ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics

Master Thesis Financial Economics

The complexity of firms and information processing: its effect on return predictability

Dirk Maarten van Duijvenbode

30 October 2023

Student Number: 533097

Supervisor: Hauwe, S van den

Second assessor: TBA

Abstract

This study explores the complications of information diffusion across US firms from 1977-2021, by looking specifically at standalone firms (easy-to-analyse) and conglomerates (complex). I re-examine the premise that return predictability is affected by firm complexity. A long-short strategy yields a positive and significant excess return of 60 basis points per month before transaction costs. This study also finds that this effect is more pronounced for smaller firms and that analysts also suffer from information processing constraints. While the effects of firm complexity on return predictability still persists, their magnitude and significance have transformed over the years. Based on the sub-period analysis, it seems that this effect seems to decrease from the 2000's onwards, even showing negative returns during 2000-2010.

Contents

1. Introduction

One of the foundational beliefs in finance is the Efficient Market Hypothesis first worked out by Fama in 1970. This theory states that security prices should, at all times, reflect all available information. Yet, many studies document that there have been deviations from this ideal. These anomalies have started an extensive literature on what factors can be of influence when trying to determine return predictability. The ability to understand these frictions can lead to a more comprehensive understanding on how information diffuses across markets and how investors process this information. Merton (1987) was among the first to address the issue as to why news and information might be incorporated later or not completely. He stated that investors, despite their best intentions, might not be cognitively equipped to process all available information for each security in the market. This limitation together with the natural inclination towards attention grabbing stocks found by Barber and Odean (2008) and information immobility (Van Nieuwerburgh and Veldkamp, 2009), meaning that investors tend to prefer local over foreign equities (home bias), has led to interesting patterns among stock returns.

An interesting area of research is the study of information processing in firms of varying complexity. Presumably, it is easier for investors to analyse information regarding a firm operating in a single industry segment than for a firm with operations in multiple industries. Frankel, Kothari, and Weber (2006) found that analyst reports are less informative when the processing costs of the information are high, particularly in cases in which a firm has multiple business segments. Cohen and Lou (2012) elaborate further on the difference in information diffusion between conglomerate firms and standalone firms and their effect on the stock price. This paper will largely follow the approach they took. This means that conglomerates are defined as firms who operate in more than one industry, while standalone firms operate in only one industry. It is expected that because of investors' limited attention and processing capacity, a delay in the updating of prices to include industry wide information is due to the complexity of processing information in more segmented firms.

A decade has passed since their paper was published and since then multiple studies have suggested the questionability of replicating a paper. Marquering, Nisser and Valla (2006) suggest that anomalies are gradually disappearing over time with a trend towards zero. McLean and Pontiff (2016) suggest that investors learn about the mispricing stated in academic papers and estimate a decline of 32% in returns post-publication. Hou, Xue and Zhang (2020) in their more recent paper state that around 50% of the anomalies found in the literature fail the 10% significance level when replicated, deeming the results insignificant. If the results can be replicated, the economic magnitudes are much more modest than previously found.

Considering these developments, this paper revisits the concept of return predictability through the channel of firm complexity, covering US stocks listed on the major exchanges from 1977-2021. This timeframe is picked as companies are required under the Statement of Financial Accounting Standards to publish segment data starting in 1976 and extends up until now. To ensure that segment information is publicly known, a six-month gap is imposed after the fiscal year end. Then the segment data is merged with CRSP monthly stock files in order to run the stock-return tests.

The choice for monthly data is embedded in the complexity of conglomerates. While daily data can be more precise, it can also be too volatile and therefore not fully capture the gradual adjustment process of stock prices due to new information. A lower frequency such as quarterly data might overly smooth out the nuances in stock price reactions to new information. Furthermore, investors who do not have the computational capabilities of institutional investors might need a few weeks to comprehend new information regarding complex firms. By using monthly data, the findings reflect a more realistic scenario of how investors respond to new data about complex firms.

The overarching question that is examined through this paper is:

Does return predictability due to variations in firm complexity and information processing persist through time?

Firstly, pseudo-conglomerates will be paired to conglomerates based on the returns of standalone firms operating in the same industry as the conglomerate. These conglomerates are then placed into decile portfolios based on the past return from the pseudo-conglomerate. Excess returns are calculated from the return of the portfolio and the risk-free rate. I do this each month. The portfolios are rebalanced monthly to ensure equal/value weights, following Cohen and Lou (2012). Then a long-short strategy is implemented, meaning a hypothetical situation in which I go long in the winner portfolio while shorting the loser portfolio. This strategy can yield up to 7.44% per year. Next to this complicated processing portfolios strategy, I also implement regression tests to define the relationship between the lagged pseudo-conglomerate returns and the returns from the conglomerate itself. With this analysis, I find that a one standard deviation increase in pseudo-conglomerate returns last month leads to a 54-basis point increase in this month’s paired conglomerate return. These results do not include transaction costs.

In addition, the level of complexity per firm and the information embedded in analyst forecasts are also examined. This study reveals that a firm's complexity positively influences return predictability. In other words, returns of a highly segmented conglomerate are better predicted using its corresponding pseudo-conglomerate return. This result, however, is only significant at the 13% level. I do find that the predictability is greater for smaller firms. Next, for the analyst forecasting, this study finds that there is a small predictability of forecast revisions for the conglomerate based on its paired pseudo-conglomerate forecast revision, indicating that analysts have similar information processing constraints as investors. Finally, the Fama and French (2015) five factor model is used as another specification to check for robustness. By using the profitability and investment factor, the excess returns found by long-short strategy remain unchanged.

The sample is also divided into four sub-periods to see whether the effect is present throughout the time period or just in certain periods. The sub-periods I choose are 1977-1988, 1989-1999, 2000-2010, and 2011-2021. This segmentation has two incentives: ensuring an even division of the data and examining a timeframe that was not yet explored by Cohen and Lou (2012). Additionally, by segmenting the analysis also around two major global events, namely the Dot Com bubble and the Global Financial Crisis, this study seeks to understand the influence of such significant disruptions on information diffusion. This study finds that the effect largely decreases from the 2000's onwards. Reasons for this might be the economic turmoil caused by the two aforementioned crises and the decline in post-publication predictability. By researching this, I not only build upon prior research but also offer fresh insights into the evolving financial markets and the effects of academic research on those markets.

2. Literature review

2.1 The Capital Asset Pricing Model

Asset pricing is an integral part of financial economics and has been a widely researched topic in finance for decades. At its core, it revolves around the process of determining the fair value of an investable asset. Researchers have been trying to determine common underlying risk factors to try and explain asset returns. In the 1960s, the field of financial economics experienced a scientific leap forward with the start of modern portfolio theory by Harry Markowitz (1959). Following his work, other economists made significant contributions to Markowitz's work including Sharpe (1964), Lintner (1965) and Mossin (1966). Even though these economists worked independently, they arrived at similar conclusions, therefore laying the cornerstone of the Capital Asset Pricing Model (CAPM). This model was deemed revolutionary as it introduced a method to quantify the relationship between expected risk and return of an asset in a systematic manner.

The following formula expresses the CAPM:

$$E(r_i) = r_f + \beta_i\big(E(r_m) - r_f\big)$$

Where:	$E(r_i)$	=	Expected return on asset i
	r_f	=	The risk-free rate
	β_i	=	The sensitivity of asset i's returns to the returns of the market as a whole; calculated as: $\beta_i = \frac{cov(r_i,r_m)}{var(r_m)}$
	$E(r_m)$	=	Expected return of the market itself

The formula shows that the expected return of an asset relies on three components: the risk-free rate, the asset's beta (its sensitivity to market movements), and the market risk premium (the difference between market and risk-free returns). Essentially, investors will demand a higher return for taking on more risk.

However, while CAPM was groundbreaking in linking risk to returns, it relies on several assumptions such as perfect markets, unlimited borrowing and lending at the risk-free rate, and rational investors. In practice, these assumptions do not always hold. Anomalies – situations in which investors are able to systematically achieve excess returns over the market – are notable limitations of the CAPM model. Fama and French (1996) showed that some CAPM anomalies vanish when adding factors like 'size' and 'value' to the CAPM model.

In the context of this study, firm complexity and the complications in information processing represent additional layers that can affect return predictability. Traditional models like the CAPM might not perfectly account for the implications introduced by firm complexity. The challenge of incorporating all relevant and available information into asset prices might lead to greater return predictability as firms become more complex, something not explicitly covered by CAPM. This study seeks to examine this aspect, potentially adding to the understanding of risk and return provided by CAPM.

The next section will look closer at the Fama and French (1996) three-factor model and extensions of this model.

