premium of conglomerates and on their abnormal returns during recessions and find no evidence that in the past 40 years conglomerates performed better than single-segment firms during recessions.

2 Data

The data on conglomerates are collected from Compustat segments file. Conglomerates are defined as firms with segments in two or more industries with different two-digit SIC codes; all other firms are referred to as single-segment. Conglomerates constitute about 30% of the sample, though their share is different in different years, varying from about 25% in the more recent sample to roughly 45% in the earlier years.

I use several measures of complexity, the simplest being the conglomerate dummy, Conglo (1 for conglomerates, 0 for single-segment firms). Another measure is Comp, or 1-HHI, which looks at concentration of sales in the conglomerate segments using Herfindahl index (HHI) and then deducts this index from 1. The logic of the Comp measure is that a two-segment conglomerate with 95% of sales concentrated in one segment is very close to a single-segment firm (its 1-HHI measure equals 0.095, and any single-segment firm has 1-HHI equal to 0 by definition), while a two-segment conglomerate with 50-50 split of sales between two industries is significantly more complex (its 1-HHI measure equals 0.5).[1]

The rest of the data come from standard data sources: stock returns and prices are from CRSP, firm-level financial information (used to compute market-to-book, profitability, etc.) is from Compustat annual file, analyst following and analyst forecasts are from IBES, institutional ownership is from Thomson 13F database.

All tests in the paper use the maximum available sample: for example, if a regression uses institutional ownership, the sample includes only firms with non-missing institutional ownership, but if the next test requires only market-to-book and Comp to be non-missing, then firms with missing institutional ownership will be included in this test. The sample in the paper is from January 1978 to December 2016 (based on availability of Compustat segments data). Detailed definitions of other variables used in the paper are in the Data Appendix.

3 Firm Complexity and Firm Characteristics

3.1 Descriptive Statistics

Table 1 presents descriptive statistics for conglomerates (defined as firms that report business segments with different two-digit SIC codes) and single-segment firms (all other firms on Compustat segment files). The descriptive statistics in Panel A and B are size-adjusted, since conglomerates are almost three times larger, on average, than single-segment firms ($3.34 billion vs. $1.26 billion).

Panel A looks at standard asset-pricing controls. On size-adjusted basis, conglomerates have significantly smaller market-to-book and investment-to-assets ratios. They also seem

[1] In robustness tests, I also use NSeg, the number of segments with different two-digit SIC codes (conglomerates with more segments in different industries are assumed to be more complex) and RSZ, coefficient of variation (the ratio of standard deviation to average) of segments' imputed market-to-book, as suggested in Rajan, Servaes, and Zingales (2000). By definition, single-segment firms have the RSZ measure equal to zero.

less profitable and have lower past returns, though economically the latter two differences are rather small. Similar, though less significant differences exist between conglomerates below and above median in terms of their complexity - more complex conglomerates have lower market-to-book and somewhat lower investment-to-assets and profitability.

Panel A underscores the importance of thoroughly controlling for known priced factors/characteristics when comparing expected returns to conglomerates and single-segment firms, in particular, for the new Fama and French (2015) factors, CMA (investment) and RMW (profitability), something that has not been done by prior studies of conglomerates expected returns, such as Lamont and Polk (2001) and Hannet al. (2013), which only controlled for size and market-to-book.

Panel B takes the first look at the information environment of conglomerates vs. single-segment firms. All variables except for idiosyncratic volatility suggest that conglomerates face more uncertainty than single-segment firms of comparable size. Controlling for size, conglomerates have weaker analyst coverage and are followed by less analysts who would be specialists in their main industry.[2] The analysts covering conglomerates make larger earnings forecast errors and disagree more in their forecasts, and earnings of conglomerates are more volatile. The same is true if we compare low vs. high complexity conglomerates: consistent with my hypothesis, more complex conglomerates have more uncertain information environment.

3.2 Multivariate Analysis

Gilson et al. (2001) look at non-random sample of conglomerates that chose to break up and show that the break-up increases analyst coverage, because after the break-up analysts get access to detailed financials of each division and also can choose which division to

[2] A specialist is an analyst who follows at least five firms with the same two-digit SIC code as the firm in question. For a conglomerate, specialists are defined using the two-digit SIC code of the biggest segment.

follow based on their expertise. Coupled with the hypotheses in Cohen and Lou (2012) and Barinov et al. (2020) that information about conglomerates is harder and costlier to process and the processing takes longer, I predict that in the full sample that includes all conglomerates, conglomerates have less analyst following and the analysts make larger mistakes and disagree more about conglomerates.

Table 2 looks at the information environment of conglomerates vs. single-segment firms using panel regressions as suggested in Peterson (2009). The dependent variables are analyst following (number of analysts following the firm and number of analysts specializing in the industry of the firm, institutional ownership, earnings forecast error, and analyst disagreement (dispersion of analyst forecast). The analysis in Panel A follows a similar analysis in Barinov et al. (2020) and uses the standard controls suggest by the literature, for example, in Gompers and Metrick (2001), Barth, Kasznik, and McNichols (2001), Bhushan (1989). The main regressor is the Comp measure based on sales concentration, which is the complexity measure used to divide conglomerates into low- and high-complexity ones in Table 1.

Panel A looks at the full sample and confirms the results in Table 1 and in Barinov et al. (2020): holding other variables constant, conglomerates have low institutional ownership, smaller analyst following, less precise earnings forecasts, and more disagreement among analysts, consistent with my hypothesis that conglomerates face higher uncertainty and less transparent information environment.

The Comp measure, by definition, equals zero for all single-segment firms, which constitute more than 70% of my sample. Thus, the Comp measure has a large mass at zero, which may be driving the results. In other words, the results in Panel A can mean that only conglomerate status, and not the degree of complexity, matters: conglomerates can differ from single-segment firms in terms of their information environment, but not from each other.

To test this hypothesis, Panel B extends the Barinov et al. (2020) analysis by looking only at conglomerates and repeats the regressions in Panel A for conglomerates only. Panel B strongly supports the idea that the degree of complexity matters and low-complexity conglomerates face a better information environment than high-complexity conglomerates. The slopes on the Comp variable in the conglomerate-only sample are all significant and are generally even bigger than in the full sample.[3]

4 The Complexity Effect

4.1 Portfolio Sorts

The evidence in the previous section shows that less information about conglomerates is produced due to increased cost of information production brought about by firm complexity. In the relative absence of information, investors' beliefs about the conglomerate value will be more diffuse (as compared to a similar single-segment firm) and there will be more disagreement between investors as to the conglomerate value. According to Miller (1997), disagreement between investors leads to overpricing in the presence of short-sale constraints, since with higher disagreement optimists are more optimistic, pessimists are more pessimistic, but short-sale constraints preclude (some of) the pessimists from trading.

Table 3 takes the first look at the complexity effect by splitting the sample into three portfolios. Single-segment firms, which constitute more than 70% of my sample, are pooled into one portfolio, labeled Zero. Conglomerates are then split into two portfolios: Low/High portfolio includes all conglomerates with firm complexity, Comp, below/above median.

In the first row of Panel A1, the five-factor Fama-French (2015) model (FF5) discovers

[3] One can also notice that the conglomerate-only sample is similar to the full sample, since the slopes on all control variables have the same sign and similar magnitude/significance in Panels A and B.

that single-segment firms earn a value-weighted alpha of 19.4 bp per month, which is significantly different from both the negative alpha of the more complex conglomerates (-16.2 bp per month, t-statistic -2.7) and the small negative alpha of low-complexity conglomerates. The second row of Panel A1 adds the momentum factor to the FF5 model and finds little difference in results, other than reduced significance of the alpha differential between low and high-complexity conglomerates. I conclude that all conglomerates have significantly lower expected returns than single-segment firms, and the difference in expected returns is related to conglomerate complexity: more complex conglomerates have lower expected returns.[4]

While the argument in the paper centers around mispricing of (complex) conglomerates, Panel A1 finds that the complexity effect comes equally from the long and short sides, i.e., from positive alphas of single-segment firms as much as from negative alphas of (complex) conglomerates. I believe that positive alphas of single-segment firms are a mechanical implication of negative alphas of conglomerates, as by definition, alphas have to sum to zero. Conglomerates are large firms, and their total capitalization exceeds that of the single-segment group, making it certain that if conglomerates are mispriced, single-segment firms will also have a significant alpha of the opposite sign.[5]

Since previous studies that did not use the FF5 model, such as Lamont and Polk (2001) and Hann, Ogneva, and Ozbas (2013), did not find significant difference in expected returns between conglomerates and single-segment firms, the next two rows in Panel A1 attempt to find out which of the two new factors in the FF5 model, CMA (investment) or RMW

[4] In Section 1 of online Robustness Appendix, I repeat Panel A1 using two alternative complexity measures, NSeg (number of segments with different two-digit SIC codes) and RSZ (variability of segment-level market-to-book) and find that the results in Panel A1 are robust to using those measures.

[5] Several papers on disagreement effects (such as Diether et al., 2002) also find positive alphas of low-disagreement firms; others (such as Ang et al., 2006) only find significant negative alphas of high disagreement firms, but the reason for the latter is that in these studies the high-disagreement group consists of very small firms and takes 2-5% of total capitalization of the market.

(profitability) create the difference in the alphas. The third row adds CMA to the older three-factor Fama-French (1993) model (FF3) and discovers little difference in the alphas between the three groups of firms. In the fourth row though, when RMW is added to the FF3 model, the alphas are similar to the FF5 alphas, which suggests that the main driver of the complexity effect in the FF5 model is the profitability factor, RMW.

The importance of RMW for measuring the abnormal performance of conglomerates is also confirmed in Panel A2, which tabulates the FF5 betas of each of the three groups of firms. The difference between MKT, SMB, HML betas of single-segment firms and high/low complexity conglomerates is minimal. The real difference is in the CMA and especially RMW betas, which are significantly positive for conglomerates and significantly negative for single-segment firms. Conglomerates (especially complex ones) are profitable, low-investment firms, which are supposed to earn superior returns (e.g., positive FF3 alphas), but they do not.

The spread in RMW betas between single-segment firms and conglomerates is visibly wider than the same spread in CMA betas (0.48 vs. 0.3). What is even more important is that the three-factor alpha of CMA is 23 bp per month in my sample period (1978-2016), and the three-factor alpha of RMW is 43 bp per month. Thus, it is not surprising that Panel A1 shows that controlling for RMW has a major role in discovering the complexity effect, while controlling for CMA also contributes, but slightly.[6]

In 1997, Statement of Financial Accounting Standards No. 131 (SFAS 131), Disclosures about Segments of an Enterprise and Related Information, was issued, and subsequent research (e.g., Herrmann and Thomas, 2000, and Berger and Hann, 2003) found that

[6]While Table 1 finds that size-adjusted profitability of conglomerates is lower than that of single-segment firms, this comparison only controls for size. FF5 model betas in Table 3 look at conglomerates profitability controlling for market-to-book and investment (HML and CMA) as well. Conglomerates are value firms, and while value firms usually have low profitability (see, e.g., Fama and French, 1995), Table 1 finds that conglomerates are as profitable as an average Compustat firm of their size.

SFAS 131 improved segment disclosure quality and made it easier for analysts to forecast conglomerates earnings. It is interesting therefore to evaluate whether the mispricing that creates the complexity effect became smaller after SFAS 131 was implemented.

The last two rows of Panel A1 add the SFAS 131 dummy (1 in 1999-2016, 0 in 1978-1998)[7] to the FF5 model. The intercept is then the pre-SFAS alpha ($\alpha_{pre-SFAS}$) and the slope on the dummy ($\Delta\alpha_{post-SFAS}$) measures the change in the alpha brought about by SFAS 131.

I find that the impact of SFAS 131 on the complexity effect is minimal: FF5 alpha shifts by just a few bps per month after 1999, and the pre-SFAS 131 alphas is very close to the full-sample alphas and its significance is preserved. While all slopes on the SFAS 131 dummy are insignificant, positive signs of point estimates are more common, suggesting that the complexity effect is, if anything, stronger post-SFAS 131. [8]

Lastly, in Panel B of Table 3, I look at tradability of the complexity effect by removing from the sample all firms with below-median liquidity and reporting the FF5 alphas across the Comp measure groups in these subsamples. All tables in the paper already have the price filter (stocks priced below $5 at the portfolio formation date are omitted). In Panel B, I omit firms with below-median effective bid-ask spread, or below-median price impact, or below-median market cap, as indicated by the name of the row.

The main point of the subsample analysis in Panel B is to remove firms with truly high trading costs. The liquidity measures may not be perfect and one can disagree that those measures being below median is enough to make a firm liquid, but as long as the subsamples keep trading costs in check, the goal is achieved.

[7] In analyzing the effect of SFAS 131, I define the post-SFAS 131 period as 1999-2016, leaving the transition year of 1998 in the pre-SFAS sample due to initial uncertainty how to implement the new rule. The results are robust to re-defining post-SFAS 131 period as 1998-2016 or 2000-2016.

[8] Another way to interpret the results with the SFAS 131 dummy is to conclude from them that the complexity effect does not deteriorate with time and is just as strong in the second half of the sample as in the first half.

The complexity effect in Panel B survives the exclusion of illiquid firms from the sample. The first line of Panel A1 estimated it at 35 bp per month; Panel D pegs it between 22 and 41 bp per month, with five out of six estimates being greater than 30 bp, and all six being statistically significant. The difference in alphas between low and high complexity conglomerates declines relatively more and loses significance, but the negative alpha of high complexity conglomerates invariably stays significant, in contrast to the negative alpha of low-complexity conglomerates.