2.2 Extensions of the CAPM

2.2.1 Fama and French three factor model

Following the inception of the CAPM in the 1960s, considerable research has been dedicated to examining the valuation of supplementary risk factors. Fama and French (1992) found that two simply measured variables, size, and book-to-market equity, appear to illustrate the cross-section of average stock returns based on non-financial firms listed on the NYSE, AMEX, and NASDAQ between 1962 and 1989. Fama and French (1993, 1995) confirm these results. They find that the book-to-market equity and the size of a firm can be considered proxies for common risk factors based on their US sample. The value effect refers to the long-term tendency of companies with high book-to-market ratios to outperform those with low book-to-market ratios. Similarly, the size effect suggests that smaller cap companies outperform larger cap companies over the long term.

Fama and French (1995) try to find if the size and value factor are actual risk factors. Their work was guided by two hypotheses: If the relation between average stock return and the proposed size and value factor are due to rational pricing then (i) there must be shared risk factors in returns related to size and book-to-market ratios, and (ii) the patterns of returns based on size and book-to-market ratios must be elucidated by the earnings behaviour. They determined that the earnings of companies exhibit size and book-to-market factors similar to those observed in returns. The market and size factors in earnings help explain the corresponding factors in returns. However, they found no evidence supporting a relationship between the value factor in earnings and returns. The following regression shows the additional factors incorporated in the CAPM, also known as the Fama and French three-factor model:

$$\mathrm{r}_{i,t} - \mathrm{r}_{f,t} = \alpha_i + \beta_i\left(\mathrm{r}_{m,t} - r_{f,t}\right) + s_i\mathrm{SMB}_t + h_i\mathrm{HML}_t + \epsilon_{i,t}$$

Where: β, s, h	=	Sensitivity of stock's excess return to the respective risk factor
$\mathrm{r}_{m,t} - \mathrm{r}_{i,t}$	=	Market risk premium
SMB	=	Excess returns of small-cap companies over big-cap companies
HML	=	Excess returns of value stocks over growth stocks

The size factor (SMB) is based on the observation that smaller firms tend to have higher average returns than larger firms, even after adjusting for market risk. The value factor (HML) relates to the tendency for shares in high-market-value companies (value stocks) to outperform shares in low-market-value companies (growth stocks). Both factors are based on historical observations. With the stated regression above in mind, the excess return of a stock is based on its sensitivity to movements in the market as a whole, the size and value. The inclusion of these factors addresses some of the empirical inconsistencies of the CAPM. For instance, the size and value factors help explain why in reality, we sometimes observe that higher-risk (high-beta) stocks do not necessarily generate higher returns, as predicted by the CAPM.

2.2.2 Four factor models

In 1993, Jegadeesh and Titman were among the first to find return continuation also known as momentum in traditional assets. In their paper, they held the belief that if prices exhibited overreactions or underreactions to the available information, it would be possible to generate profits by trading stocks based on their past returns. They would rank stocks in deciles based on their performance in a previously specified period. Then a strategy of going long in the best past performers and shorting the worst past performers was implemented, which resulted in positive excess returns.

Using this research as his inspiration, Carhart (1997) looked at return continuation with regards to mutual fund returns. He found that buying last year's top-decile mutual fund and shorting last year's worst decile mutual fund resulted in a return of 8% per year. His work proposes three important guidelines for mutual fund investors aiming to maximise their wealth: (1) Avert funds with steadily poor performance; (2) funds that achieved high returns in the previous year are expected to have above-average returns in the following year, but not in subsequent years; and (3) investment costs such as expense ratios, transaction costs, and load fees directly and negatively affect performance.

The Carhart model is an extension of the Fama and French three factor model by including a cross-sectional momentum factor that strengthens the explanatory power of the three-factor model. Therefore, the model presented in the following regression resembles the previous three-factor regression:

$$\mathrm{r}_{i,t} - \mathrm{r}_{f,t} = \alpha_i + \beta_i(\mathrm{r}_{m,t} - r_{f,t}) + s_i\mathrm{SMB}_t + h_i\mathrm{HML}_t + u_i\mathrm{UMD}_t + \epsilon_{i,t}$$

Where: UMD = Excess returns of past high performing firms over past low performing firms

Another factor that has shown importance in financial markets is liquidity, as it denotes the ability to trade large quantities of an asset quickly, at low cost, and without significantly affecting the asset's price. Research done by Amihud and Mendelson (1986) and Brennen, Chordia and Subrahmanyam (1998) found that less liquid stocks have higher average returns, implying a certain liquidity risk premium offered to investors to compensate for the higher costs and/or risks associated with trading less liquid stocks. The work by Chordia, Roll and Subrahmanyam (2000, 2001) acknowledged the potential usefulness of market-wide liquidity as a state variable affecting expected stock returns.

This is similar to the work of Pastor and Stambaugh (2003). They argue that liquidity could be a state variable that affects expected stock returns. The reason for this is because it could have pervasive effects on an investor's overall welfare. Namely, a security whose returns drop when liquidity conditions worsen would need higher expected returns to compensate for this risk. The authors of this paper found that the sensitivity of an asset's returns to fluctuations in aggregate market liquidity (liquidity beta) is indeed an important factor in asset pricing. Stocks with higher liquidity betas, meaning that they are more sensitive to changes in overall market liquidity, tend to have higher expected returns. Between 1966 and 1999, a difference between the top and bottom deciles in predicted liquidity betas yielded an excess return of 9 percent per year with respect to the Fama and French three factor model and an excess return of 7.5 percent per year after accounting for the Carhart four factor model.

The regression of this model would be as follows:

$$\mathrm{r}_{i,t} - \mathrm{r}_{f,t} = \alpha_i + \beta_i(\mathrm{r}_{m,t} - r_{f,t}) + s_i\mathrm{SMB}_t + h_i\mathrm{HML}_t + l_i\mathrm{LIQ}_t + \epsilon_{i,t}$$

Where: LIQ = The product of absolute return and dollar volume of trading

2.2.3 Fama and French five factor model

The Fama and French three-factor model marked a significant advancement in our understanding of asset pricing, but the assumptions and limitations discussed before showed it was not yet sufficient. Subsequent empirical observations suggested that the three-factor model was not capturing all relevant risks affecting the cross-section of average stock returns. In the Fama and French paper from 2015, there was an attempt to address the limitations of their previous model by adding two new factors to their model, namely profitability and investment: creating the Fama and French five factor model. The model can be quantified as follows:

$$\mathrm{r}_{i,t} - \mathrm{r}_{f,t} = \alpha_i + \beta_i(\mathrm{r}_{m,t} - r_{f,t}) + s_i\mathrm{SMB}_t + h_i\mathrm{HML}_t + r_iRMW_t + c_i\mathrm{CMA}_t + \epsilon_{i,t}$$

Where: RMW = The difference between the returns of firms with robust and weak operating profitability

CMA = The difference between the returns of firms that invest conservatively and aggressively

The profitability factor (RMW) was introduced based on the observation that firms with higher profits have tended to outperform those with lower profits, while controlling for other factors. This aligns with economic theory, which suggests that companies that are more profitable are more likely to produce higher returns for investors. The investment factor (CMA) addresses the empirical anomaly that firms investing heavily in new capital have historically underperformed. Firms that are aggressive investors may be taking on higher levels of risk, for instance, by funding projects that may not be profitable in the future. Conversely, conservative companies may be seen as less risky and thus command lower expected returns.

The Fama and French five factor model offers a more comprehensive explanation of asset pricing, capturing a wider array of risks that are priced in the market. With this additional information, it is easier to understand why certain stocks seem to consistently outperform others. However, as with any model, the five-factor model is still a simplification of reality. It does not capture all sources of risk and only looks at risk factors while leaving out the ever-growing field of behavioural finance. Despite these limitations, the model is an important tool in the toolbox of modern finance, helping us to understand and quantify the sources of risk in asset prices.

2.3 Information Processing in Financial Markets

2.3.1 Effect of information on stock prices

Information forms the basis upon which investors make decisions about buying or selling assets such as stocks. The timeliness, accuracy and availability of certain information can therefore significantly impact stock prices and the decisions of investors. In 1970, Fama wrote the paper "Efficient Capital Markets: A Review of Theory and Empirical Work" and with it introduced the concept Efficient Market Hypothesis (EMH). EMH is crucial as is postulates that all available information is already incorporated into asset prices. This is what he considered as the Strong Form and called the market therefore efficient. There is also the weak form and the semi-strong form, in which prices are solely based on historical prices and only based on all available public information respectively. From the EMH stem some important implications. Namely, if markets are efficient, it implies that outperforming the market through stock selection or market timing is unlikely, it implies that stock prices are an accurate reflection of a firm's underlying value, facilitating an efficient allocation of capital and lastly it implies a decreased use for regulation aimed at preventing the use of insider information.

However, there is also a large body of literature challenging the EMH. Researchers such as Shiller (1981) and Thaler (1999) introduced psychological factors into financial models, indicating that investors do not always behave rationally, and markets are not always efficient. Additionally, there are market anomalies which can affect stock prices. Under- and overreaction to news has been found by Barberis, Shleifer and Vishny (1998). Their model of investor sentiment, of how investors form their beliefs, is based on psychological evidence and they find underreaction to earnings announcements and overreaction to a series of good or bad news.

There has also been evidence of historical prices predicting the performance of a stock in the future, which is known as the momentum effect. Jegadeesh and Titman (1993) found that buying a decile portfolio of the best performers and selling a decile portfolio of the worst performers, earned them an average positive excess return of 1%. Lastly, the phenomenon of Post-Earnings Announcement Drift (PEAD) has been established and discussed by Bernard and Thomas (1989). PEAD refers to the tendency of stocks to continue drifting in the same directions as an earnings surprise for some time following the announcement. Explanations include investor underreaction or behavioural biases such as conservatism bias or confirmation bias.

2.3.2 Information complexity in more complex firms

As discussed above, the real-world application of the EMH often bumps into practical complications. The intricate structure and operations of certain firms can be one of these practical complications. Unlike simpler firms that operate for instance in a singular industry or single geographic location, more complex firms can operate in multiple sectors, international markets and may even have multi-layered organisational structures. The complexity is not inherently negative, it can arise from the need to diversify certain business risks, enter new markets, or innovation. However, with the increased complexity, new challenges to how information is processed are introduced. The information that is perceived form such complex firms can be tough to interpret potentially leading to inefficiencies in how stock prices reflect this information.

Ali and Hirshleifer (2020) highlight that the complexity of a firm's linkages, measured by the number of connections a firm has to other firms, can strengthen momentum spillovers, due to the increased cognitive processing required to update information. This idea is supported by Dong, Li, Lin, and Ni (2016), who argue that high information-processing costs are more pronounced in complex firms with opaque financial reporting. Additionally, Blankespoor, deHaan and Marinovic (2020) emphasize that when the costs of acquiring, monitoring, and analysing firm disclosures outweigh expected trading gains, investors might disregard the information, resulting in underpricing—especially within the framework of the semi-strong EMH. I expect to find a similar result as these papers. Namely, the spillovers mentioned as conglomerates are linked to their respective stand-alone firms based on similar industries.

Adding to this complexity, Miller (2010) finds that more complex financial reporting correlates with lower trading volumes, disproportionately affecting small investors. Further studies by You and Zhang (2009) and Dolde and Mishra (2002) add another layer of complexity. They illustrate how investors' reactions to 10-K filings are sluggish, especially for firms with complex reports, and that such firms are more likely to manage foreign exchange exposures due to their inherent complexity. Collectively, these studies suggest that while complexity may arise from strategic needs like diversification and innovation, it introduces new challenges to how information is processed, potentially leading to inefficiencies that depart from the idealized theories posited by EMH.

However, there have been advancements in financial technology designed to decrease information complexity. This digital evolution in financial reporting has also been studied for its impact on information processing. Research by Kim, Li, and Liu (2019) and Huang, Shan, and Yang (2021) examine the mandate of eXtensible Business Reporting Language (XBRL). This is a standardised language for digitally communicating business and financial data. It allows for an automated exchange of financial information between organisations, regulators, and analysts which in turn significantly reduces information-processing costs. Kim et al. (2019) found that the XBRL mandate led to an increase in the number of shareholders for a firm, while Huang et al. (2021) found that XBRL speeds up the information incorporation process and facilitates the market in learning about younger firms. As the adoption of XBRL started in 2009, I expect that the predictability due to firm complexity and complicated information processing will be less from this time onwards.

2.3.3 Investor inattention

There is also the phenomenon called investor inattention, which is an aspect of the limitations in human information processing. Investor inattention has significant implications for asset pricing due to certain cognitive biases, the use of heuristics and memory constraints. Barberis (2018) made a seminal contribution to this area by critiquing the traditional financial models that assume investors can instantly and perfectly process any new information that hits the market. This assumption seems unrealistic with the real-world limitations of human cognition. Barberis cites the phenomenon of inattention, where stock prices initially underreact to good earnings news and only adjust fully after a lag as more investors come to know and act on the new information. Empirical support for this is provided by DellaVigna and Pollet (2009), who find that post-earnings announcement drift is more significant when firms announce their earnings on a Friday—a day when investors are thought to be less attentive. Blankespoor et al. (2020) continue this by discussing the very real costs involved in processing firm disclosures. Their review suggests that even as technology advances, the cost of processing complex financial disclosures continues to be a barrier for investors. These costs effectively turn firm disclosures into a form of private information, thereby creating a lag in information incorporation into asset prices.

The concept of "equilibrium of disequilibrium" presented by Dong et al. (2016) offers another angle. They argue that the high cost of information processing affects share price informativeness and leads to an underinvestment in expensive firm-specific information. In line with this, Veldkamp (2006) suggests that the purchase of low-price common information can lead to asset price comovement, further highlighting the consequences of information processing limitations. Duffie (2010) also delves into the behavioural implications of limited attention span in investment decisions. He notes that asset prices at any given time are likely determined by a small subset of investors who are actively trading, reflecting the real-world situation where many investors are not continually focused on the market. Duffie's observations are supported by empirical evidence indicating that most investors adjust their portfolios remarkably infrequently, which could be attributed to cognitive limitations or the high costs—both in time and effort—of keeping up with the constantly changing financial markets. Based on this, I expect that the turnover of a stock, which implies the amount of attention from investors given to a stock, is negatively correlated with the return predictability as investors are actively monitoring the company and therefore the information it discloses.

2.3.4 Analyst coverage

Financial analysts play a complex role in financial markets. Not only do they distribute information on certain stocks, but their activities and choices can have a predictive value in asset pricing and can in turn influence the perceived firm fundamentals. According to Lee and So (2017), analyst coverage contains invaluable insights into expected returns. They show that firms with abnormally high analyst coverage outperform those with abnormally low coverage by approximately 80 basis points per month. They suggest that the standard security analyst is trained in delivering information to the market, incurs significant costs when deciding to change which companies they cover, and benefits greatly from identifying stocks with higher growth potential. Considering their motivations and their relative sophisticated understanding of company outlooks, they argue that the analysts' decisions on which companies to focus on offer valuable insights for predicting the future performance of those firms. This view finds historical support in works by McNichols and O'Brien (1997), Scherbina (2008), and Das, Guo, and Zhang (2006).

Doukas, Kim and Pantzalis (2005) investigate the effects of analyst coverage as well. They find that excessive analyst coverage can lead to overvaluation and subsequently to lower future returns. They point out that this can be attributed to conflicts of interest, such as investment banking incentives, that may misalign analyst activity with accurate firm valuation. It seems to be different in emerging markets. Chan and Hameed (2006) provide an interesting insight by investigating those emerging markets. Contrary to the belief that analysts focus on firm-specific information, they find that greater analyst coverage leads to higher stock price synchronicity. This means that firms with more analyst coverage tend to have stock prices that move more in line with market trends, rather than based on firm-specific news.

2.4 Return predictability and information processing

2.4.1 Firm linkages

The financial market is a complex web of interrelated entities in which the performance of a business not only has effect on itself but can have detrimental effects on its linked counterparts. Whether it is the similarity of the industry, strategic alliances or social ties, the assumption is that information moves across these links, albeit a slow information diffusion. Cohen and Lou's (2012) paper is the basis for this literature review. In their paper, they introduced the idea of return predictability due to the complexity of information processing. They use so called 'stand-alone firms' and 'conglomerates' to categorise firms into easy-to-analyse and complex-to-analyse. By looking at how straightforwardly the information impacts their stock prices, they found a significant return predictability of 118 basis points per month. Additionally, they showed that the more complicated the firms was, the clearer the return predictability was. Even sell-side analysts are not immune to the constraints of complex information processing, as their forecast revisions of less complex firms predict their revisions of more complex firms in the future.

Previous research by Cohen and Frazzini (2008) finds confirmation of return predictability across economically linked firms. In their paper they focus on 'well-defined customer-supplier links' between firms and investigate how shocks to one firm affect the linked firm. They examine the theory that when investors are subject to attention constraints, stock prices are slow to reflect news about companies that are economically linked. This in turn can lead to predictable returns. A study by Cao, Chordia, and Lin (2016) comes to a similar conclusion. They examine return predictability across alliance partners and find that a long-short portfolio based on lagged returns of those partners earns a return of 89 basis points per month. They believe, in accordance with Cohen and Frazzini (2008), that investor inattention and limits to arbitrage may be the source of this underreaction between alliance partners. Lastly, a rather new paper by Barinov, Park and Yildizhan (2022) builds on the idea of conglomerates and single-segment firms. They show that the post-earnings announcement drift (PEAD) is greater for conglomerates than for single-segment firms. As conglomerates are more challenging to understand firm-specific information, the information processing is slowed down. Resulting in a longer over- or underreaction from investors on an earnings announcement.