An interesting by-product of Table 3 left for future research is the relation between the profitability effect and complexity effect. As Table 3 shows, the complexity effect is stronger once I control for RMW, the profitability factor. According to RMW betas, conglomerates seem to be profitable firms that have relatively low returns despite being profitable. Hence, if I drop conglomerates from the sample, the profitability effect will increase. In untabulated findings, I find that the Carhart alpha differential between top and bottom profitability quintile is indeed 25 bp per month higher in the single-segment firm subsample than among conglomerates, and this difference cannot be attributed to differences in size between single-segment firms and conglomerates.

4.2 Robustness in Cross-Sectional Regressions

Table 4 performs several robustness checks on the results in Table 3. First, in its first four columns, Table 4 verifies that the complexity effect is visible in cross-sectional regression and robust to using two alternative measures of complexity - number of segments with different two-digit SIC codes, NSeg, and variability of segment-level market-to-book, the RSZ measure. The regressions use DGTW-adjusted returns (that is, returns adjusted for size, market-to-book, and momentum as in Daniels, Grinblatt, Titman, and Wermers, 1997) on the left-hand side and a long list of standard controls, including investment and profitability, on the right-hand side.

In the first four columns of Table 4, all four complexity variables (the two described above plus Comp and the conglomerate dummy, Conglo) are negative and strongly statistically significant, with t-statistics above 3, thus clearing the higher threshold Harvey, Liu, and Zhu (2016) recommend for new anomalies. Other control variables also have expected signs: size and momentum are not significant due to the DGTW-adjustment on the left-hand side, previous month return (reversal) and investment are negative and significant, and profitability is positive and significant, market beta is positive, but insignificant.

The first four columns of Table 4, as well as all other tests in the paper, use the sample of all firms covered by Compustat segment files. Single-segment firms, however, do not have to file segment-level data, and some conglomerates with small segments (segment sales less than 5% of total sales) do not have to report segment-level data as well. Many of them do report voluntarily (and then they show up on Compustat segment files), but some do not (and show up only on the standard Compustat annual files). Thus, firms that are on Compustat, but not on Compustat segment files, are either single-segment firms or multi-segment firms with small segments.

In my analysis, I do not exclude small segments if they are reported on Compustat segment files. Firms with small segments are coded as having Conglo dummy equal to one, and I compute the other complexity measures for them (their RSZ and Comp measures come out to be very close to RSZ and Comp of single-segment firms, which are zero by definition).

The next four columns (five to eight) of Table 4 test the robustness of the complexity effect to labeling all Compustat firms that are not on Compustat segment files, as single-segment firms. These firms are assumed to have Conglo equal to zero, NSeg equal to one, and Comp and RSZ equal to zero. This is not exactly correct, because some of these firms may have more than one business segment which they do not report, but one can argue that such small segments do not materially increase firm complexity.

Columns five to eight of Table 4 show that all complexity variables remain negative and highly significant in the larger sample, though their slopes decline by about 20%, which is not surprising given that classification errors are unavoidable for Compustat firms that are not on Compustat segment files (e.g., a small number of these firms can have multiple small segments, each one of which does not produce sales greater than 5% of total sales, but taken together these segments can take a large fraction of the firm and materially increase its complexity - such firms will be misclassified by columns five to eight of Table 4 as zero-complexity firms).

Another potential problem with the baseline analysis in the first four columns in Table 4 is that NSeg has a large mass at 1, and Comp and RSZ have a large mass at 0 (Comp and RSZ are 0 for all single-segment firms, which constitute close to 80% of my sample). Thus, it is not clear from the regressions in columns one to four whether the negative link between, e.g., Comp and future returns is driven by the mass at zero (thus saying that conglomerates have lower future returns than single-segment firms, but all conglomerates are essentially the same irrespective of their complexity) or whether this negative link presents a negative association between firm complexity and expected returns, which also exists among conglomerates.

Table 3 deals with this problem by putting all single-segment firms into one portfolio (Zero complexity), but splitting conglomerates into the ones with complexity below median (Low portfolio) and the ones with complexity above median (High portfolio). Then the rightmost column of Panel A1in Table 3 tests the difference in the alpha of High and Low portfolios and finds that the alphas of High portfolios are indeed more negative, but the difference is marginally significant at 10%.

Column nine of Table 4 applies a similar solution in the multiple regression setting by creating two dummy variables for low and high complexity conglomerates. LoComp/HiComp equals one if Comp is below/above median Comp for conglomerates in a given year and zero otherwise. Column nine regresses future returns on LoComp and HiComp and standard controls. The slopes on LoComp/HiComp measure the difference in (abnormal) future returns between low/high complexity conglomerates and single-segment firms.[9]

I observe that the difference in future returns between high complexity conglomerates and single-segment firms is strongly significant and twice larger than the similar difference between low complexity conglomerates and single-segment firms, with the latter difference being at most marginally significant. This evidence is consistent with my hypothesis that complexity matters and high-complexity conglomerates are more likely to be overpriced. This evidence is also consistent with portfolio sorts in Table 3, which finds (in Z-H and Z-L columns in Panel A1) 75% smaller difference in the alpha differential between high/low complexity conglomerates and single-segment firms, but finds that the alpha differential between low-complexity conglomerates and single-segment firms is statistically significant.

Column ten in Table 4 tests if the difference in the future return differentials between low/high complexity conglomerates and single-segment firms is statistically significant by using Conglo dummy instead of LoComp. In this setup, the Conglo dummy (1 for all conglomerates, 0 otherwise) captures the difference in future returns between low-complexity conglomerates and single-segment firms, and the high complexity dummy captures the extra (negative) abnormal return to high-complexity conglomerates. The slope on HiComp dummy roughly double the slope on Conglo dummy. The lack of power to reject the

[9] In Section 1 of online Robustness Appendix, I show that the results in columns nine to eleven of Table 4 are robust to forming low/high complexity dummies based on number of segments (NSeg) or the RSZ measure).

equality of future returns of high and low complexity conglomerates was also an issue in portfolio sorts (Table 3, column L-H in Panel A1), where this difference was marginally significant.

The rightmost column of Table 4 (column eleven) tries a different approach of restricting the sample only to conglomerates. The rightmost column finds marginally significant difference between future returns to low and high complexity conglomerates HiComp dummy.

Overall, columns nine to eleven in Table 4 confirm the finding in Table 3 that the degree of complexity matters, even though in some cases we see the difference in expected returns between high and low complexity conglomerates, but do not have enough power to reject the null.

4.3 Persistence of Complexity Effect

Table 5 looks at how long the complexity effect remains visible. On the one hand, a longer life of mispricing means lower trading costs and higher potential profits for someone trading on it. On the other hand, one would expect the mispricing to be eventually resolved: if a pattern in expected returns persists for a decade without weakening, one would be more inclined to attribute it to difference in risk.

Panel A of Table 5 repeats the cross-sectional regressions from Table 4, lagging the Comp measure of complexity by the number of years indicated in the column heading. All control variables are always lagged by the same number of periods as in Table 4. The results in Panel A stay the same if I replace the Comp measure by the Conglo dummy, NSeg, or the RSZ measure.

I find in Panel A that the complexity effect stays at the same level for four years and then weakens in the fifth year. This evidence suggest that trading on the complexity effect would require minimum turnover and the total strength of the complexity effect is much larger than what the one-year tests in Tables 3 and 4 suggest.

The result that the complexity effect seems to never completely disappear may be troubling for the mispricing story (though LaPorta et al., 1997, for example, show that the value effect does not deteriorate for at least five years, but favor its mispricing explanation), but this result is likely to be caused by the fact that conglomerates rarely break up, and once a firm has positive Comp, it will have positive Comp for at least several years.

To distinguish between the effect of the conglomerate status and the potentially less persistent effect of complexity per se, in Panel B of Table 5 I look at the three FF5 alpha differentials from the left part of Table 3. The top two rows in Panel B of Table 5 report the alpha differential between high/low complexity conglomerates and single-segment firms (Z-H/Z-L) in five years after portfolio formation. These differentials are likely to be mostly ´´single-segment vs. conglomerate" and, similar to Panel A, they stay roughly the same (Z-L differential even increases) for at least five years.

The bottom row of Panel B looks at the FF5 alpha differential between low and high complexity conglomerates (L-H). This differential is due to differences in complexity and it has a much shorter life, only two years. Hence, the pure complexity effect, net of its long-term part related to the organizational form, is a medium-term effect - it brings roughly 15 bp per month for two years, for the total of roughly 3.6%.

Panel C of Table 5 restricts the sample to conglomerates only and regresses future returns on the HiComp dummy and controls, similar to Table 5. As in Panel A of Table 5, the controls are always lagged by the same number of periods, and the HiComp dummy is lagged by the number of years indicated in the column heading. Panel C finds that the pure complexity effect, net of the impact of the organizational form, is still visible after four years (though it loses significance in the third year) and then seems to disappear.

Panel D of Table 5 attempts to look at persistence of the complexity effect by considering conglomerate age, i.e., the number of consecutive years a firm has been reporting multiple two-digit SIC segments. The median conglomerate in my sample is five years old; 25% of conglomerates are older than 10 years and 10% of all conglomerates are older than 18 years.[10] In column one of Panel D, I run, in the conglomerate-only subsample, a regression of conglomerate returns on the log of conglomerate age and the standard asset-pricing controls. Table 3 pegs the complexity effect at roughly 35 bp per month; the slope on the conglomerate age variable suggests that this effect will be completely gone for a conglomerate roughly 20 years old, and will be reduced to insignificance even earlier than that.

The next three columns of Panel D add the complexity variables (one per column, either Comp=1-HHI, or NSeg, or RSZ) and their interaction with the log of conglomerate age. Similar to the first column, where the slope on conglomerate age is roughly three times smaller than the complexity effects, in columns two to four the slope on the interaction term is also three times smaller than the slope on the complexity variable itself, suggesting that complexity will not matter for a 20-year-old conglomerate. Also, the slope on the complexity variable in the presence of the interaction term reflects the complexity effect of a new, one-year-old conglomerate, and the comparison of this slope with a similar slope in Table 4 suggest that the complexity effect triples for new conglomerates - the result that will be further discussed in the next subsection.

Overall, Table 5 suggests that the complexity effect has two parts: an extremely long-lived ´´single-segment vs. conglomerates" part, which lasts for at least five years due to the strong persistence of the conglomerate status, and the medium-term ´´low vs. high

[10]Some conglomerates in my sample stop reporting multiple segments for several years and then start reporting them again; in this case, their age as a conglomerate restarts at one after the break. Since some short breaks (e.g., two years) are probably a reporting issue rather than the sign of breaking up a conglomerate and then forming a new one, the age statistics I report are likely to be underestimates.

complexity" part, which lasts for two or four years in the conglomerate-only sample.

4.4 New Conglomerates and Post-M&A Effects

One way a firm can become a conglomerate is through mergers and acquisitions. As known since at least Agrawal, Jaffe, and Mandelker (1992), bidders have negative abnormal returns in up to three post-merger years, which can be partially responsible for the overall negative abnormal returns to all conglomerates. To test what part of the complexity effect can be attributed to new conglomerates that arise out of mergers, I create three ´´new conglomerate" dummies (NewConglo1/NewConglo2/NewConglo3), which equal one in one/two/three years after the company expanded into another industry with a different two-digit SIC code.

In untabulated results, I find that roughly 8% of conglomerates in my sample (and roughly 2.5% of all firms on Compustat segment files) are new conglomerates according to NewConglo1 dummy (one if the firm expanded into another industry in the past year). Not all those cases are mergers: in fact, between one-half and two-thirds of them can be tracked to M&A activity using SDC data, while the rest of the new conglomerates seem to grow ´´from within" (e.g., a manufacturing company can develop a financing arm or a retail segment and start reporting it as a new line of business).

My explanation of the complexity effect suggests an alternative interpretation of post-merger underperformance: investors are more confused and disagree more about new conglomerates, and thus new conglomerates underperform more. As time passes, investors collect more information on the conglomerate and the disagreement subsides; this is also consistent with the result in Panel D of Table 5 that older conglomerates (15-20 years old) do not underperform.

The left part of Table 6 adds the ´´new conglomerate" dummies to the regression of future returns on the conglomerate dummy and standard controls. As the previous paragraph explains, under my explanatio of the complexity effect I also expect the slopes on the NewConglo dummies to be negative.

The main result in Table 6 is that the conglomerate dummy remains negative and significant: compared to Table 3, its slope declines by roughly 20%, which means that only 20% of the complexity effect I document in the paper is due to the new conglomerates and presumably to the bidders' post-merger underperformance. I do find the bidders' underperformance in Table 6: all ´´new conglomerate" dummies come out negative and significant, and their slopes are 2 to 3.5 times larger than the slope on the Conglo dummy, indicating that ´´new conglomerates" do significantly worse than older conglomerates. The roughly three-times stronger complexity effect for new conglomerates in Table 6 is similar with what Panel D of Table 5 finds using the conglomerate age variable.

In the right part of Table 6, I also consider increases in the number of segments: even a company that is already a conglomerate can be involved in M&A and add a segment by buying a company from a different industry (thereby adding to its complexity and consequent disagreement). Such increases in firm complexity and consequences of such mergers would not be considered in the left part of Table 6: a conglomerate that adds another segment would not be a ´´new conglomerate" and would have the Conglo dummy equal to 1 and the NewConglo dummies equal to 0.

The right part of Table 6 creates the dummy variables (SegInc1/SegInc2/SegInc3) for conglomerates that have experienced an increase in the number of segments with different two-digit SIC codes in the past one/two/three years. SegInc dummies equal one for all new conglomerates, but they also equal one for conglomerates that added another segment. The fraction of conglomerates that added another segment in the past year is considerable and constitutes around 5% of all conglomerates (roughly 1.5% of all firms on Compustat segment files), which is in addition to 8% of conglomerates and 2.5% of all Compustat segment firms that turned from single-segment firms into conglomerates.