2.4.2 Alternative linkages

More recent papers also tend to look at other linkages between firms. Lee, Sun, Wang, and Zhang (2019) find return predictability among technology-linked firms. The so-called focal firms whose tech-peers gain higher returns will in turn earn themselves a higher return in the following months. They believe that their findings better align with the idea that stock prices are slow to adapt to subtle news impacting closely related companies in the tech sector. Furthermore, research done by Müller (2019) examines if there is a similar pattern of gradual information diffusion across stocks when portfolios are formed based on anomalies in the literature such as size, value, and asset growth. He found that when a company reports earnings surprises it can serve as a signal for what might happen to other companies that share the same characteristics. Surprisingly, the usual factors that might explain these patterns such as industry trends or risk factors did not account for what he observed.

Scherbina and Schlusche (2013) looked at linked stocks through co-mentions in news stories, resulting in the possibility to cross-predict one another's future returns. This information flow is not only in the same industry but can go from small to large stocks and even across different industries. Their results indicate that both limited attention and the processing of complicated information are causing this gradual information diffusion. Even social ties between firms can affect their stock returns and fundamentals like earnings. Peng, Titman, Yönac and Zhou (2022) researched this idea and found that firms in the same industry located in socially connected areas tend to move in sync, albeit it not instantaneously. They would form portfolios that buy stocks of companies when their socially connected industry peers have done well in the previous month and vice versa, resulting in an excess return of 84 basis points per month. This social connection effect is stronger for companies that are not in the spotlight.

3. Data

3.1 Data Gathering

In this paper, I will be using several different databases, most of which are combined in the Compustat database from Wharton Research Data Services. The time period that is of interest for this study is 1977-2021. This is because all firms are mandated by the Statement of Financial Accounting Standards (SFAS) to disclose all relevant financial information of an industry segment within their firm that accounts for over 10% of the company's annual consolidates sales from 1976 onwards. By starting from 1977, the study ensures a full year of the mandated disclosures is included.

I will be looking at the firms that are listed on the main exchanges of the United States, namely the New York Stock Exchange (NYSE), the American Stock Exchange (AMEX, until 2008) and the NASDAQ. As the US capital market is among the largest and most liquid, and is a hub for diverse firms with various industries and sizes, it provides a large dataset that is representative and can offer insights that are broadly applicable in other developed markets. Both active and inactive stocks are used, as to diminish survivorship bias in the sample. The main datasets obtained from Compustat are the following:

- CCM Linking Table

  This dataset is acquired for all companies in the North American database. This will be used for merging the annual sale files with the CRSP dataset. Data include: LPERMNO, gvkey, linkdate etc.
- Fundamental Annual Data

  This dataset is downloaded for all firms included in the North American Database. The data included is: Annual Sales, Book Value per Share, Common Shares Outstanding etc.
- History Segment Data

  This dataset is obtained for all firms included in the North American Database. The data that is included is: Sales (per industry) and the 4-digit SIC code for each industry.
- CRSP Data

  The CRSP monthly stock files are acquired for all firms in the North American Database from Compustat. Data included in this dataset is: Price, Volume, Return, Shares Outstanding etc.
- *IBES Analyst data*

  The summary history files from the Institutional Brokers' Estimate System (IBES) are downloaded from Compustat for the US file. This data includes the number of estimates per firm per month, but also the mean of those consensus estimates.

Additional data that is required for the analyses later in this paper are obtained from the Kenneth R. French Data Library[1]. This includes:

- $r_{m,t} - r_{f,t}$ – The market factor which represents the excess return of the market over the risk-free rate.
- SMB_t – The size factor. Excess returns of small-cap companies over big-cap companies.
- HML_t – The value factor. Excess returns of value stocks over growth stocks.
- RMW_t – The profitability factor. The difference between the returns of firms with robust and weak operating profitability.
- CMA_t – The investment factor. The difference between the returns of firms that invest conservatively and aggressively.

3.2 Data filtering

Once all data was obtained, the data needs to be checked for errors, missing data, and certain constraints that I want to apply. First, I require firms to have both market value and book value available at the end of the previous fiscal year, as is done by Cohen and Lou (2012). I also exclude the stocks that are below the five-dollar mark at the beginning of the period to weaken the effect of micro-cap stocks.

[1] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html#Research

Firms are characterised either as conglomerates or standalone firms, the same way as in Cohen and Lou (2012). A standalone firm is determined as a firm who operates in one segment that comprises at least 80% of total sales reported in Compustat annual files. This is to remove firms that do operate in more industries but fail to report data for some industry segments. Conglomerate firms are determined similarly, those that operate in more than one segment and those segments combined comprise of more than 80% of total sales. This ensures that the sum of the segments is a fair representation of the entirety of the firm. As I am using the percentage of segment sales of total sales to define a firm either as a conglomerate or standalone, I remove all observations that do not have annual total sales from the sample. I exclude all missing segment sales from the sample as well. Lastly, I remove segments with missing 4-digit SIC codes from the sample.

Table 1: Descriptive statistics, 1977 - 2021

	Min	Median	Max	Mean	Std Dev
Panel A: Time Series (1977 – 2021)					
Number of Congl firms per year	946	1367	1748	1331	215
Number of standalones per year	922	2123	4465	2312	747
Full sample % coverage of CRSP universe (EW)	52.91	69.52	88.94	69.66	10.14
Full sample % coverage of CRSP universe (VW)	65.17	74.76	79.44	74.00	3.10
Congl firms % of CRSP universe (EW)	16.34	26.29	39.83	25.96	6.15
Congl firms % of CRSP universe (VW)	30.76	42.47	48.49	41.81	4.82
Standalones % of CRSP universe (EW)	26.46	42.44	67.31	43.70	10.34
Standalones % of CRSP universe (VW)	19.97	27.77	41.51	28.70	5.90
Panel B: Pooled firm-year observations					
No. of industries per conglomerate	2	3	11	3.05	1.17
Percent of sales per industry segment	0	0.33	1	0.43	0.35

Notes: This table presents descriptive statistics of each year for the entire sample period, 1977-2021. Percent coverage of the CRSP stock universe (EW) is the number of stocks, either in the conglomerate dataset, standalone dataset, or both in a given year, divided by the total number of stocks in the CRSP dataset. Percent coverage of the CRSP stock universe (VW) is the total market capitalisation of stocks, either in the conglomerate dataset, standalone dataset, or both in a given year, divided by the total market capitalisation of the CRSP dataset.

After all screening procedures, the sample contains over 160.000 distinct firm-year observations, with around 100.000 attributed to standalone firms, and 60.000 to conglomerate firms. In table 1, the descriptive statistics are shown of the sample, covering the time period from 1977-2021. In Panel A, the number of firms per year is shown, but also the percentage of my sample compared to the CRSP universe, both equal weighted (based on the number of firms) and value weighted (in terms of market capitalisation). The sample covers on average almost 70% of the CRSP universe in terms of the number of firms included, and 74% of the CRSP universe based on total market capitalisation. In Panel B, the number of industries per conglomerate each year is depicted, as well as the percentage of total sales that are earned by each industry. The average number of industries per conglomerate is 3.05, while for some firms it can be as high as 11 industries. Additionally, the average sales contributed by a segment to the total sales of a firm is around 43%.

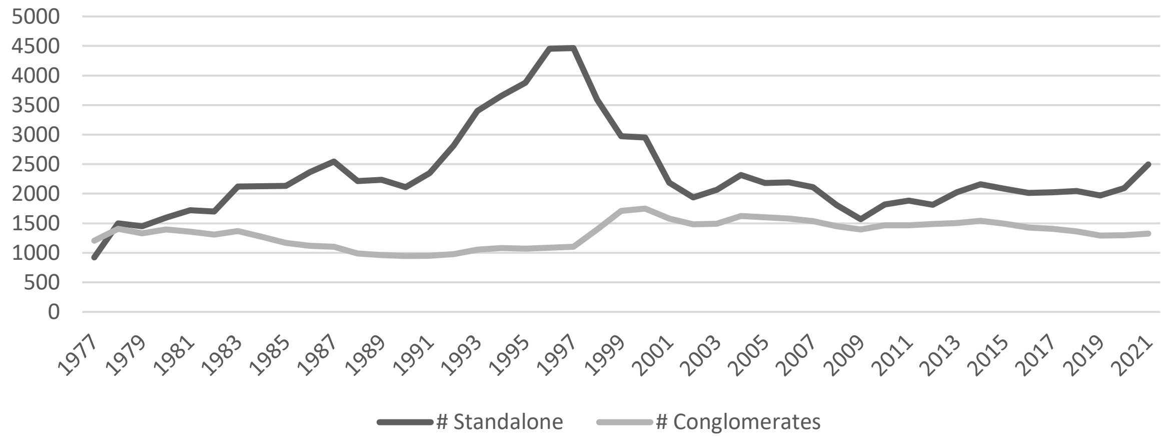

Figure 1: Number of firms per year

In figure 1, the number of standalones and conglomerates are plotted over the time period from 1977-2021. Interestingly, the number of standalones is growing massively during the 1990's. This is probably due to the surge in technology startups. The presence of speculative investing and the great amount of venture capital being available for technology companies can be the reason for the spike during the 90's in standalone firms. As Wheale and Amin (2003) mention, investors were not rational and too optimistic about the ideas and potential of the new technology sector that emerged due to the access to internet. In the beginning of 2000 and even a little before that, there is a huge decrease in standalone firms which coincides with the burst of the Dot Com bubble. There is also a dip in the number of standalones when the global financial crisis broke out in 2007.

4. Methodology

The main assumption of this paper, based on the research presented above, is that investors often face challenges in gathering, managing, and incorporating information about more complex firms, leading to varying delays in reflecting the same information in company valuations. This is specifically examined based on industries present in certain firms. Industries are defined through their two-digit Standard Industrial Classification (SIC) code, which in turn is derived from the four-digit SIC code obtained from the Compustat Segment data.

4.1 Portfolio formation

Firstly, while it is fairly straightforward to adjust the valuation of a company operating solely in one industry given certain industry information, for a firm with operations in several industries the task becomes more complex. In order to test the idea, I follow a similar methodology as Cohen and Lou (2012). At the end of June in each year, for each conglomerate in the sample, a corresponding 'pseudo-conglomerate' is formed. This 'pseudo-conglomerate' is essentially a portfolio of the various industries a conglomerate operates in, but only using the standalone returns from each industry.

Then these segment portfolios are weighted based on the sales percentage of each industry segment within the actual conglomerate. Meaning that if a conglomerate operates in four different industry segments, and the sales percentages of total sales are 10%, 20%, 30% and 40%, then a pseudo-conglomerate is formed as follows:

$$PCRET = 0.1 \times (returns\ of\ industry\ 1) + 0.2 \times (returns\ of\ industry\ 2) + 0.3 \times (returns\ of\ industry\ 3) + 0.4 \times (returns\ of\ industry\ 4) \quad (1)$$

with the returns calculated based solely on standalone firms.

I repeat this step at the beginning of every month starting in July, based on segment information from the previous fiscal year. In order to form winning and losing portfolios, all conglomerate firms are sorted into deciles using the pseudo-conglomerate returns from the month before. With decile 10 being the winner portfolio and decile 1 the loser portfolio. Then the decile portfolios are rebalanced every month to maintain either equal or value weights. This strategy is called 'complicated processing' by Cohen and Lou (2012).

4.2 Complicated Processing Portfolios

Following this, I calculate the excess returns by using the risk-free return from the Kenneth R. French Library, and I add the risk factors one at a time. First, I include the market risk premium, then the SMB and HML factor from Fama and French (1995), then the momentum factor (Carhart, 1997) and lastly the liquidity factor (Pastor and Stambaugh, 2003). The alpha from this regression are the excess returns after accounting for these various risk factors. The regression equation is as follows:

$$Excess\ Returns_t = \alpha + \beta_1 * (\mathrm{r}_{m,t} - r_{f,t}) + \beta_2 * SMB_t + \beta_3 * HML_t + \beta_4 * MOM_t + \beta_5 * LIQ_t + \epsilon \quad (2)$$

with the excess returns being either equal weighted or value weighted.

All regressions are run for every decile. Furthermore, I create a Winner-Loser portfolio. Meaning that the excess returns from the loser portfolio are subtracted from the winner portfolio, simulating the situation in which an investor would go long in the previous winning firms, while shorting the past losing firms. The winner-loser portfolio is also regressed using both five factor models.

4.3 Forecasting regressions

Next, I conduct Fama Macbeth (1973) forecasting regressions, in order to highlight the effect of the main variable, namely pseudo-conglomerate returns from the previous period, while controlling for other variables that can affect firm returns. The other variables that I am controlling for are the lagged returns of the conglomerate itself, which is to account for the short-term reversal effect highlighted by Jegadeesh (1990). In addition, as referenced in Moskowitz and Grinblatt (1999), the value-weighted primary industry return of the conglomerate from the previous month is added to the regression. On top of this, extra controls such as lagged size, book-to-market ratio, price momentum and the turnover of the conglomerate are added. Cross-sectional regressions are then conducted every month. Subsequently, I average the estimates from each period over time to determine the anticipated risk premium for each control. Newey-West standard errors (Newey and West, 1987) are used for up to 12 lags of autocorrelation.

As Fama-MacBeth regressions require a two step-approach, I will highlight the formulas used below:

$$RET_t = \alpha_i + \lambda_{t-1}^{PCRET}\beta_{i,t-1}^{PCRET} + \lambda_{t-1}^{size}\beta_{i,t-1}^{size} + \lambda_{t-1}^{B/M}\beta_{i,t-1}^{B/M} + \lambda_{t-1}^{MOM}\beta_{i,t-1}^{MOM} + \lambda_{t-1}^{turnover}\beta_{i,t-1}^{turnover} + \epsilon_{i,t-1} \quad (3)$$

This is the first step in which I use the characteristics as explanatory variables in *T* cross-sectional regressions, in my case 534 months. The variables of interest are the λ_t^f for the exposure to each risk factor or explanatory variable.

If there is a linear relationship between the returns and the variable in a month, then $\lambda_t^f \neq 0$. Then the time-series average is taken of the estimates.

$$\frac{1}{T}\sum_{t=1}^{T}\hat{\lambda}_t^f \tag{4}$$

The estimates that come out of this equation can be interpreted as the risk attributes for the risk factors that are of interest for this paper.

4.4 Firm Complexity and Analyst Coverage

These Fama-MacBeth forecasting regressions are used for further analysis as well. As I want to see whether characteristics that come from the complexity of the firm itself has influence on the previous results. One of the measures I am using is the Herfindahl-Hirschman Index (HHI). This index was developed independently by Hirschman (1945) and Herfindahl (1950) to measure the market concentration and can therefore be used to determine market competitiveness. The formula is as follows:

$$HHI = s_1^2 + s_2^2 + s_3^2 + .. + s_n^2 \tag{5}$$

Other factors will be market capitalization with regards to the NYSE median, the average daily turnover of the firm and the number of estimates from analysts.

Furthermore, I want to look at the information embedded in analyst forecasts. Analysts are professionals who analyse both simple and more complex firms. Therefore, these analysts are required to predict what will happen for the entirety of the firm and thus have the use the same information as other traders. Meaning that the complexity of the information processing capacity is identical. However, as analysts are predicting, they do not have to engage in actual trading. Thus, if information processing complexity is the primary driver of the observed results, a similarity between analysts forecasts for both simple and conglomerate firms is expected.

To test this hypothesis, I use the annual earnings forecasts from these analysts. The mean estimates are used per firm per month. Then I calculate the revision using a similar formula as Lys and Sohn (1990).

$$\Delta FEPS_{ijt} = \left(FEPS_{ijt} - FEPS_{ijt-1}\right)/\left(P_{ijt-1}\right) \tag{6}$$

With FEPS being the forecasted EPS and P_{ijt-1} being the lagged price of the stock.

The idea is similar to that of before, meaning the average forecast per industry per month, while only using the standalone firms, is extracted from the Compustat files. After that, I create the pseudo-conglomerate forecasts (PCF) for each conglomerate. Then Fama-MacBeth forecast regressions are run in the same way as equation 3 and 4, which are described above. With this, I investigate whether forecast revisions for simple standalone firms can predict subsequent revisions for their more complex conglomerate pairing.

4.5 Robustness and sub-period analysis

Moreover, I am running some robustness tests. First, the complicated processing portfolios mechanism is also done using the Fama and French (2015) five factor model. As the paper from Cohen and Lou came out in 2012, they did not use the newer factors as controls. Meaning that equation 2 will be modified to the following equation:

$$Excess\ Returns_t = \alpha + \beta_1 * (\mathrm{r}_{m,t} - r_{f,t}) + \beta_2 * SMB_t + \beta_3 * HML_t + \beta_4 * RMW_t + \beta_5 * CMA_t + \epsilon \quad (7)$$

Finally, I divide the sample into four distinct sub-samples based on time. I create sub-samples for the periods 1977-1988, 1989-1999, 2000-2010 and 2011-2021. These time periods are chosen as they all resemble 132 months of data, except the first time period which has 138 months of data. This division is interesting as it can become clear whether the effect is there in general, or that it is specific to a certain time period. Additionally, it is interesting to find out how the strategy works during the period of 2000-2010 as those years saw two big events happening to the market, namely the Dot Com bubble of 2001 and the great financial crisis of 2008. Finally, the inclusion of the 2011-2021 period can offer insights into the post-publication predictability as the paper from Cohen and Lou (2012) uses data up to 2009.