The right part of Table 6 adds the SegInc dummies to the regressions of future returns on the number of segments (NSeg) and standard controls. The NSeg variable captures the effect of having another segment on expected returns; the SegInc dummies estimate the incremental reaction if this segment is also new. Similar to the left part of Table 6, the right part finds that controlling for the post-merger returns does not eliminate the main effect: the slope on NSeg declines by roughly 25% compared to Table 4, but remains statistically significant. The significance is not strong (the t-statistics range between -1.84 and -2.12), but the SegInc dummies have the same problem: on the one hand, their slopes are 2 to 3.5 times larger than the slope on NSeg, meaning that new segments affect future returns much more than existing segments (consistent with similar evidence in the left part of Table 6), on the other hand, SegInc dummies also lack significance, similar to the weakly significant NSeg variable.

Overall, Table 6 confirms that the complexity effect this paper records is, for the larger part, independent of post-merger underperformance. In fact, this paper suggests a different view on the post-merger underperformance: the underperformance can be the product of interaction between relatively larger disagreement about complex firms and short-sale constraints, rather than investors' inability to understand the value-destroying character of mergers.[11]

[11] In untabulated results, I add NewConglo and SegInc dummies to regressions in Table 2 and find that in most cases new conglomerates and firms that have experienced an increase in the number of segments have less analyst following, larger analyst forecast errors and lower institutional ownership than other conglomerates.

5 Complexity Effect and Mispricing

5.1 Diversification Discount and Firm Complexity

The complexity effect in returns suggests a new interpretation of the diversification discount as ´´slow bleeding" of a firm after it becomes a conglomerate. The complexity effect suggests that conglomerates trade at lower multiples at least partly because that conglomerates post inferior returns for a long time and their value slowly deteriorates.

Panel A of Table 7 starts with looking at size, age, and industry adjusted[12] market-to-book (MB) and market-to-sales (MS) ratios of firms in the three years before they become a conglomerate (i.e., add another business segment with a different two-digit SIC code). If the conglomerate is formed from a merger of two publicly traded companies, I only look at the company that retains the pre-merger permno on CRSP.

I find that MB and MS of future conglomerates are significantly higher pre-conglomeration than those of their peers, indicating potential overpricing. This is consistent with the evidence Villalonga (2004b) obtains from a smaller sample of 154 conglomerates that she tracks five years before they are formed and the evidence in Asquith (1983) and Agarwal et al. (1992) that bidders tend to perform abnormally well before they decide to engage in M&A.

The rightmost column of Panel A looks at the difference between actual and imputed MB (MS) of newly formed conglomerates in the year when the conglomerate first appeared as such on Compustat segment files.[13] Starting with Berger and Ofek (1995), imputed MB (MS) assigns to each segment the average MB (MS) ratio of all single-segment firms in the industry (in my case, defined by two-digit SIC code) of the segment. Imputed MB (MS)

[12] The adjustment takes all single-segment firms with the same two-digit SIC code that have market cap and firm age within 10% of those of the future conglomerate. MB and MS of those matching firms are averaged and deducted from MB and MS of the future conglomerate.

[13] Following Villalonga (2004b), I omit excess MB and MS greater than 10 in absolute magnitude as obvious outliers.

of a conglomerate is the weighted average of imputed MB (MS) ratio of its segments, with weights being segment-level assets (sales).

Hund et al. (2021) argue that since MB and MS negatively depend on firm size and age, and single-segment firms tend to be smaller and younger than conglomerates, creating a pseudo-conglomerate out of smaller and younger firms in place of the actual business segments will always produce the impression that conglomerates trade at a discount. Thus, when I compute segment-level MB (MS), I only take those single-segment firms from the same industry that are within 10% of market cap and firm age (years on CRSP) of the conglomerate (not the segment).[14]

The rightmost column of Panel A shows that conglomerates still have positive excess value in their formation year and thus appear overpriced. The first column of Panel C also tests the significance of the difference in excess MB (MS) between the year before (year -1) and the year of conglomerate formation (year 0) and finds both changes to be negative, with the change in excess MB being statistically significant, which can be interpreted as conglomeration being value-destroying.

Panel B of Table 7 look at excess MB (MS) of conglomerates in five-year periods after conglomeration. In the first five years, the evidence on whether conglomerates remain overpriced is mixed: excess MB becomes negative, indicating beginning of underpricing (or value destruction), while excess MS remains positive. What is clear, however, is that both MB and MS significantly decline between years -1 and years 1-5, as well as between years 0 and years 1-5 (see Panel C for formal tests). These declines are consistent with

[14]Hund et al. (2021) use a more complicated matching scheme that involves estimating likelihood of conglomeration as a function of age and size and matching conglomerates and single-segment firms on this likelihood. Hund et al.'s matching reveals no significant diversification discount; in this paper, I do not take a stand on whether the diversification discount exists, my focus in this section is the link between conglomerate valuation and conglomerate age. If my size-age matching is inferior to Hund et al.'s matching, that will only work against me finding that conglomerates are overpriced at the beginning of their life as conglomerates.

Table 6 that reveals particularly strong underperformance of conglomerates in the first three years after they are formed.

Another thing I observe in Panel B is that excess MB and MS of conglomerates significantly drop as conglomerate ages, consistent with my ´´slow bleeding" view of the diversification discount. This is not the issue of MB and MS dropping with firm age, because I match each conglomerate to single-segment firms of the same firm age to compute conglomerate's imputed MB and MS.[15] Rather, it seems that at conglomerate formation, conglomerates are overpriced because of the combination of investor disagreement and short sale constraints: some investors are overoptimistic about synergies, coinsurance, and other benefits of diversification, and the voices of pessimistic investors are not heard in the market to the same extent because of short sale constraints. As the time passes and information about conglomerate performance arrives, the overpricing is slowly resolved and conglomerates begin to trade at lower multiples than comparable single-segment firms, which reflect value-destroying consequences of conglomeration.

Panel B shows that MB and MS of conglomerates slowly drop for at least fifteen years after the conglomerate is formed, consistent with findings in Panel D of Table 5 that the underperformance of conglomerates lasts for about as long. The rightmost column of Panel C presents a formal test of significance between MB and MS of conglomerates of 1-5 years of age and conglomerates of 21+ years of age (as conglomerates) and finds that the difference is significant, thus confirming my main prediction that young conglomerates start overpriced and then subsequently lose value.

Lastly, in Panel D of Table 7, I study whether complexity impacts the magnitude of

[15]Note the distinction of firm age (the matching variable) and conglomerate age (CongAge in Table 5 and the variable I break down the conglomerate sample on in Panels B and D in Table 7). Firm age refers to the number of years the firm has been actively traded, while conglomerate age is the number of years the firm has been a conglomerate. For example, if a firm started as a single-segment firm, then became a conglomerate after ten years and has been a conglomerate for another 15 years, its age is 25 years, but its conglomerate age is 15 years.

the diversification discount. My prediction is that in the first years of their life as conglomerates, more complex conglomerates are going to be more overpriced. Since, as Tables 3 and 4 reveal, more complex conglomerates have stronger underperformance, old complex conglomerates can end up trading at lower valuation multiples than less complex conglomerates (probably also because forming complex conglomerates is more value-destroying). Thus, for all conglomerates taken together the relation of complexity and valuation ratios has an ambiguous sign.

In Panel D, I regress excess MS ratio on controls from Hund et al. (2021) and cluster standard errors by firm-year. In the first column, I find that in the full sample that includes both conglomerates and single-segment firms, 1-HHI measure of complexity is negatively related to market-to-sales ratio.[16] Since 1-HHI measure has a large mass at zero, this result can simply reflect the existence of the diversification discount.[17] The second column of Panel D shows that this is exactly the case: in the sample of conglomerates only, complexity appears unrelated to excess MS ratio.

The next five columns of Table D break down all conglomerates by conglomerate age, i.e., the number of years passed since they became conglomerates, just like Panel B does. As expected, I find that in the first five to ten years of a conglomerate life, more complexity means more overpricing (stemming, according to my explanation of the complexity effect, from the interaction of short sale constraints and disagreement created by complexity). More complexity also means prolonged underperformance and potentially more value destruction; for this reason the relation between excess MS of a conglomerate and the conglomerate's complexity slowly flips the sign to negative and marginally significant at

[16]Untabulated results with MB instead of MS as the left-hand side variable and with NSeg instead of 1-HHI look very similar.

[17]Hund et al. (2021) do not find a significant diversification discount in similar regressions; as the footnote on the previous page explains, the difference in results is due to different matching techniques used in Hund et al. and this paper.

the 10% level in years 11-15 and visibly stronger, negative and significant in years 16-20. Complexity then stops mattering for very old conglomerates (21 years and older), which are likely to be fairly priced and may even be the optimal firm structure if they survived as conglomerates for that long.

5.2 Complexity Effect, Limits to Arbitrage, and Short-Sale Constraints

My explanation of the complexity effect is the interaction of investor disagreement and short-sale constraints along the lines of the Miller (1977) theory: keeping pessimistic investors out of the market creates overpricing, and the overpricing is more severe if there is larger disagreement and the average optimistic investor remaining in the market is more optimistic. The natural implication of that is a stronger complexity effect among short-sale constrained firms. In Panel A of Table 8, I follow a long tradition of using institutional ownership as a proxy for short-sale constraints (see, e.g., Nagel, 2005, Asquith, Pathak, and Ritter, 2005). Since institutions are the main lenders in the short-sale market, low institutional ownership indicates low supply of shares for shorting and, consequently, high shorting fees (and probably also higher search costs). Following Nagel (2005), I also orthogonalize institutional ownership to firm size to make sure I am not capturing any size effects.[18]

Panel A performs three-by-three sorts on the orthogonalized/residual institutional ownership and firm complexity. The complexity groups are defined similarly to Table 3: Zero group includes only single-segment firms (close to 80% of my sample), for which all complexity measures are zero by definition, and Low and High groups include conglomerates with complexity measure (different for each part of the panel) below and above median

[18] Section 3 of online Robustness Appendix repeats Table 8 replacing 1-HHI first with number of segments, NSeg, and then with variability of segment-level market-to-book, RSZ, and arrives at results similar to what is presented in Table 8.

for the portfolio formation year. I do not perform finer sorts on institutional ownership, because the high and low complexity groups are less than 15% of my sample each, and the whole sample for Panel A (firms on Compustat segment files with non-missing institutional ownership on 13F files) has on average 2500 firms per year, and as little as 1500 firms in some years.

Panel A shows that the FF5 alpha differential between single-segment firms and high-complexity conglomerates is particularly large (59 bp per month, t-statistic 5.01) for low institutional ownership group, smaller, but still significant in the medium institutional ownership group, and absent in the high institutional ownership group. This is exactly what my explanation of the complexity effect implies: firm complexity creates disagreement among investors, and if the disagreement is coupled with short-sale constraints (low institutional ownership), overpricing occurs and the complexity effect arises. If the short-sale constraints are not binding (high institutional ownership), then the overpricing does not occur and there is no complexity effect.

Consistent with the Miller story, the larger part of the stronger complexity effect in the low institutional ownership subsample comes from the low FF5 alphas of high-complexity conglomerates, which are 25 bp per month smaller if those high-complexity conglomerates have low institutional ownership. Under the Miller story, the overpricing of high-complexity conglomerates should only exist among short-sale constrained firms.[19]

Panel B of Table 8 performs similar sorts with idiosyncratic volatility instead of institutional ownership. Idiosyncratic volatility is a very popular measure of limits to arbitrage,

[19] The positive alphas of single-segment firms, which are the flip side of the conglomerates overpricing, also increase by 21 bp per month as one goes from high to low institutional ownership subsample. Miller (1977) makes no prediction about the underpricing of low disagreement (zero complexity) firms, though they are bound to have positive alphas, since all alphas sum up to zero, and high disagreement (high complexity) firms have negative alphas. The more positive alphas of single-segment, lower institutional ownership firms can be attributed to the fact that institutional ownership can be a proxy for investor sophistication, and both overpricing and underpricing should thus be stronger for low institutional ownership firms.

used as the main variable in the seminal Shleifer and Vishny (1997) paper. The evidence presented in Panel B is consistent with the complexity effect being mispricing and coming primarily from the short side. The FF5 alpha of high-complexity conglomerates increases, in absolute magnitude, by 65 bp per month when I go from low to high idiosyncratic volatility firms, and reaches -74 bp per month in the high idiosyncratic volatility subsample. The FF5 alpha of single-segment firms, on the other hand, changes from insignificantly positive to negative and marginally significant when I compare low and high idiosyncratic volatility subsamples, with the change in the FF5 alpha of low-complexity conglomerates being in between, closer to the similar change for high-complexity conglomerates. Low-complexity conglomerates also have very negative alphas in the high idiosyncratic volatility groups, but, just like high-complexity conglomerates, do not have significant alphas in any other group.

The only case when I lack power to reject the null of no relation between the complexity effect and idiosyncratic volatility is the bottom right corner, which records the difference in the FF5 alpha differential between single-segment firms and high-complexity firms in high and low idiosyncratic volatility subsample at 27 bp per month, but cannot reject the null that it is zero (and also cannot reject, for example, that it is 60 bp per month).

An interesting by-product of the Panel B analysis is the relation between firm complexity and idiosyncratic volatility effect of Ang et al. (2006), recorded in the last row of Panel B. It turns out that the idiosyncratic volatility effect is at 38 bp per month (marginally significant) for single-segment firms and at 65 bp per month (highly significant) for high-complexity conglomerates, despite single-segment firms being about three times smaller, on average, than conglomerates, and thus more prone to mispricing. This surprising pattern suggests that one can construct a refinement of the low volatility strategy that would not focus on trading in small illiquid firms. The relation between the idiosyncratic volatility effect and firm complexity is consistent with the use of firm complexity as a limits to arbitrage variable first attempted in Barinov et al. (2020) and Barinov (2020).

Panel C double-sorts firms on complexity and estimated probability to be on special using the formula in D'Avolio (2002). A firm is said to be on special if the shorting fee exceeds the risk-free rate; D'Avolio estimates the probability to be on special using proprietary data on shorting fees in 2000-2001. Ali and Trombley (2006) use the formula from D'Avolio (2002) to estimate probability to be on special in a longer period and find that the estimated probability is closely related to shorting fees in several subperiods, for which they have shorting fees data. The advantage of using estimated probability on special is that the formula (see Data Appendix) requires only standard CRSP and Compustat data and thus is available for a larger number of firms and for a longer period than short interest (starts only in 1988) or estimated shorting fee measures based on short interest (e.g., the one in Boehme et al., 2006).

The findings in Panel C are qualitatively the same as in Panel A, confirming that in Panel A institutional ownership picks up short-sale constraints. Panel C finds that the complexity effect is relatively weak for easy-to-short firms (low estimated probability to be on special) and is significantly higher for hard-to-short firms (20 bp vs. 65 bp per month, t-statistics 1.75 and 3.26). The difference is driven primarily by the large negative alpha (-56 bp per month, t-statistic -3.89) of high complexity conglomerates with high probability to be on special.

Panel D double-sorts firms on complexity and the composite overpricing measure of Stambaugh et al. (2015). Stambaugh et al. argue that Miller (1977)-type anomalies will exist only for overpriced stocks, but the pattern in the overpriced stocks subsample will dominate the cross-section because of arbitrage asymmetry. The example Stambaugh et al. use is the Ang et al. (2006) idiosyncratic volatility effect: high volatility overpriced firms are the most overpriced (and thus idiosyncratic volatility is priced negatively among overpriced firms), the reverse is true among underpriced firms, but on average idiosyncratic volatility is negatively priced because shorting volatile overpriced firms is more costly than buying volatile underpriced firms.

Similar argument can be made using firm complexity as a proxy for limits to arbitrage or disagreement/opaque information environment. The results in Panel D partially support this argument: first, the complexity effect is absent for underpriced firms, but it is stronger for overpriced firms. Second, the largest alphas in the double sorts are the negative alphas of conglomerates in the overpriced subsample. On the other hand, the complexity effect seems to be stronger in the middle overpricing group, though the difference is not significant, and there is no ´´negative complexity effect" for underpriced firms.[20] I do find, however, that the Stambaugh et al. overpricing measure is priced only for conglomerates, which is consistent with the use of complexity as a limits to arbitrage variable in Barinov et al. (2020) and Barinov (2020).

5.3 Complexity Effect and Earnings Announcements

Another way of identifying mispricing is looking at earnings announcements, a method proposed by La Porta et al. (1997) and Sloan (1996). If complex firms are overpriced, investors will realize their mistake once they see the earnings and respond by correcting the mispricing, at least partially, at earnings announcements. Since the period around earnings announcements (the day before, the day of, and the day after) is so short, the part of an anomaly (in my case, the complexity effect) that happens around earnings announcements, is almost surely due to resolution of mispricing: if the complexity effect

[20] In Section 3 of online Robustness Appendix, I find that double sorts on the overpricing measure and either NSeg or RSZ produce more monotonic and significant relation between the complexity effect and the overpricing measure.

was risk, less than 5% of the (lower) risk premium of complex firms would happen in the three days around earnings announcements, as those days are less than 5% of roughly 63 trading days in a quarter.

Table 9 regresses cumulative abnormal returns in the three days around earnings announcements on firm complexity measures and standard asset-pricing controls. The complexity measures are significantly related to the announcement returns, which confirms the hypothesis that there is significant price adjustment (correction of mispricing) of complex firms at earnings announcements.

The magnitude of the announcement effect can be estimated by dividing the coefficients in Table 9 by the respective coefficients from Table 4 times three, since the cross-sectional regressions in Table 4 have monthly returns on the left-hand side and announcement returns are effectively quarterly (there is one earnings announcement per quarter). I find, comparing Table 4 and Table 9, that roughly one-quarter to one-third of the complexity effects are concentrated around earnings announcement, depending on what measure of complexity I use. That compares favorably with the concentration of the investment effect at earnings announcements (14%, also estimated comparing Table 4 and Table 9) and is reasonably close to the concentration at earnings announcements of the profitability effect (41%) or the similar concentration of the value effect La Porta et al. (1997) report in their sample (35%).

The 23-31% of the complexity effect concentrated at earnings announcements is the lower bound of the mispricing's share in the complexity effect. Indeed, the information that can help with the resolution of mispricing can come between earnings announcements in all sorts of forms (earnings guidance, management cash flow forecasts, analyst forecast revisions). The large part of the complexity effect being concentrated at earnings announcements is suggestive that mispricing is responsible for an even bigger part of the complexity effect than what Table 9 records.

It is important to note that the result in Table 9 is distinct from the result in Barinov et al. (2020), who find that post-earnings-announcement drift is stronger for complex firms, i.e., that complex firms with good/bad earnings announcements have positive/negative abnormal returns in the post-announcement period. Table 9 looks at the announcement window and shows that complex firms, on average, have negative earnings announcement returns.

6 Alternative Explanations of Complexity Effect

6.1 Complexity Effect and the Business Cycle

Several papers (e.g., Kuppuswamy and Villalonga, 2010, Gopalan and Xie, 2011, and Matvos and Seru , 2014) suggest that conglomerates tend to do better in bad times, presumably because of coinsurance between the segments and well-developed internal capital market. These papers show that for conglomerates recessions lead to smaller distortions in investment and have a smaller negative effect on cash flows, which is consistent with conglomerates being lower risk due to smaller drop in value in recessions, and that could be an explanation of conglomerates' negative alphas.

Table 10 uses several popular business cycle variables to test directly whether conglomerates are low-risk firms. Panel A defines lower risk as a smaller increase in risk premium during recessions and regresses returns to the three complexity groups from Table 3 (single-segment firms, low- and high-complexity conglomerates) on lagged business cycle variables. The risk of all complexity groups seems predictable from the business cycle variables, and in several instances (TED spread, used by Matvos and Seru, 2014, as well as term spread and, to a smaller extent, default premium) I do observe the signs consistent with the idea that conglomerates witness less risk in bad times and are thus low-risk overall. However, similar slopes for single-segment firms are roughly as large, and in no case I observe a significant difference between the two (which would imply that the trading strategy based on the complexity effect has high risk in recessions and is therefore risky).

Panel B runs ICAPM-style regressions of the same returns to complexity groups on the market return and innovations to the business cycle variables from Panel A. The innovations are residuals from ARMA(1,1) model fitted to each business cycle variable separately. The regressions generally do not produce significant coefficients that would suggest that either conglomerates or single-segment firms tend to lose less than average in response to bad news and are thus low-risk (VIX and TED spread appear produce the opposite evidence for conglomerates). More importantly, Panel B agrees with Panel A that none of the six business cycle variables shows that the strategy of buying single-segment firms and shorting conglomerates reacts to business cycle phases in a way that would paint it more/less risky than what standard factor models would suggest.

Another way of looking at risk of conglomerates and their potential performance in recessions is to go back to Panel A2 of Table 3 and assume, as Fama and French (2015) and Hou, Xue, and Zhang (2015) do, that the Fama-French factors are risk factors. As mentioned before, the RMW factor stands out as the most important one and suggests that conglomerates are, in fact, riskier than single-segment firms and would perform worse in a recession. Controlling for this fact is important for the main conclusion of the paper that conglomerates underperform given their level of risk and thus have negative alphas.

6.2 Complexity Effect and Miller (1977) Anomalies

Several well-known anomalies, such as the idiosyncratic volatility effect of Ang et al. (2006) and the analyst disagreement effect of Diether et al. (2002), suggest that disagreement is negatively priced. The existing disagreement effects are unlikely to overlap with the complexity effect, since high idiosyncratic volatility/analyst disagreement firms are typically quite small, while conglomerates are, on average, large firms.

The Miller (1977) theory predicting that disagreement, coupled with short-sale constraints, creates overpricing, can also serve as an explanation of why institutional ownership/short interest are positively/negatively priced (see, e.g., Asquith et al., 2005).

In Section 2 of online Robustness Appendix, I revisit cross-sectional regressions from Table 4 and add, as additional controls, idiosyncratic volatility, analyst disagreement, turnover, institutional ownership and relative short interest, both one-by-one and all together. I find that the complexity effect is little changed controlling for all those variables, confirming my intuition that while the economic mechanism creating the complexity effect and all the anomalies in the two paragraphs above can be similar, the complexity effect highlights a very special case of disagreement pricing, and the firms that it identifies as overpriced were not labeled as such by previous studies.

6.3 Complexity Effect and Coinsurance

Hann et al. (2013) conclude that conglomerates have lower cost of capital by using weighted average cost of equity and cost of debt and defining cost of equity as implied rate of return that sets discounted value of earnings forecasts equal to the current stock price. The complexity effect can also be interpreted as saying that cost of equity for conglomerates is lower, though my proxy for cost of equity would be average future abnormal returns (net of profitability and investment anomalies) rather than implied rate of return.

Hann et al. argue that lower cost of capital for conglomerates arises due to coinsurance between segments and shows that the cost of capital drops as cash flow correlation between the segments declines, and the gap in cost of capital between conglomerates and single-segment firms increases for financially constrained firms.

In Section 4 of online Robustness Appendix, I split the sample on between-segment cash flow correlation and re-run cross-sectional regressions from the first four columns of Table 4 in each cash flow correlation group. I do not find that the complexity effect depends on correlation between cash-flow correlation between the segments.

I also split the sample on several financial constraints measures, as well as on credit rating, and arrive at mixed evidence: the relation between the complexity effect and financial constraints takes different signs depending on the measure of financial constraints I use, and the relation between the complexity effect and credit rating is non-monotonic.

I conclude that the complexity effect is distinct from the cost of capital effect in Hann et al. (2013) and is unrelated to benefits of coinsurance. The latter is consistent with results in Table 10 that does not find any evidence that conglomerates perform better than single-segment firms during recessions.

6.4 Complexity Effect and Relation between Diversification Discount and Expected Returns

Lamont and Polk (2001) argue that conglomerates trade at lower multiples because their expected returns are higher. Lamont and Polk do not find substantial difference in expected returns between single-segment firms and conglomerates (controlling for MKT, SMB, and HML factors, but not for CMA and RMW), but they do find that conglomerates with larger diversification discount have higher expected returns. Depending on the correlation between diversification discount and my complexity measures, this finding can either work against me (if the correlation is positive) or suggest that my result and the result of Lamont and Polk (2001) overlap.

In Section 5 of online Robustness Appendix, I split the conglomerates-only sample on the magnitude of the diversification discount and find that the complexity effect is stronger for conglomerates with a larger discount, suggesting that firm complexity and the diversification discount are positively related and Lamont and Polk (2001) results are working against me. The stronger complexity effect for conglomerates that end up being traded at lower multiples is consistent with my hypothesis that the diversification discount accumulates slowly as a result of the complexity effect (and the underperformance of conglomerates it implies) being at work for a prolonged period of time.

I also explore in Robustness Appendix the potential relation between the complexity effect and the Lamont and Polk (2001) result by adding the magnitude of the diversification discount to the conglomerates-only regression in the last column of Table 4. Consistent with the results in the previous paragraph, I find that the complexity effect becomes stronger once I control for the link between the diversification discount and expected returns.

Mitton and Vorkink (2010) suggest an explanation of the Lamont and Polk (2001) result: if investors have a skewness preference, conglomeration would increase expected returns by diversifying away idiosyncratic skewness. In online Robustness Appendix, I split the sample into high and low skewness groups and re-run regressions from the first four columns of Table 4 for each skewness group. I find that the complexity effect is similar for low and high skewness firms, confirming again that the Lamont and Polk (2001) result and the complexity effect are two independent phenomena.

6.5 Complexity Effect and Financial Constraints

One benefit of conglomeration is creation of internal capital markets (and coinsurance between segments, as Hann et al., 2013, point out), which alleviates financial constraints. Assuming that financial constraints are priced and more financially constrained firms are risker, one can argue that a possible reason for lower expected returns of conglomerates is their lower financial constraints.

In Section 6 of online Robustness Appendix, I form financial constraints factors by sorting firms into quintiles on several measures of financial constraints (such as, e.g., the Hadlock and Pierce, 2010, index) and financial distress (such as, e.g., expected default frequency, EDF) and taking the return differential between the top (most financially constrained) and the bottom quintiles. I add these factors one-by-one to the five-factor Fama and French (2015) model fitted to returns to the zero-minus-high complexity portfolio and find that the said portfolio loads negatively on all financial constraints factors, suggesting that the long leg of the portfolio (highly complex conglomerates) comoves more with the long leg of the financial constraints factors. The exception is the EDF-based factor, which the zero-minus-high complexity portfolio loads positively on, but even in this case controlling for financial constraints/distress changes the alpha of the zero-minus-high complexity portfolio by only a few bp per month.

7 Conclusion

The paper presents a long-term negative relation between firm complexity and expected returns (the complexity effect). The complexity effect is 35 bp per month in portfolio sorts and close to 60 bp per month in subsamples of firms with higher limits to arbitrage.

I show that conglomerates have lower analyst coverage, lower institutional ownership, and higher analyst disagreement than single-segment firms with comparable characteristics, and higher-complexity conglomerates have even lower information quality and higher disagreement. This extra disagreement, coupled with short-sale constraints, creates overpricing of conglomerates - under short-sale constraints, pessimistic investors (partially) stay out of the market, and the optimists, who set the prices, are on average more optimistic if there is more disagreement. The overpricing of conglomerates is stronger if a conglomerate is more complex (has more business segments in different industries, has sales more dispersed among the segments, has segments that are more different in terms of growth opportunities).