5. Results

This part of the paper will focus on the results of the different analyses described above considering the sample of firms over the period of 1977-2021. This includes the complicated processing analysis of the decile portfolio and their excess returns, the Fama-MacBeth forecasting regressions on the effect of the lagged pseudo-conglomerate returns and also the sub-period analysis.

5.1 Complicated Processing

Table 2 shows the results from the complicated processing portfolio strategy for the period of 1977-2021. If the premise holds that investors' restricted resources and capabilities, coupled with the inherent complexity of processing information for complicated firms such as conglomerates, influence how information is revealed for these companies, then there's an expectation. Specifically, updates to the values of pseudo-conglomerates, reflected in their price changes, should act as a forecast, predicting the subsequent adjustments to the values of their matched conglomerate firm and therefore their future prices. As a result, the challenges investors face in rapid processing and acting upon that information in a versatile company would be further highlighted.

In panel A of table 2, the results are given based on equal weighting. Panel A reports estimates that justify the hypothesis that complicated information processing can influence the rate at which investors seem to incorporate available information into stock prices of multifaceted firms. Looking at the excess returns, deciles 1 through 10, all seem to have significant excess returns based on their paired pseudo-conglomerate. By utilising the long-short strategy described earlier, I find monthly excess returns of 60 basis points, significant at the 1% level. This is equivalent to a return of 7.44% per year, using this strategy.

Adding controls for known return determinants, such as Fama and French (1993) three-factor, momentum and liquidity leave results largely unchanged. With momentum having the largest influence, but still leaving excess returns at 51 basis points per month, significant at the 1% level. These results are similar to the ones found by Cohen and Lou (2012). What stands out is that the returns for the decile 2 portfolio seem to be lower than those of the losing portfolio for all cases. This might be due to short-term reversals as documented before by Shiller (1981), who suggests investor overreaction as a sentiment-based explanation of short-term reversals.

Table 2: Complicated Processing Portfolios, Excess Returns 1977-2021

Decile	Excess Returns	1-Factor Alpha	3-Factor Alpha	4-Factor Alpha	5-Factor Alpha
Panel A: Equal weights					
1	1.38%	0.52%	0.38%	0.60%	0.56%
	(4.86)	(3.27)	(2.96)	(4.36)	(4.15)
2	0.8%	0.04%	-0.05%	0.03%	0.06%
	(3.44)	(0.38)	(-0.58)	(0.35)	(0.66)
3	1.00%	0.21%	0.10%	0.17%	0.14%
	(4.17)	(1.94)	(1.27)	(2.28)	(1.75)
4	1.00%	0.25%	0.14%	0.20%	0.22%
	(4.36)	(2.33)	(1.84)	(2.63)	(2.78)
5	1.05%	0.30%	0.18%	0.23%	0.31%
	(4.58)	(2.77)	(2.47)	(3.06)	(4.12)
6	1.11%	0.38%	0.28%	0.31%	0.40%
	(4.98)	(3.69)	(4.01)	(4.16)	(5.29)
7	1.18%	0.45%	0.34%	0.39%	0.44%
	(5.26)	(4.25)	(4.58)	(4.96)	(6.00)
8	1.38%	0.66%	0.56%	0.58%	0.62%
	(6.23)	(6.17)	(7.57)	(7.30)	(7.36)
9	1.42%	0.68%	0.59%	0.60%	0.65%
	(6.13)	(5.95)	(6.83)	(6.19)	(7.06)
10	1.98%	1.21%	1.12%	1.10%	1.15%
	(7.90)	(8.85)	(10.11)	(9.88)	(9.66)
L/S	0.60%***	0.69%***	0.74%***	0.51%***	0.59%***
	(3.53)	(3.94)	(4.29)	(2.98)	(3.41)
Panel B: Value weights					
1	0.93%	0.21%	0.19%	0.33%	0.18%
	(3.87)	(1.52)	(1.37)	(1.97)	(1.13)
2	1.01%	0.28%	0.27%	0.34%	0.31%
	(4.44)	(2.49)	(2.34)	(2.79)	(2.48)
3	1.13%	0.42%	0.39%	0.42%	0.25%
	(5.13)	(4.05)	(3.79)	(3.99)	(2.21)
4	1.21%	0.53%	0.49%	0.48%	0.53%
	(5.68)	(4.91)	(4.74)	(4.63)	(4.77)
5	1.20%	0.43%	0.39%	0.39%	0.37%
	(5.25)	(4.00)	(3.75)	(3.75)	(3.63)
6	1.31%	0.64%	0.60%	0.59%	0.62%
	(6.27)	(6.09)	(5.79)	(5.51)	(6.12)
7	1.14%	0.47%	0.44%	0.45%	0.46%
	(5.49)	(4.73)	(4.58)	(4.58)	(4.54)
8	1.22%	0.55%	0.53%	0.52%	0.59%
	(5.84)	(5.50)	(5.44)	(5.01)	(5.25)
9	1.36%	0.67%	0.65%	0.56%	0.53%
	(6.02)	(5.38)	(5.27)	(4.60)	(4.22)
10	1.53%	0.82%	0.80%	0.74%	0.70%
	(6.47)	(5.96)	(5.78)	(5.17)	(4.91)
L/S	0.61%***	0.61%***	0.60%***	0.41%*	0.52%**
	(2.96)	(2.96)	(2.88)	(1.76)	(2.31)

Note: This table presents the excess returns of calendar-time portfolios. At the start of every month, all conglomerate stocks are arranged in increasing order based on the past month's return of their associated pseudo-conglomerates. These ranked stocks are then distributed among one of ten decile portfolios. All stocks are equal (value) weighted and are rebalanced monthly to maintain these equal (value) weights. This table considers stocks priced over $5 at the beginning of the formation period. Alpha represents the intercept from the regressions analysis of the monthly excess return from the strategy. Included are the Fama and French (1993) factor loadings, the Carhart (1997) momentum factor and the liquidity factor from Stambaugh and Pastor (2003). L/S refers to the alpha of a no-cost portfolio of conglomerates which goes long on firms with the highest 10% pseudo-conglomerate returns in the past month, while shorting conglomerates with the bottom 10% of pseudo-conglomerate returns in the past month. Returns are expressed as monthly percentages. T-statistics are presented below the coefficient estimates. Values are rounded to two decimal places. *, **, *** denote statistical significance at the 10%, 5% and 1% level, respectively.

In panel B, the value weighted estimates are reported. The estimate of excess returns is similar, as there is a monthly excess return of 61 basis points per month, significant at the 1% level. However, when introducing the momentum factor, there is a decrease to 41 basis points per month, only significant at the 10% level. This result slightly differs from Cohen and Lou (2012) as they find significant returns at the 5% even when adding the momentum factor. Another difference is the magnitude of the estimate. My results seem to have cut the magnitude of the effect by half, as Cohen and Lou (2012) seem to have found an excess return of 118 basis points per month. Both of these changes can come from the fact that investors might have learned about the outcome from the paper and have acted upon the new information. The literature suggests that anomalies are diminishing over time with a trend towards zero (Marquering et al. 2006) and that post-publication return-predictability decreases with 35% as found by Mclean and Pontiff (2016).

Siganos (2007) in its work on momentum strategies also looked at the large downside risk associated with short selling and its limited upside potential. As a result, Foltice and Langer (2015) elaborated further on Siganos' idea and implemented a strategy that would only buy past winners. In addition, this would also avoid trading costs linked to short selling. They found positive returns, even outperforming the S&P500 benchmark by 2.44% per month for the period of 1991 to 2010. My results report that the coefficients of decile 1 in table 2 are all positive and most significantly different from zero, meaning it decreases the returns from the long-short strategy implemented in this paper. By not going short in the lowest decile portfolio, the excess returns could increase to 198 basis points per month, or 26.52% annualised return.