The complexity effect suggests a new interpretation of the diversification discount: lower valuation ratios at which conglomerates trade appear to be a product of many years of underperformance captured by the complexity effect. I find that firms that later become conglomerates has positive excess valuation ratios for three years before conglomeration; the excess valuation ratios remain positive in the first years after conglomerate formation and strongly decline as conglomerate ages. Even more, I find that conglomerates' valuation ratios positively depend on complexity in the first decade of a conglomerate life (more complex conglomerates seem to be more overpriced), but this relation switches to significantly negative in the second decade after a conglomerate is formed (more complex conglomerates underperform more).

Consistent with the mispricing explanation, the complexity effect is stronger for firms with low institutional ownership (that is, lower supply of shares to short), higher idiosyncratic volatility (higher limits to arbitrage), and higher estimated probability of high shorting fees. Roughly 30% of the complexity effect is concentrated at earnings announcements, when investors are more likely to correct mispricing.

I control for several seemingly related effects that are often attributed to the interaction of short sales and disagreement (the idiosyncratic volatility effect of Ang et al., 2006, the analyst disagreement effect of Diether et al., 2006, the short interest effect of Asquith et al., 2005), as well as for the relation between diversification discount and expected returns from Lamont and Polk (2001), and find that the complexity effect is, if anything, stronger, when all those effects are controlled for. The complexity effect is also distinct from the long-run underperformance of bidders post-acquisition.

Firm complexity seems to be a special disagreement variable that is positively related to size. Trading on the complexity effect, in contrast to trading on other disagreement anomalies, does not include trading in very small, very volatile, illiquid or distressed companies, which makes the complexity effect easy to exploit and harder to explain. In fact, removing from the sample firms with below median liquidity does not materially affect the magnitude of the complexity effect.

I do not find any signs that the complexity effect is related to risk: the long-short strategy that buys single-segment firms and shorts conglomerates does not seem to respond in a negative way to bad news about a long list of business-cycle variables, and predictive regressions for the long-short strategy do not find that its risk shifts over the business cycle. While prior work (see, e.g., Hann et al., 2013) suggested that conglomerates can have low risk because of coinsurance between segments, I do not find that the magnitude of the complexity effect is unambiguously related to cross-segment cash-flow correlation or financial constraints.

References

[1] Agrawal, Anup, Jeffrey F. Jaffe and Gershon N. Mandelker, 1992, The Post-Merger Performance of Acquiring Firms: A Re-Examination of an Anomaly, Journal of Finance 47, 1605-1621.

[2] Amihud, Yakov, 2002, Illiquidity and Stock returns: Cross-Section and Time-Series Effects, Journal of Financial Markets 5, 31–56.

[3] Ali, Ashiq, and Mark A. Trombley, 2006, Short Sales Constraints and Momentum in Stock Returns, Journal of Business Finance and Accounting 33, 587–615.

[4] Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2006, The Cross-Section of Volatility and Expected Returns, Journal of Finance 61, 259–299.

[5] Asquith, Paul, 1983, Merger Bids, Uncertainty, and Stockholder Returns, Journal of Financial Economics 11, 51–83.

[6] Asquith, Paul, Parag A. Pathak, and Jay R. Ritter, 2005, Short Interest, Institutional Ownership, and Stock Returns, Journal of Financial Economics 78, 243–276.

[7] Barinov, Alexander, 2020, Firm Complexity and Limits to Arbitrage, Working paper.

[8] Barinov, Alexander, Shawn S. Park, and Celim Yildizhan, 2020, Firm Complexity and Post Earnings Announcement Drift, Working paper.

[9] Barth, Mary E., Ron Kasznik, and Maureen McNichols, 2001, Analyst Coverage and Intangible Assets, Journal of Accounting Research 39, 1—34.

[10] Berger, Philip G., and Rebecca Hann, 2003, The Impact of SFAS No. 131 on Information and Monitoring, Journal of Accounting Research 41, 163—223.

[11] Berger, Philip G., and Eli Ofek, 1995, Diversification's Effect on Firm Value, Journal of Financial Economics 37, 39–66.

[12] Bhushan, Ravi, 1989, Firm Characteristics and Analyst Following, Journal of Accounting and Economics 11, 255-–74.

[13] Boehme, Rodney D., Bartley R. Danielsen, and Sorin M. Sorescu, 2006, Short Sale Constraints, Differences of Opinion, and Overvaluation, Journal of Financial and Quantitative Analysis 41, 455–488.

[14] Carhart, Mark M., 1997, On the Persistence in Mutual Funds Performance, Journal of Finance 52, 57–82.

[15] Chordia, Tarun, Sahn-Wook Huh, and Avanidhar Subrahmanyam, 2007, The Cross-Section of Expected Trading Activity, Review of Financial Studies 20, 709–741.

[16] Cohen, Lauren, and Dong Lou, 2012, Complicated Firms, Journal of Financial Economics 104, 383–400.

[17] Custodio, Claudia, 2014. Mergers and Acquisitions Accounting and the Diversification Discount. Journal of Finance 69, 219–240.

[18] Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers, 1997, Measuring Mutual Fund Performance with Characteristic-Based Benchmarks, Journal of Finance 52, 1035–1058.

[19] D'Avolio, Gene, 2002, The Market for Borrowing Stock, Journal of Financial Economics 66, 271–306.

[20] Diether, Karl, Christopher Malloy, and Anna Scherbina, 2002, Differences of Opinion and the Cross-Section of Returns, Journal of Finance 57, 2113–2141.

[21] Fama, Eugene F., and Kenneth R. French, 1993, Common Risk Factors in the Returns on Stocks and Bonds, Journal of Financial Economics 33, 3–56.

[22] Fama, Eugene F., and Kenneth R. French, 1995, Size and Book-to-Market Factors in Earnings and Returns, Journal of Finance 50, 131–155.

[23] Fama, Eugene F., and Kenneth R. French, 2015, A Five-Factor Asset Pricing Model, Journal of Financial Economics 116, 1–22.

[24] Fama, Eugene F., and MacBeth, James, 1973, Risk, Return, and Equilibrium: Empirical Tests, Journal of Political Economy 81, 607–636.

[25] Gao, Xiaohui, and Jay R. Ritter, 2010, The Marketing of Seasoned Equity Offerings, Journal of Financial Economics 97, 33–52.

[26] Gilson, Stuart C., Paul M. Healy, Christopher F. Noe, and Krishna G. Palepu, 2001, Analyst Specialization and Conglomerate Stock Breakups, Journal of Accounting Research 39, 565–582.

[27] Gopalan, Radhakrishnan, and Kangzhen Xie, 2011, Conglomerates and Industry Distress, Review of Financial Studies 24, 3642–3687.

[28] Hadlock, Charles J., and Joshua R. Pierce, 2010, New Evidence on Measuring Financial Constraints: Moving Beyond the KZ Index, Review of Financial Studies 23, 1909—1940.

[29] Hann, Rebecca N., Maria Ogneva, and Oguzhan Ozbas, 2013, Corporate Diversification and Stock Returns, Journal of Finance 68, 1961–1999.

[30] Harvey, Campbell, Yan Liu, and Heqing Zhu, 2016, ... and the Cross-Section of Expected Returns, Review of Financial Studies 29, 5–68.

[31] Herrmann, Don, and Wayne B. Thomas, 2000, An Analysis of Segment Disclosures under SFAS No. 131 and SFAS No. 14, Accounting Horizons 14, 287-–302.

[32] Hoberg, Gerard, and Gordon M. Phillips, 2014, Product Market Uniqueness, Organizational Form and Stock Market Valuations, Working Paper

[33] Holden, Craig W., 2009, New Low-Frequency Spread Measures, Journal of Financial Markets 12, 778–813.

[34] Hou, K., Xue, C., and Zhang, L., 2015, Disgesting Anomalies: An Investment Approach, Review of Financial Studies 28, 650–705.

[35] Hund, John E., Donald Monk, and Sheri Tice, 2021, The Berger-Ofek Diversification Discount Is Just Poor Firm Matching, Critical Finance Review, forthcoming.

[36] Kaplan, Steven N., and Luigi Zingales, 1997, Do Investment-Cash Flow Sensitivities Provide Useful Measures of Financing Constraints? Quarterly Journal of Economics 115, 169–215.

[37] Kuppuswamy, Venkat, and Belen Villalonga, 2016, Does Diversification Create Value in the Presence of External Financing Constraints? Evidence from the 2007–2009 Financial Crisis, Management Science 62, 905–923.

[38] Lamont, Owen A., and Christopher Polk, 2001, The Diversification Discount: Cash Flows versus Returns, Journal of Finance 56, 1693–1721.

[39] Larry H. P. Land, and Rene M. Stulz, 1994, Tobin's q, Corporate Diversification, and Firm Performance, Journal of Political Economy 102, 1248–1280.

[40] La Porta, Rafael, Josef Lakonishok, Andrei Shleifer, and Robert Vishny, 1997, Good News for Value Stocks: Further Evidence on Market Efficiency, Journal of Finance 52, 859–874.

[41] Lesmond, David A., Ogden, Joseph P., Trzcinka, Charles A., 1999, A New Estimate of Transaction Costs, Review of Financial Studies 12, 1113–1141.

[42] Matvos, Gregor, and Amit Seru, 2014, Resource Allocation within Firms and Financial Market Dislocation: Evidence from Diversified Conglomerates, Review of Financial Studies 27, 1143–1189.

[43] Miller, Edward M., 1977, Risk, Uncertainty, and Divergence of Opinion, Journal of Finance 32, 1151–1168.

[44] Mitton, Todd, and Keith Vorkink, 2010, Why Do Firms with Diversification Discounts Have Higher Expected Returns? Journal of Financial and Quantitative Analysis 45, 1367–1390.

[45] Nagel, Stefan, 2005, Short Sales, Institutional Investors, and the Cross-Section of Stock Returns, Journal of Financial Economics 78, 277–309.

[46] Newey, Whitney K., and Kenneth D. West, 1987, A Simple Positive Semi-Definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, Econometrica 55, 703–708.

[47] Peterson, Mitchell A., 2009, Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches, Review of Financial Studies 22, 435–480.

[48] Rajan, Raghuram, Henry Servaes, and Luigi Zingales, 2000, The Cost of Diversity: The Diversification Discount and Inefficient Investment, Journal of Finance 55, 35–80.

[49] Roll, Richard, 1984, A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market, Journal of Finance 39, 1127–1139.

[50] Sloan, Richard G., 1996, Do Stock Prices Fully Reflect Information in Accruals and Cash Flows about Future Earnings? Accounting Review 71, 289–315.

[51] Shleifer, Andrei, and Robert Vishny, 1997, The Limits of Arbitrage, Journal of Finance 52, 35–55.

[52] Stambaugh, Robert F., Jianfeng Yu, Yu Yuan, 2015, Arbitrage Asymmetry and the Idiosyncratic Volatility Puzzle, Journal of Finance 70, 1903–1948.

[53] Villalonga, Belen, 2004a, Diversification Discount or Premium? New Evidence from the Business Information Tracing Series, Journal of Finance 59, 479–506.

[54] Villalonga, Belen, 2004a, Does Diversification Cause the ´´Diversification Discount"? Financial Management 2004, 5–27.

[55] Whited, Toni M., and Guojun Wu, 2006, Financial Constraints Risk, Review of Financial Studies 19, 531–559.

Table 1. Descriptive Statistics, Size-Adjusted

The table presents size-adjusted averages of firm characteristics for single-segment firms and conglomerates, as well as the difference between the two (S-C). Conglomerates are defined as firms with business segments in more than one industry, industries are defined using two-digit SIC codes. Single-segment firms are firms that report business segments in only one industry on Compustat segment files.

The table also splits conglomerates into low and high complexity groups, depending on whether Complexity variable (Comp) is above or below the median. The second right-most column (H-L) reports the difference in firm characteristics between high and low-complexity conglomerates, and the rightmost column (L-S) compares low-complexity conglomerates with single-segment firms. Complexity variable is defined as the sales concentration across segments, 1-HHI, where HHI is the sum of squared shares of segment sales in total firm revenue.

Panel A looks at standard asset-pricing controls - market-to-book (MB), investment-to-assets (Inv), profitability (Prof), returns in the portfolio formation year, excluding the last month return (Mom).

Panel B reports information environment proxies - idiosyncratic volatility (IVol), analyst forecast dispersion (Disp), analyst forecast error (Error), earnings variability (CVEarn), number of analysts and specialist analysts following the firm (# An and # Spec), institutional ownership (IO).

Detailed definitions of all variables are in Data Appendix. The t-statistics use Newey-West (1987) correction for heteroscedasticity and autocorrelation. The sample period is from January 1978 to December 2016. The sample excludes stocks priced below $5 on the portfolio formation date.

Panel A. Complexity and Asset Pricing Controls

	Single	Conglo	S-C	t-stat	LComp	HComp	H-L	t-stat	L-S	t-stat
MB	0.373	-1.047	1.419	*6.33*	-0.658	-1.283	-0.625	*-2.14*	-1.031	*-3.71*
Inv	0.056	-0.143	0.199	*2.96*	-0.129	-0.167	-0.038	*-1.12*	-0.185	*-2.56*
Prof	0.029	0.005	0.024	*8.95*	0.015	0.001	-0.014	*-7.24*	-0.014	*-5.23*
Mom	0.013	-0.014	0.026	*2.92*	-0.004	-0.019	-0.015	*-3.68*	-0.017	*-2.24*

Panel B. Complexity and Information Environment

	Single	Conglo	S-C	t-stat	LComp	HComp	H-L	t-stat	L-S	t-stat
IVol	0.213	-0.119	0.333	*12.4*	-0.045	-0.146	-0.101	*-7.42*	-0.259	*-12.9*
Disp	0.432	2.002	-1.569	*-2.71*	1.657	2.232	0.575	*0.78*	1.224	*2.16*
Error	-2.878	5.040	-7.917	*-3.33*	0.382	7.334	6.951	*1.91*	3.260	*1.22*
# An	0.412	-0.655	1.068	*19.7*	-0.240	-0.941	0.337	*3.04*	0.426	*6.35*
# Spec	*0.360*	*-0.825*	*1.185*	*20.4*	*-0.264*	*-1.269*	*-1.005*	*-22.2*	*-0.624*	*-12.8*
CVEarn	-0.196	0.481	-0.677	*-9.52*	0.230	0.566	-0.701	*-19.2*	-0.653	*-12.8*

Table 2. Firm Complexity and Information Environment

The table presents panel regressions of information environment proxies - number of analysts following the firm (# An), number of analysts that are specialists in the firm's (main) business (# Spec), institutional ownership (IO), analyst forecast error (Error), and analyst forecast dispersion (Disp) - on the sales-based Comp variable and the controls from the literature. Comp variable is sales concentration across segments, 1-HHI, where HHI is the sum of squared shares of segment sales in total firm revenue. Panel A looks at the full sample (firms with Compustat segment data), Panel B restricts the sample to conglomerates only. Conglomerates are defined as firms with business segments in more than one two-digit SIC industry. The controls include the dummy variable for Nasdaq-traded stocks (Nasdaq), market beta (Beta), reciprocal of stock price (Price), stock turnover (Turn, dollar trading volume over market cap), total return volatility (TVol), the dummy variable for S&P 500 stocks (S&P500), firm leverage (Lev), R&D expenditures over total assets (R&D), intangible assets over total assets (Intan), firm age (Age, the number of months since the firm first appeared on CRSP), dividend yield (Div), cumulative returns in the past quarter/year (RetQ1/RetYR1). Detailed definitions of all variables are in Data Appendix. All control variables are percentage ranks. The t-statistics use standard errors clustered by firm-year-quarter. The sample period is from January 1978 to December 2016. The sample excludes stocks priced below $5.

Panel A. Firm Complexity and Firm Information Environment: Full Sample

# An	**1**	**# Spec**	**2**	**IO**	**3**	**Error**	**4**	**Disp**	**5**
Const	11.65	**Const**	-18.33	**Const**	16.96	**Const**	81.33	**Const**	11.69
t-stat	*2.66*	**t-stat**	*-3.35*	**t-stat**	*5.34*	**t-stat**	*5.11*	**t-stat**	*3.80*
Comp	-27.60	**Comp**	-60.86	**Comp**	-15.08	**Comp**	21.78	**Comp**	15.30
t-stat	*-8.72*	**t-stat**	*-13.8*	**t-stat**	*-7.20*	**t-stat**	*2.09*	**t-stat**	*5.76*
Nasdaq	11.37	**Nasdaq**	1.347	**Age**	-0.027	**Intan**	-0.014	**Intan**	-0.057
t-stat	*7.87*	**t-stat**	*0.72*	**t-stat**	*-2.01*	**t-stat**	*-0.22*	**t-stat**	*-2.68*
Beta	-0.115	**Beta**	-0.196	**Div**	0.021	**BLev**	0.279	**BLev**	0.149
t-stat	*-6.32*	**t-stat**	*-8.13*	**t-stat**	*0.59*	**t-stat**	*3.80*	**t-stat**	*8.04*
MB	0.039	**MB**	-0.062	**MB**	-0.065	**R&D**	0.164	**R&D**	0.166
t-stat	*1.85*	**t-stat**	*-2.25*	**t-stat**	*-4.68*	**t-stat**	*2.36*	**t-stat**	*8.71*
Size	2.638	**Size**	2.410	**Size**	0.160	**Size**	-1.084	**Size**	-0.312
t-stat	*61.0*	**t-stat**	*40.6*	**t-stat**	*5.78*	**t-stat**	*-7.21*	**t-stat**	*-12.0*
Price	0.118	**Price**	-0.228	**Price**	0.316	**TVolD**	0.544	**TVolD**	0.298
t-stat	*3.08*	**t-stat**	*-4.20*	**t-stat**	*14.6*	**t-stat**	*3.85*	**t-stat**	*11.5*
RetYR1	-0.482	**RetYR1**	-0.534	**RetQ1**	0.026				
t-stat	*-35.2*	**t-stat**	*-32.0*	**t-stat**	*6.00*				
RetYR2	-0.309	**RetYR2**	-0.320	**RetYR1**	-0.009				
t-stat	*-28.9*	**t-stat**	*-23.3*	**t-stat**	*-1.42*				
Turn	0.611	**Turn**	0.442	**S&P500**	1.895				
t-stat	*25.9*	**t-stat**	*13.2*	**t-stat**	*1.51*				
TVol	0.031	**TVol**	0.892						
t-stat	*0.82*	**t-stat**	*15.9*						

Panel B. Firm Complexity and Firm Information Environment: Conglomerates Only

# An	1	# Spec	2	IO	3	Error	4	Disp	5
Const	-5.277	**Const**	-22.62	**Const**	18.25	**Const**	108.39	**Const**	5.360
t-stat	*-0.71*	**t-stat**	*-2.32*	**t-stat**	*3.85*	**t-stat**	*4.65*	**t-stat**	*0.85*
Comp	-33.53	**Comp**	-77.57	**Comp**	-19.22	**Comp**	30.17	**Comp**	17.00
t-stat	*-8.25*	**t-stat**	*-13.3*	**t-stat**	*-7.17*	**t-stat**	*2.48*	**t-stat**	*4.57*
Nasdaq	18.81	**Nasdaq**	4.705	**Age**	-0.149	**Intan**	0.002	**Intan**	-0.128
t-stat	*8.33*	**t-stat**	*1.45*	**t-stat**	*-6.73*	**t-stat**	*0.02*	**t-stat**	*-3.64*
Beta	-0.067	**Beta**	-0.102	**Div**	0.166	**BLev**	0.221	**BLev**	0.157
t-stat	*-2.08*	**t-stat**	*-2.40*	**t-stat**	*3.36*	**t-stat**	*1.73*	**t-stat**	*4.76*
MB	0.081	**MB**	0.004	**MB**	-0.013	**R&D**	0.269	**R&D**	0.127
t-stat	*2.18*	**t-stat**	*0.08*	**t-stat**	*-0.59*	**t-stat**	*1.96*	**t-stat**	*3.32*
Size	2.874	**Size**	2.638	**Size**	0.133	**Size**	-1.395	**Size**	-0.211
t-stat	*40.1*	**t-stat**	*25.4*	**t-stat**	*2.98*	**t-stat**	*-5.69*	**t-stat**	*-3.48*
Price	0.094	**Price**	-0.432	**Price**	0.273	**TVolD**	0.311	**TVolD**	0.401
t-stat	*1.50*	**t-stat**	*-4.93*	**t-stat**	*7.91*	**t-stat**	*1.59*	**t-stat**	*7.53*
RetYR1	-0.446	**RetYR1**	-0.573	**RetQ1**	0.029				
t-stat	*-17.9*	**t-stat**	*-19.6*	**t-stat**	*4.57*				
RetYR2	-0.300	**RetYR2**	-0.351	**RetYR1**	0.003				
t-stat	*-16.1*	**t-stat**	*-13.8*	**t-stat**	*0.28*				
Turn	0.585	**Turn**	0.275	**S&P500**	-0.114				
t-stat	*15.1*	**t-stat**	*4.47*	**t-stat**	*-0.06*				
TVol	-0.069	**TVol**	1.046						
t-stat	*-1.26*	**t-stat**	*13.1*						

Table 3. Firm Complexity and Expected Returns: Portfolio Sorts

The table sorts firms into three groups on firm complexity - the Zero group includes all single-segment firms, the Low/High group includes conglomerates with Comp=1-HHI below/above median. Conglomerates are defined as firms with business segments in more than one industry, industries are defined using two-digit SIC codes. Single-segment firms are firms that report business segments in only one industry on Compustat segment files. HHI is the sum of squared shares of segment sales in total firm revenue.

Panel A1 reports value-weighted alphas of the complexity groups and the differences between them using several models: the five-factor Fama-French (2015) model (α_{FF5}), the three-factor Fama-French (1993) model (FF3) augmented with the profitability/investment factor from the FF5 model ($\alpha_{FF3+RMW}/\alpha_{FF3+RMW}$), and the FF5 model augmented with the momentum factor ($\alpha_{FF5+MOM}$). The last two rows of the left part add the dummy for post-SFAS131 period (1 in 1999-2016, 0 in 1978-1998) to the FF5 model and present the FF5 alpha pre-SFAS131 (the intercept) and the change in the alpha post-SFAS131 (the slope on the dummy). Panels A2 reports the factor betas of the complexity groups from the FF5 model.

Panel B reports alphas from the FF5 model fitted to complexity groups based on the Comp measure, but with firms with below-median size or below-median liquidity at the portfolio formation date omitted from the sample. The measure of liquidity used is indicated in the left-most column. The liquidity measures are: Roll (1984) bid-ask spread measure, effective bid-ask spread measure of Corwin and Schultz (2012), and effective tick of Holden (2009). Amihud (2002) measure estimates price impact (in percent of stock price per \$1 million trade) by dividing absolute daily return by trading volume and averaging the ratio within a firm-year. Zero return frequency (Zero) is the fraction of days with no price change (and no trade) in a year.

Detailed definitions of all variables are in Data Appendix. The t-statistics use Newey-West (1987) correction for heteroscedasticity and autocorrelation. The sample period is from January 1978 to December 2016. The sample excludes stocks priced below \$5 on the portfolio formation date.

Panel A1. Sales-Based Complexity and Alphas

	Zero	Low	High	Z-H	Z-L	L-H
α_{FF5}	0.194	-0.009	-0.162	0.354	0.203	0.154
t-stat	*3.04*	*-0.17*	*-2.70*	*4.01*	*2.46*	*1.85*
$\alpha_{FF5+MOM}$	0.240	0.014	-0.114	0.353	0.226	0.129
t-stat	*3.50*	*0.29*	*-1.82*	*4.06*	*2.61*	*1.47*
$\alpha_{FF3+CMA}$	0.044	0.013	-0.075	0.119	0.031	0.090
t-stat	*0.64*	*0.26*	*-1.26*	*1.14*	*0.35*	*1.13*
$\alpha_{FF3+RMW}$	0.162	0.020	-0.115	0.276	0.142	0.136
t-stat	*2.75*	*0.37*	*-2.05*	*3.41*	*1.85*	*1.71*
$\alpha_{pre-SFAS}$	0.169	-0.037	-0.184	0.353	0.205	0.147
t-stat	*2.27*	*-0.65*	*-2.79*	*3.11*	*2.10*	*1.71*
$\Delta\alpha_{post-SFAS}$	0.054	0.061	0.049	0.004	-0.006	0.017
t-stat	*0.50*	*0.67*	*0.49*	*0.03*	*-0.05*	*0.12*

Panel A2. Sales-Based Complexity and FF5 Betas

	Zero	Low	High	Z-H	Z-L	L-H
β_{MKT}	0.962	1.004	1.084	-0.122	-0.042	-0.080
t-stat	*49.6*	*46.8*	*77.6*	*-4.61*	*-1.58*	*-3.25*
β_{SMB}	0.007	-0.047	-0.076	0.083	0.053	0.029
t-stat	*0.23*	*-1.91*	*-2.82*	*1.93*	*1.42*	*0.91*
β_{HML}	-0.096	-0.036	0.032	-0.128	-0.059	-0.067
t-stat	*-2.73*	*-0.98*	*0.96*	*-2.84*	*-1.26*	*-1.76*
β_{CMA}	-0.121	0.184	0.177	-0.298	-0.306	0.006
t-stat	*-2.09*	*2.84*	*3.34*	*-4.69*	*-4.59*	*0.08*
β_{RMW}	-0.309	0.117	0.174	-0.483	-0.426	-0.057
t-stat	*-7.17*	*3.09*	*3.94*	*-9.43*	*-8.00*	*-1.27*

Panel B. FF5 Alphas of Sales-Based Complexity Groups, Firms with Liquidity below Median Removed

	Zero	Low	High	Z-H	Z-M	L-H		Zero	Low	High	Z-H	Z-M	L-H
Roll	0.174	-0.116	-0.134	0.309	0.290	0.019	**Amihud**	0.175	-0.094	-0.178	0.353	0.269	0.084
t-stat	*3.03*	*-1.46*	*-2.19*	*3.10*	*3.25*	*0.18*	t-stat	*2.77*	*-1.24*	*-2.92*	*3.47*	*2.94*	*0.82*
Spread	0.026	-0.104	-0.190	0.216	0.129	0.087	**Zero**	0.192	-0.102	-0.191	0.383	0.294	0.089
t-stat	*0.47*	*-1.34*	*-3.05*	*2.38*	*1.30*	*0.83*	t-stat	*3.00*	*-1.34*	*-3.06*	*3.77*	*3.21*	*0.89*
Efftick	0.201	-0.090	-0.175	0.377	0.292	0.085	**Size**	0.228	-0.103	-0.175	0.403	0.332	0.071
t-stat	*3.13*	*-1.20*	*-2.85*	*3.68*	*3.25*	*0.83*	t-stat	*3.28*	*-1.28*	*-2.82*	*3.72*	*3.37*	*0.65*

Table 4. Firm Complexity and Expected Returns: Cross-Sectional Regressions

The table presents firm-level Fama-MacBeth (1973) regressions of characteristic-adjusted (DGTW) returns on complexity measures and standard asset-pricing controls. DGTW adjustment (described in detail in the Data Appendix) adjusts returns for size, market-to-book, and momentum. The complexity measures include the conglomerate dummy (Conglo, 1 for conglomerates, zero for single-segment firms), the Comp variable, defined as the sales concentration across segments, 1-HHI, where HHI is the sum of squared shares of segment sales in total firm revenue, the number of business segments with different two-digit SIC codes (NSeg), and the RSZ complexity measure, which is the standard deviation of segment asset-weighted market-to-books for the conglomerate divided by the equal-weighted average market-to-book of the segments. Conglomerates are defined as firms with more than one segment with different two-digit SIC codes. Single-segment firms are the rest of the firms on Compustat segment files. HiComp/LoComp variable is 1 if the Comp measure is above/below median, and 0 otherwise. The median complexity is determined separately in each quarter using only conglomerates. The right-most column of Panel C restricts the sample to conglomerates only. The standard asset pricing controls include the market beta (Beta), market capitalization (Size), market-to-book (MB), cumulative returns in months t-2 to t-12 (Mom), returns in the previous month (Rev), investment (Inv), gross profitability (Prof). Detailed definitions of all variables are in Data Appendix. The t-statistics use Newey-West (1987) correction for heteroscedasticity and autocorrelation. The sample period is from January 1978 to December 2016. The sample excludes stocks priced below $5.