Table 3: Complicated Processing Portfolios, factor loadings 1977-2021

Decile	Excess Returns	5-Factor Alpha	MKT	SMB	HML	MOM	LIQ
Panel A: Equal weights							
1	1.38%	0.56%	1.093	0.704	0.199	-0.272	-0.016
	(4.86)	(4.15)	(30.64)	(9.53)	(2.65)	(-4.08)	(-0.67)
2	0.8%	0.06%	1.025	0.513	0.159	-0.108	0.011
	(3.44)	(0.66)	(40.44)	(8.18)	(2.84)	(-2.90)	(0.59)
3	1.00%	0.14%	1.012	0.507	0.218	-0.090	-0.011
	(4.17)	(1.75)	(44.33)	(8.53)	(4.86)	(-2.85)	(-0.65)
4	1.00%	0.22%	0.989	0.508	0.231	-0.075	0.007
	(4.36)	(2.78)	(45.79)	(11.23)	(6.03)	(-2.89)	(0.46)
5	1.05%	0.31%	0.969	0.551	0.257	-0.065	0.028
	(4.58)	(4.12)	(41.67)	(13.02)	(6.06)	(-2.97)	(1.86)
6	1.11%	0.40%	0.948	0.533	0.195	-0.045	0.035
	(4.98)	(5.29)	(44.73)	(15.65)	(6.18)	(-2.07)	(1.79)
7	1.18%	0.44%	0.957	0.537	0.204	-0.054	0.020
	(5.26)	(6.00)	(42.86)	(11.24)	(5.62)	(-1.78)	(1.04)
8	1.38%	0.62%	0.939	0.539	0.196	-0.025	0.016
	(6.23)	(7.36)	(43.78)	(15.04)	(5.08)	(-0.85)	(0.96)
9	1.42%	0.65%	0.958	0.596	0.194	-0.028	0.018
	(6.13)	(7.06)	(36.96)	(11.84)	(3.89)	(-0.56)	(0.96)
10	1.98%	1.15%	1.013	0.614	0.160	0.014	0.016
	(7.90)	(9.66)	(34.11)	(13.06)	(2.88)	(0.33)	(0.59)
L/S	0.60%***	0.59%***	-0.080*	-0.090	-0.039	0.286***	0.031
	(3.53)	(3.41)	(-1.86)	(-1.12)	(-0.37)	(3.86)	(0.95)
Panel B: Value weights							
1	0.93%	0.18%	1.002	0.001	0.001	-0.159	-0.055
	(3.87)	(1.13)	(26.05)	(0.01)	(0.01)	(-1.95)	(-1.83)
2	1.01%	0.31%	1.030	-0.076	0.050	-0.085	-0.008
	(4.44)	(2.48)	(30.94)	(-1.14)	(0.77)	(-1.69)	(-0.30)
3	1.13%	0.25%	1.065	-0.148	0.148	-0.039	-0.065
	(5.13)	(2.21)	(36.11)	(-2.75)	(2.82)	(-1.09)	(-2.39)
4	1.21%	0.53%	0.985	-0.096	0.171	0.009	0.018
	(5.68)	(4.77)	(35.38)	(-2.21)	(3.53)	(0.29)	(0.79)
5	1.20%	0.37%	0.977	0.003	0.167	-0.003	-0.008
	(5.25)	(3.63)	(31.73)	(0.05)	(2.79)	(-0.08)	(-0.40)
6	1.31%	0.62%	0.974	-0.031	0.157	0.016	0.013
	(6.27)	(6.12)	(30.49)	(-0.65)	(2.89)	(0.49)	(0.52)
7	1.14%	0.46%	0.974	-0.087	0.132	-0.015	0.002
	(5.49)	(4.54)	(38.93)	(-2.00)	(2.83)	(-0.38)	(0.11)
8	1.22%	0.59%	0.966	-0.074	0.090	0.005	0.025
	(5.84)	(5.25)	(36.08)	(-1.58)	(1.85)	(0.13)	(1.13)
9	1.36%	0.53%	0.999	0.057	0.087	0.116	-0.011
	(6.02)	(4.22)	(27.68)	(0.66)	(1.13)	(1.87)	(-0.41)
10	1.53%	0.70%	1.051	-0.092	0.145	0.078	-0.014
	(6.47)	(4.91)	(28.32)	(-1.43)	(2.53)	(1.20)	(-0.48)
L/S	0.61%***	0.52%**	0.050	-0.093	0.145	0.237*	0.041
	(2.96)	(2.31)	(0.88)	(-0.94)	(1.30)	(1.91)	(0.90)

Note: This table presents the factor loadings of calendar-time portfolios. At the start of every month, all conglomerate stocks are arranged in increasing order based on the past month's return of their associated pseudo-conglomerates. These ranked stocks are then distributed among one of ten decile portfolios. All stocks are equal (value) weighted and are rebalanced monthly to maintain these equal (value) weights. This table considers stocks priced over $5 at the beginning of the formation period. Alpha represents the intercept from the regressions analysis of the monthly excess return from the strategy. Included are the Fama and French (1993) factor loadings, the Carhart (1997) momentum factor and the liquidity factor from Stambaugh and Pastor (2003). L/S refers to the alpha of a no-cost portfolio of conglomerates which goes long on firms with the highest 10% pseudo-conglomerate returns in the past month, while shorting conglomerates with the bottom 10% of pseudo-conglomerate returns in the past month. Returns are expressed as monthly percentages. T-statistics are presented in parentheses. *, **, *** denote statistical significance at the 10%, 5% and 1% level, respectively.

In table 3, instead of looking at the excess returns for different factor models, I examine the factor loadings of the pseudo-conglomerate returns on the actual conglomerate returns for the five-factor model with factors including the market (MKT), size (SMB), value (HML), momentum (MOM) and liquidity (LIQ). Again, panel A and panel B represent equal weights and value weights respectively. From panel A it is evident that the only significant factor to the excess return of the long-short strategy is momentum while the market factor is only significant at the 10% level. Momentum seems to be losing its significance when moving to value weights in panel B as it is positive and significant only at the 10% level. This in turn explains why excess returns in the four-factor model from table 2 is the lowest for equal weights, and second lowest for the value weights with only a 10% significance. As a result, the momentum effect is present in our sample and as such influences the excess returns found from the complicated processing portfolios strategy.

5.2 Regression tests

I now turn to test the hypothesis of investor inabilities to process complex information in terms of the Fama MacBeth forecasting regressions. This enables me to examine the effect of the independent variable, which is the lagged pseudo-conglomerate return, more accurately while controlling for other factors of firm returns. These regressions are done in a similar fashion as Cohen and Lou (2012). The results of this regression are reported in table 4.

The independent variables used in the regression are the lagged pseudo-conglomerate returns paired to the conglomerate, the return of the conglomerate in the previous month and the returns of the industry portfolio associated with the primary industry of the conglomerate. Additional controls such as size, book-to-market, momentum, and turnover are added. Columns 1 and 2 of table 4 show the results of the Fama-MacBeth forecasting regressions on the returns of the conglomerate. They can be viewed as the basic results.

As can be seen from the table below, the lagged pseudo-conglomerate returns appear to be a large and significant determinant of the paired conglomerate return in the next month. Even after controlling for size, book-to-market, momentum and turnover, the basis coefficient is equal 7.187 and significant at the 1% level. Meaning that a one-standard-deviation increase in the lagged pseudo-conglomerate return results in a 54-basis point increase in its paired conglomerate return this month. By adding the lagged returns of the conglomerate itself and the industry returns, there is more control for short-term stock reversal and industry momentum. However, in column 2 the coefficient remains largely unchanged as it is still positive and significant at the 1% level. These results largely coincide with the results found by Cohen and Lou (2012).

The analysis in and of itself is the same for columns 3 and 4. However, instead of using the returns of the conglomerate as the dependent variable, now I use the difference between the conglomerate return and the paired primary industry return value weighted, as used in Moskowitz and Grinblatt (1999). I do this to diminish the industry momentum effect from our predictor variable, namely the lagged pseudo-conglomerate returns. Specifically, the stock return continuation is now free form industry-wide return autocorrelation, making our variable of interest isolated from this industry effect. Examining column 3, the coefficient attributable to the lagged pseudo-conglomerate returns has decreased in magnitude, to a value of 0.969 with it only being significant at the 10% level. Column 4 indicates a positive and significant value for the predictor $PCRET_{t-1}$ of 4.314. Even though my significance and magnitude are lower for column 3, this is still in line with the results from Cohen and Lou (2012). Yet they mention that if industry-wide return continuation should not be evident anymore in column 4, the independent variable $INDRET_{t-1}$ should be insignificant. I find a different result. The coefficient in table 4 shows that my value is equal to -5.113 and significant at the 1% level. Meaning that there is still a level of industry-wide return continuation present in the sample.

An alternative approach to remove industry returns from the conglomerate returns is also given by Cohen and Lou (2012). They argue that subtracting the value weighted industry returns can be insufficient as this only reflects information from the conglomerates primary industry. In this sample, the primary industry of a conglomerate can represent as low as 35% of total sales for the conglomerate. To account for this discrepancy, in columns 5 and 6 the dependent variable is switched to the return of the conglomerate in this month minus the paired pseudo-conglomerate return in this month. This should incorporate information from all operating segments of the firm and the variable of interest should isolate just the complicated information processing mechanism. As a result, the coefficient in column 5 turns negative while in column 6 it is positive but insignificant, even at the 10% level. These results differ from Cohen and Lou (2012), as they find positive and significant results. This decrease in predictability was already occurring in columns 3 and 4, but by specifying the industry wide returns even further, the coefficients turn either negative or insignificant. Meaning that looking at column 6, our predictor variable can be considered as a refined measure of industry returns.