	Compustat Segment Firms				All Compustat Firms				High vs. Low Complexity			
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**		**9**	**10**	**11**
Beta	0.115	0.115	0.114	0.117	0.080	0.081	0.080	0.080	**Beta**	0.094	0.094	0.058
t-stat	*1.14*	*1.14*	*1.14*	*1.16*	*0.88*	*0.88*	*0.87*	*0.87*	**t-stat**	*0.93*	*0.93*	*0.53*
Size	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**Size**	0.000	0.000	0.000
t-stat	*-1.31*	*-1.36*	*-1.12*	*-1.37*	*-2.53*	*-2.57*	*-2.18*	*-2.72*	**t-stat**	*-0.27*	*-0.27*	*0.20*
MB	-0.028	-0.028	-0.028	-0.025	-0.005	-0.005	-0.005	-0.005	**MB**	-0.038	-0.038	-0.045
t-stat	*-3.91*	*-3.95*	*-3.85*	*-3.70*	*-2.46*	*-2.47*	*-2.46*	*-2.45*	**t-stat**	*-4.56*	*-4.56*	*-3.24*
Mom	-0.073	-0.073	-0.074	-0.072	-0.072	-0.071	-0.072	-0.070	**Mom**	-0.091	-0.091	-0.086
t-stat	*-1.44*	*-1.44*	*-1.46*	*-1.40*	*-1.65*	*-1.64*	*-1.65*	*-1.62*	**t-stat**	*-1.66*	*-1.66*	*-0.97*
Rev	-0.047	-0.047	-0.047	-0.047	-0.043	-0.043	-0.043	-0.043	**Rev**	-0.048	-0.048	-0.049
t-stat	*-11.0*	*-11.0*	*-11.0*	*-11.0*	*-11.28*	*-11.29*	*-11.29*	*-11.29*	**t-stat**	*-10.6*	*-10.6*	*-9.81*
Inv	-0.918	-0.923	-0.921	-0.936	-0.282	-0.283	-0.283	-0.280	**Inv**	-1.038	-1.038	-1.104
t-stat	*-3.99*	*-4.00*	*-4.01*	*-3.97*	*-2.69*	*-2.70*	*-2.72*	*-2.69*	**t-stat**	*-4.28*	*-4.28*	*-3.91*
Prof	0.484	0.485	0.481	0.476	0.441	0.443	0.439	0.442	**Prof**	0.507	0.507	0.792
t-stat	*3.31*	*3.33*	*3.29*	*3.26*	*3.34*	*3.35*	*3.32*	*3.35*	**t-stat**	*3.41*	*3.41*	*4.07*
Conglo	-0.125				-0.100				**HiComp**	-0.167	-0.079	-0.086
t-stat	*-3.10*				*-2.50*				**t-stat**	*-3.17*	*-1.47*	*-1.72*
Comp		-0.324				-0.259			**LoComp**	-0.088		
t-stat		*-3.44*				*-2.75*			**t-stat**	*-1.62*		
NSeg			-0.067				-0.054		**Conglo**		-0.089	
t-stat			*-3.14*				*-2.56*		**t-stat**		*-1.62*	
RSZ				-0.009				-0.008				
t-stat				*-3.58*				*-3.14*				

Table 5. Persistence of Complexity Effect

The table looks at the complexity effect in the five years after portfolio formation. Panel A performs cross-sectional regressions of DGTW-adjusted returns on Comp measure of complexity and the control variables from Tables 4. Comp is lagged by the number of years indicated in the first row, all other variables are lagged by one month/year.

Panel B looks at the FF5 alpha differential between single-segment firms with Comp=0 and conglomerates with Comp measure above/below median for the year (Z-H/Z-L), as well as the FF5 alpha differential between conglomerates with Comp measure above/below median (L-H). FF5 is the five-factor Fama-French (2015) model (with MKT, SMB, HML, and two new CMA (investment) and RMW (profitability) factors). The first row of Panel B reports the number of years by which Comp measure is lagged before forming the portfolios.

Panel C looks only at conglomerates (firms with two or more business segments with different two-digit SIC codes) and performs cross-sectional regressions of DGTW-adjusted returns on the HiComp dummy and control variables from Tables 5. HiComp is 1 if the Comp measure is above median, and 0 otherwise.

Panel D runs the regression of conglomerate returns on conglomerate age (CongAge), the number of consecutive years a firm is reported on Compustat segment files as having business segments with different two-digit SIC codes and then also interacts CongAge with three measures of complexity (1-HHI, NSeg, RSZ).

Detailed definitions of all variables are in Data Appendix. The t-statistics use Newey-West (1987) correction for heteroscedasticity and autocorrelation. The sample period is from January 1978 to December 2016. The sample excludes stocks priced below $5 on the portfolio formation date.

Panel A. Persistence of Complexity Effect: Cross-Sectional Regressions

	Year 1	**Year 2**	**Year 3**	**Year 4**	**Year 5**
Beta	0.115	0.118	0.118	0.110	0.084
t-stat	*1.14*	*1.24*	*1.21*	*1.07*	*0.76*
Size	0.000	0.000	0.000	0.000	0.000
t-stat	*-1.36*	*-1.90*	*-1.36*	*-1.61*	*-2.09*
MB	-0.028	-0.026	-0.027	-0.030	-0.026
t-stat	*-3.95*	*-3.75*	*-4.18*	*-4.45*	*-3.74*
Mom	-0.073	-0.085	-0.088	-0.098	-0.117
t-stat	*-1.44*	*-1.98*	*-1.93*	*-1.98*	*-2.28*
Rev	-0.047	-0.044	-0.041	-0.043	-0.042
t-stat	*-10.97*	*-11.31*	*-10.75*	*-11.06*	*-10.92*
Inv	-0.923	-0.920	-1.000	-1.059	-0.986
t-stat	*-4.00*	*-3.91*	*-4.15*	*-4.27*	*-3.83*
Prof	0.485	0.491	0.522	0.601	0.619
t-stat	*3.33*	*3.41*	*3.58*	*4.17*	*4.26*
Comp	-0.324	-0.352	-0.341	-0.369	-0.280
t-stat	*-3.44*	*-3.65*	*-3.40*	*-3.60*	*-2.68*

Panel B. Persistence of Complexity Effect: Portfolio Sorts

	Year 1	**Year 2**	**Year 3**	**Year 4**	**Year 5**
$\boldsymbol{\alpha_{FF5}^{Z-H}}$	0.354	0.275	0.329	0.335	0.299
t-stat	*4.01*	*2.78*	*3.42*	*3.32*	*3.10*
$\boldsymbol{\alpha_{FF5}^{Z-L}}$	0.203	0.149	0.283	0.282	0.284
t-stat	*2.46*	*1.76*	*3.64*	*3.58*	*3.59*
$\boldsymbol{\alpha_{FF5}^{L-H}}$	0.154	0.127	0.046	0.053	0.015
t-stat	*1.85*	*1.93*	*0.59*	*0.69*	*0.20*

Panel C. Persistence of Complexity Effect: Conglomerates Only Sample

	Year1	**Year2**	**Year3**	**Year4**	**Year5**
Beta	0.058	0.075	0.040	0.129	0.067
t-stat	*0.53*	*0.73*	*0.37*	*1.08*	*0.57*
Size	0.000	0.000	0.000	0.000	0.000
t-stat	*0.20*	*-0.85*	*-1.10*	*-0.22*	*-1.34*
MB	-0.045	-0.049	-0.033	-0.038	-0.018
t-stat	*-3.24*	*-3.47*	*-2.31*	*-2.67*	*-1.19*
Mom	-0.086	-0.023	-0.092	-0.158	-0.178
t-stat	*-0.97*	*-0.24*	*-0.94*	*-1.57*	*-1.74*
Rev	-0.049	-0.049	-0.047	-0.049	-0.044
t-stat	*-9.81*	*-8.24*	*-7.58*	*-8.59*	*-9.41*
Inv	-1.104	-0.856	-0.999	-0.806	-0.744
t-stat	*-3.91*	*-3.13*	*-3.50*	*-2.61*	*-2.55*
Prof	0.792	0.705	0.719	0.700	0.761
t-stat	*4.07*	*4.43*	*4.45*	*4.24*	*4.34*
HiComp	-0.086	-0.101	-0.080	-0.116	-0.078
t-stat	*-1.72*	*-1.92*	*-1.56*	*-2.20*	*-1.45*

Panel D. Conglomerate Age and Complexity Effect

	CongOnly	Var=NSeg	Var=Comp	Var=RSZ
Beta	0.075	0.079	0.078	0.085
t-stat	*0.69*	*0.83*	*0.82*	*0.89*
Size	-0.002	-0.002	-0.002	-0.001
t-stat	*-2.13*	*-2.62*	*-2.58*	*-2.49*
MB	-0.013	-0.007	-0.007	-0.008
t-stat	*-1.31*	*-2.61*	*-2.60*	*-2.06*
Mom	0.025	-0.039	-0.039	-0.030
t-stat	*0.31*	*-0.87*	*-0.86*	*-0.66*
Rev	-0.048	-0.044	-0.044	-0.043
t-stat	*-8.44*	*-11.31*	*-11.30*	*-11.28*
Inv	-0.616	-0.388	-0.385	-0.388
t-stat	*-2.99*	*-3.06*	*-3.03*	*-2.98*
Prof	0.678	0.421	0.419	0.417
t-stat	*3.67*	*3.10*	*3.08*	*3.06*
CongAge	0.111	-0.011	-0.098	-0.034
t-stat	*2.24*	*-0.37*	*-2.55*	*-1.29*
Var		-0.985	-0.240	-0.039
t-stat		*-3.55*	*-3.76*	*-4.04*
Var·CongAge		0.305	0.083	0.015
t-stat		*2.37*	*3.10*	*2.95*

Table 6. New Conglomerates and Segment Additions

The table repeats the firm-level cross-sectional regressions from the previous tables (with the same control variables and same Conglo and NSeg) with one additional variable added. NewCong is the dummy variable for a new conglomerate (NewCong1 is 1 if the conglomerate formed in the past year and 0 otherwise, NewCong2 is 1 if the conglomerate formed in the past two years and 0 otherwise, etc). SegInc is the dummy variable for the increase in the number of firm segments with different two-digit SIC codes (SegInc1 is 1 if the number of segments increased in the past year and 0 otherwise, SegInc2 is 1 if the number of segments increased in the past two years and 0 otherwise, etc). Detailed definitions of all variables are in Data Appendix. The t-statistics use Newey-West (1987) correction for heteroscedasticity and autocorrelation. The sample period is from January 1978 to December 2016. The sample excludes stocks priced below $5.

	1	**2**	**3**		**4**	**5**	**6**
Beta	0.092	0.095	0.095	**Beta**	0.090	0.094	0.093
t-stat	*1.06*	*1.08*	*1.02*	**t-stat**	*1.05*	*1.07*	*1.00*
Size	-0.006	-0.006	-0.005	**Size**	-0.005	-0.005	-0.004
t-stat	*-2.81*	*-2.75*	*-2.68*	**t-stat**	*-2.40*	*-2.34*	*-2.16*
MB	-0.007	-0.005	-0.008	**MB**	-0.007	-0.005	-0.009
t-stat	*-2.10*	*-1.61*	*-2.16*	**t-stat**	*-2.10*	*-1.50*	*-2.22*
Mom	-0.062	-0.076	-0.090	**Mom**	-0.063	-0.076	-0.089
t-stat	*-1.13*	*-1.34*	*-1.48*	**t-stat**	*-1.15*	*-1.34*	*-1.47*
Rev	-0.044	-0.042	-0.044	**Rev**	-0.044	-0.042	-0.044
t-stat	*-12.4*	*-11.8*	*-11.9*	**t-stat**	*-12.4*	*-11.8*	*-11.9*
Inv	-0.332	-0.499	-0.579	**Inv**	-0.337	-0.504	-0.583
t-stat	*-3.07*	*-4.06*	*-4.30*	**t-stat**	*-3.12*	*-4.11*	*-4.34*
Prof	0.462	0.482	0.561	**Prof**	0.459	0.481	0.559
t-stat	*3.96*	*4.08*	*4.64*	**t-stat**	*3.93*	*4.06*	*4.62*
Conglo	-0.096	-0.101	-0.101	**NSeg**	-0.049	-0.050	-0.049
t-stat	*-2.20*	*-2.16*	*-2.01*	**t-stat**	*-2.12*	*-2.02*	*-1.84*
NewCong1	-0.354			**SegInc1**	-0.181		
t-stat	*-2.45*			**t-stat**	*-1.85*		
NewCong2		-0.195		**SegInc2**		-0.098	
t-stat		*-1.91*		**t-stat**		*-1.30*	
NewCong3			-0.213	**SegInc3**			-0.103
t-stat			*-2.24*	**t-stat**			*-1.46*

Table 7. Diversification Discount, Conglomerate Age, and Firm Complexity

Panel A reports size-age-industry adjusted market-to-book (MB) and market-to-sales (MS) ratios in three years (-3 to -1) prior to firms becoming conglomerates in year 0. Panel B reports excess MB/MS ratios across conglomerate age (CongAge) groups. CongAge is the number of consecutive years a firm is reported on Compustat segment files as having business segments with different two-digit SIC codes. Excess MB/MS ratio of a conglomerate is the difference between MB/MS of a conglomerate and weighted average of imputed segment MB/MS (with total assets/sales used as weights). Imputed segment MB/MS is average MB/MS of single-segment firms in the same two-digit SIC industry that are within 10% of the conglomerate market cap and firm age (the number of years the firm appears on CRSP). Panel C reports difference in means tests for several pairs from Panels A and B. Panel D presents panel regressions of excess MS ratio on controls from Hund et al. (2021) and firm complexity: the first column includes all firms, the second one - only conglomerates, the remaining columns break down conglomerates into conglomerate age groups.

Panel A. Pre-Conglomeration

Time	**-3**	**-2**	**-1**	**0**
MBdisc	0.172	0.201	0.154	0.061
t-stat	*5.25*	*6.35*	*5.35*	*2.53*
MSdisc	0.320	0.420	0.409	0.346
t-stat	*7.29*	*9.80*	*10.6*	*9.53*

Panel B. Discount in Age Groups

Time	**1-5**	**6-10**	**11-15**	**16-20**	**21+**
MBdisc	-0.049	-0.058	-0.157	-0.289	-0.235
t-stat	*-5.50*	*-7.30*	*-12.7*	*-17.8*	*-14.0*
MSdisc	0.067	-0.038	-0.103	-0.210	-0.217
t-stat	*5.36*	*-3.58*	*-6.39*	*-9.92*	*-9.41*

Panel C. Differences

Time	**0 vs. -1**	**1-5 vs -1**	**1-5 vs. 0**	**21+ vs 1-5**
MBdisc	-0.093	-0.203	-0.110	-0.186
t-stat	*-2.45*	*-6.77*	*-4.28*	*-9.74*
MSdisc	-0.063	-0.342	-0.279	-0.284
t-stat	*-1.19*	*-8.50*	*-7.28*	*-10.8*

Panel D. Panel Regressions

	All Firms	All Conglos	1-5	6-10	11-15	16-20	21+
Intercept	0.027	-0.095	0.015	-0.135	-0.304	-0.501	-0.771
t-stat	*0.47*	*-1.92*	*0.32*	*-2.08*	*-4.19*	*-4.57*	*-5.92*
Age	-0.002	-0.002	-0.002	-0.001	0.000	0.001	0.000
t-stat	*-5.07*	*-3.81*	*-4.75*	*-1.97*	*-0.44*	*0.85*	*-0.53*
CapEx	0.011	0.203	0.155	0.190	0.217	-0.333	-0.280
t-stat	*0.30*	*2.67*	*2.39*	*1.23*	*0.75*	*-1.30*	*-1.10*
log(Sales)	-0.267	-0.246	-0.269	-0.195	-0.184	-0.216	-0.193
t-stat	*-15.9*	*-13.5*	*-13.7*	*-6.39*	*-4.50*	*-3.76*	*-3.17*
log(TA)	0.259	0.246	0.253	0.181	0.203	0.261	0.270
t-stat	*18.2*	*13.6*	*14.7*	*5.91*	*4.85*	*4.66*	*4.45*
Margin	0.172	0.228	0.106	1.013	0.834	1.011	0.591
t-stat	*5.48*	*1.58*	*1.18*	*4.48*	*1.91*	*2.33*	*1.58*
Comp	-0.085	0.006	0.109	0.099	-0.155	-0.274	-0.008
t-stat	*-2.46*	*0.15*	*2.89*	*1.99*	*-1.85*	*-2.07*	*-0.05*

Table 8. Complexity Effects and Limits to Arbitrage

The table reports five-factor Fama-French (2015) alphas in the double sorts on complexity and either residual institutional ownership, RInst (Panel A), or idiosyncratic volatility, IVol (Panel B), or probability to be on special, Special (Panel C), or the overpricing measure from Stambaugh et al. (2015), OverPrice (Panel D). The complexity groups (names in column headings) are the Low group, which includes all single-segment firms, and the Med/High groups, which include conglomerates with complexity below/above median. RInst/IVol groups (names in row headings) are top 30%, middle 40%, bottom 30%. The breakpoints for the RInst/IVol groups are determined using NYSE only firms. RInst is institutional ownership orthogonalized to size. Detailed definitions of all variables are in Data Appendix. The t-statistics use Newey-West (1987) correction for heteroscedasticity and autocorrelation. The sample period is from January 1978 to December 2016. The sample excludes stocks priced below $5 on the portfolio formation date.

Panel A. Comp/RInst Sorts

	Zero	**Low**	**High**	**Z-H**
Low	0.291	-0.296	-0.296	0.587
t-stat	*2.84*	*-1.81*	*-3.58*	*5.01*
Med	0.193	-0.205	-0.096	0.290
t-stat	*2.64*	*-2.05*	*-0.88*	*2.55*
High	0.079	0.026	-0.043	0.122
t-stat	*0.90*	*0.24*	*-0.42*	*1.02*
L-H	-0.212	0.320	0.253	0.465
t-stat	*-1.73*	*1.72*	*1.86*	*3.16*

Panel B. Comp/IVol Sorts

	Zero	**Low**	**High**	**Z-H**
Low	0.112	-0.154	-0.084	0.196
t-stat	*1.32*	*-1.81*	*-1.26*	*1.77*
Med	0.140	-0.062	-0.096	0.236
t-stat	*1.78*	*-0.57*	*-0.94*	*1.82*
High	-0.272	-0.676	-0.736	0.463
t-stat	*-1.77*	*-2.30*	*-3.53*	*2.19*
H-L	0.384	0.523	0.652	0.267
t-stat	*2.08*	*1.59*	*3.01*	*1.14*

Panel C. Comp/Special Sorts

	Zero	**Low**	**High**	**Z-H**
Low	0.173	-0.211	-0.029	0.202
t-stat	*2.29*	*-2.73*	*-0.36*	*1.75*
Med	0.141	0.054	-0.269	0.410
t-stat	*1.56*	*0.43*	*-2.01*	*2.95*
High	0.092	-0.365	-0.558	0.650
t-stat	*0.62*	*-2.44*	*-3.89*	*3.26*
H-L	0.081	0.154	0.529	0.448
t-stat	*0.56*	*0.88*	*3.56*	*2.19*

Panel D. Comp/OverPrice Sorts

	Zero	**Low**	**High**	**Z-H**
Low	0.174	0.037	0.064	0.111
t-stat	*2.06*	*0.39*	*0.68*	*0.74*
Med	0.350	-0.095	-0.191	0.541
t-stat	*4.13*	*-0.78*	*-2.19*	*4.48*
High	-0.014	-0.455	-0.346	0.333
t-stat	*-0.08*	*-2.87*	*-2.26*	*1.93*
H-L	0.188	0.492	0.410	0.222
t-stat	*0.96*	*2.28*	*2.74*	*1.02*

Table 9. Complexity Effects and Earnings Announcements

The table regresses earnings announcement returns on complexity measures and standard asset-pricing controls from Tables 4-6. Earnings announcement returns are cumulative returns in the three days around the announcement. The complexity measures include the conglomerate dummy (Conglo, 1 for conglomerates, zero for single-segment firms), the Comp variable, defined as the sales concentration across segments, 1-HHI, where HHI is the sum of squared shares of segment sales in total firm revenue, the number of business segments with different two-digit SIC codes (NSeg), and the RSZ complexity measure, which is the standard deviation of segment asset-weighted market-to-books for the conglomerate divided by the equal-weighted average market-to-book of the segments. Conglomerates are defined as firms with business segments in two or more industries with different two-digit SIC codes. Detailed definitions of all variables are in Data Appendix. The t-statistics use Newey-West (1987) correction for heteroscedasticity and autocorrelation. The sample period is from January 1978 to December 2016. The sample excludes stocks priced below $5.

	1	**2**	**3**	**4**
Beta	-0.211	-0.213	-0.213	-0.220
t-stat	*-5.65*	*-5.68*	*-5.71*	*-5.74*
Size	0.000	0.000	0.000	0.000
t-stat	*-2.60*	*-2.64*	*-2.30*	*-2.67*
MB	-0.071	-0.071	-0.072	-0.062
t-stat	*-6.77*	*-6.72*	*-6.80*	*-5.55*
Mom	0.051	0.053	0.052	0.049
t-stat	*0.85*	*0.88*	*0.87*	*0.79*
Rev	0.006	0.006	0.006	0.006
t-stat	*4.01*	*4.02*	*4.03*	*4.13*
Inv	-0.383	-0.377	-0.385	-0.348
t-stat	*-2.36*	*-2.31*	*-2.36*	*-2.05*
Prof	0.597	0.599	0.587	0.594
t-stat	*6.60*	*6.58*	*6.40*	*6.33*
Conglo	-0.086			
t-stat	*-3.00*			
Complexity		-0.223		
t-stat		*-3.03*		
NSeg			-0.062	
t-stat			*-3.99*	
RSZ				-0.018
t-stat				*-4.55*

Table 10. Complexity Effects and the Business Cycle

Panel A presents the estimates from regressions that predict the risk premium to portfolios named in the headings of the columns (see Table 3 for description of the portfolios). The predictive variables include default premium (DEF), dividend yield of the market portfolio (DY), 1-month Treasury bill rate (TB), term premium (TERM), VIX index, and TED spread. Panel B presents estimates from regressions of returns to the same portfolios on the market return and innovations to the variables above. The innovations are residuals from ARMA(1,1) models fitted separately to each of the variables. Detailed definitions of all variables are in Data Appendix. The t-statistics use Newey-West (1987) correction for heteroscedasticity and autocorrelation. The sample period is from January 1986 to December 2016. The sample excludes stocks priced below $5 on the portfolio formation date.

Panel A. Predictive Regressions

	Zero	Low	High	Z-H	Z-L	L-H
Constant	-1.349	-1.122	-0.845	-0.500	-0.227	-0.263
t-stat	*-1.09*	*-1.05*	*-0.77*	*-0.78*	*-0.44*	*-0.59*
$\boldsymbol{DEF_{t-1}}$	-1.768	-1.718	-1.469	-0.301	-0.051	-0.256
t-stat	*-1.76*	*-1.85*	*-1.51*	*-0.68*	*-0.13*	*-0.93*
$\boldsymbol{DY_{t-1}}$	2.261	2.049	1.893	0.369	0.212	0.161
t-stat	*3.41*	*3.99*	*3.67*	*1.04*	*0.68*	*0.88*
$\boldsymbol{TB_{t-1}}$	-6.244	-4.577	-4.318	-1.939	-1.667	-0.302
t-stat	*-2.43*	*-2.26*	*-2.00*	*-1.51*	*-1.44*	*-0.44*
$\boldsymbol{TERM_{t-1}}$	-1.109	-0.872	-1.028	-0.082	-0.237	0.152
t-stat	*-2.94*	*-2.52*	*-2.84*	*-0.50*	*-1.60*	*1.32*
$\boldsymbol{VIX_{t-1}}$	0.162	0.134	0.139	0.023	0.027	-0.005
t-stat	*3.71*	*3.65*	*3.79*	*0.98*	*1.05*	*-0.27*
$\boldsymbol{TED_{t-1}}$	-2.464	-2.625	-2.554	0.090	0.160	-0.068
t-stat	*-1.55*	*-1.87*	*-1.91*	*0.17*	*0.34*	*-0.21*

Panel B. Exposure to Macroeconomic News

	Zero	Low	High	Z-H	Z-L	L-H
Constant	-0.083	0.016	0.111	-0.193	-0.099	-0.093
t-stat	*-0.96*	*0.27*	*1.52*	*-1.41*	*-0.84*	*-1.01*
$\boldsymbol{\beta_{MKT}}$	1.128	0.885	0.951	0.177	0.243	-0.066
t-stat	*24.6*	*26.8*	*26.1*	*2.72*	*3.84*	*-2.30*
$\boldsymbol{\Delta DEF_t}$	-1.057	0.271	0.119	-1.186	-1.328	0.132
t-stat	*-1.35*	*0.23*	*0.16*	*-0.93*	*-0.85*	*0.11*
$\boldsymbol{\Delta DY_t}$	1.024	-2.657	1.174	-0.156	3.681	-3.846
t-stat	*0.67*	*-2.11*	*0.88*	*-0.06*	*1.62*	*-2.46*
$\boldsymbol{\Delta TB_t}$	-2.052	0.306	-1.242	-0.798	-2.358	1.574
t-stat	*-0.71*	*0.13*	*-0.47*	*-0.16*	*-0.52*	*0.66*
$\boldsymbol{\Delta TERM_t}$	-0.019	-0.602	0.427	-0.442	0.582	-1.021
t-stat	*-0.03*	*-1.05*	*0.75*	*-0.44*	*0.64*	*-1.60*
$\boldsymbol{\Delta VIX_t}$	0.041	-0.053	-0.056	0.097	0.094	0.003
t-stat	*1.19*	*-1.90*	*-1.98*	*1.87*	*2.39*	*0.09*
$\boldsymbol{\Delta TED_t}$	0.195	-1.058	0.308	-0.112	1.253	-1.364
t-stat	*0.46*	*-2.56*	*0.80*	*-0.17*	*2.36*	*-2.90*