Table 4: Complicated Processing Returns, cross-sectional regressions 1977-2021

Dep variable *100	RET_t		$RET_t - INDRET_t$		$RET_t - PCRET_t$	
	(1)	(2)	(3)	(4)	(5)	(6)
$PCRET_{t-1}$	7.187***	6.405***	0.969*	4.314***	-2.957***	1.835
	(6.77)	(5.07)	(1.71)	(4.55)	(-3.13)	(1.26)
RET_{t-1}		-4.095***		-4.193***		-4.534***
		(-9.47)		(-9.57)		(-10.01)
$INDRET_{t-1}$		5.311**		-5.113***		-7.420***
		(2.46)		(-2.77)		(-3.41)
SIZE	-0.025**	-0.024**	-0.023**	-0.022**	-0.021**	-0.021**
	(-2.32)	(-2.23)	(-2.40)	(-2.28)	(-2.42)	(-2.39)
B/M	-0.014	-0.012	-0.007	-0.006	-0.009	-0.009
	(-1.57)	(-1.40)	(-1.01)	(-0.87)	(-1.37)	(1.25)
MOM	0.393***	0.398***	0.352***	0.364***	0.354***	0.371***
	(3.84)	(3.80)	(3.77)	(3.79)	(3.76)	(3.88)
TURNOVER	-0.016	-0.018*	-0.019**	-0.019**	-0.018**	-0.019**
	(-1.62)	(-1.67)	(-2.12)	(-2.07)	(-2.13)	(-2.18)
Adj R^2	0.02	0.03	0.01	0.03	0.02	0.03

Note: In this table the Fama-MacBeth forecasting regressions of stock returns are reported. In columns 1 and 2, the monthly return of the conglomerate is the dependent variable, in columns 3 and 4 it is the value-weighted industry returns subtracted from the conglomerate return and in columns 5 and 6 it is the excess return of the conglomerate over its paired pseudo-conglomerate. The independent variables are the pseudo-conglomerate returns of the previous month (PCRET), the lagged return of the conglomerate (RET) and the lagged return of the industry portfolio of the conglomerates primary industry (INDRET). All regressions include control variables which are measured at the end of June each year. These variables include size, book-to-market, momentum, and turnover. Every calendar month, cross-sectional regressions are run. Newey-West (1983) standard errors are used for up to 12 lags. Fama-MacBeth t-statistics are reported in parentheses. *, **, *** denote statistical significance at the 10%, 5% and 1% level, respectively.

Looking at the control variables, it appears that both size, book-to-market and momentum are all significant at the 5% through all iterations of the regression analysis. Turnover is only significant in column 6. The size factor is negative and significant for all columns, meaning that if a conglomerate is larger, the returns tend to be smaller. This is in line with the extensive literature, with one being from Fama and French (1992), that found that small- and mid-cap stocks tend to outperform large-cap stocks. In table 2 there are negative values for the book-to-market ratio, however they are insignificant and therefore not different from zero. In this case, the book-to-market ratio has had no effect on the return of conglomerates. Momentum seems to have a positive and significant effect on returns. This observation is consistent with the findings in Table 2, where the four-factor model that incorporated momentum showed the smallest excess returns among all models. Lastly, the turnover of a stock seems to have a small but negative impact on returns. This means that if the turnover of a share is higher, then the returns tend to be lower. This can happen due to investors trading on short-term information or overreaction to information, which both drive the price away from its intrinsic value. That in turn can lead to price reversals when the market corrects itself.

Next, I examine the effect of the characteristics of a conglomerate firm on the complicated information processing. One of the factors is how complicated a conglomerate is. As the pseudo-conglomerate pairing is based on the share of the industry inside a conglomerate, the number of industries can then be a measure of complexity of a conglomerate. The Herfindahl-index is used to define this complexity of the firm. Although the Hefindahl-index is usually used to examine market competitiveness, I use it in a similar fashion to measure the complexity of a firm based on the number of industries and its share within the firm. The hypothesis is that if a firm has more industries, the embedding of information into its price needs a more complicated analysis.

Another interesting phenomenon is limits-to-arbitrage. This theory defined in 1997 by Shleifer and Vishny, describes that pricing inefficiencies may persist due to the constraints on traders representing professional money management firms. These traders, when investing on behalf of clients, can face reputational risks if they attempt to arbitrage stock mispricings that might continue for an extended period of time. A client can then perceive incompetence and decide to withdraw its money from the fund, resulting in hesitation from these traders to engage in such arbitrage opportunities and therefore the persistence of mispricing occurs. When clients withdraw their money, the fund manager has to usually unwind its position at a loss if it does not retain enough funds. In this paper, I use the size of the firm as a proxy for the limits-to-arbitrage. The idea is that if a firm is larger then it does not have to unwind its position as quick as other firms, and therefore the limits-to-arbitrage effect should be less visible.

Finally, the investor inattention phenomenon is of interest as well. Is the effect present because investors are not aware of certain information or a particular stock? Proxies that are used to determine whether investors are aware of certain companies are the turnover and number of estimates from analysts. The turnover represents the total number of shares traded over a certain period divided by the number of shares outstanding. The turnover measure for attention has been used in the literature before by Hou, Peng and Xiong (2006) and Loh (2010). He argued that prior trading activity is a proxy for the amount of attention given to a firm by active investors. Additionally, the number of estimates from analysts is also a proxy for attention. Kelly and Ljungqvist (2012) found that the closure of 43 brokerage firms' research operations increased information asymmetry while share prices and uninformed investor's demand fell. The prospect here is that if a firm has more analysts covering the stock, the inattention problem should be attenuated more.

Table 5: Level of complexity in complicated firms, Fama MacBeth 1977-2021

Dep variable	**RET_t**			
***100**	**(1)**	**(2)**	**(3)**	**(4)**
$PCRET_{t-1}$	6.964***	7.816***	5.886***	6.538***
	(5.07)	(5.87)	(4.73)	(5.17)
$PCRET_{t-1}$ *Herfindahl > median*	-1.719			
	(-1.50)			
$PCRET_{t-1}$ *Mkt Cap > NYSE median*		-3.182***		
		(-3.64)		
$PCRET_{t-1}$ *Turnover > median*			1.130	
			(1.02)	
$PCRET_{t-1}$ *#Estimates > median*				-0.375
				(-0.34)
Controls	YES	YES	YES	YES
Adj R^2	0.04	0.04	0.04	0.03

Note: In this table are the Fama-MacBeth forecasting regressions of individual stock returns reported. The monthly return of the conglomerate is the dependent variable. The independent variables are the pseudo-conglomerate returns of the previous month (PCRET) and a number of interactions with this variable. These interactions terms are the Herfindahl index, which is based on the segment sales of a firm in a fiscal year, the market capitalisation of the firm at the end of June, the average daily turnover of the firm in the previous year and the number of estimates from different analysts at the end of June. These interaction terms are based on a dummy variable with one representing larger than the median, otherwise zero. All regressions include the control variables similar to table 4: lagged RET, INDRET, size, B/M, momentum, Turnover but also the dummy itself is included. Every calendar month, cross-sectional regressions are run. Newey-West (1983) standard errors are used for up to 12 lags. Fama-MacBeth t-statistics are reported in parentheses. *, **, *** denote statistical significance at the 10%, 5% and 1% level, respectively.

The results from the Fama-MacBeth regressions with the use of these characteristics are given in table 5. The dependent variable is the return of the conglomerate (RET_t), and the independent variable is the return of the pseudo-conglomerate pairing in the previous month ($PCRET_{t-1}$). In addition, an interaction term between the lagged pseud-conglomerate return and the factors described above are added one at a time. I do this through a dummy variable which has value one if the firm is above the median and zero otherwise. Besides this, the control variables given in table 4 and the dummy variable itself are used in each regression but are left out of the table to make it briefer. In column 1, the results are reported for the complexity of a firm through the Herfindahl-index. The coefficient estimate of the interaction term between the complexity of a firm and the previous pseudo-conglomerate returns is negative. This is what was expected as a higher Herfindahl-index indicates a less complicated firm and therefore a smaller return predictability based on past pseudo-conglomerate returns. However, the coefficient is only just insignificant, at the 13% level.

In column 2, the results for the proxy for limits-to-arbitrage are displayed. I create a dummy variable that equals one if the firm is above the NYSE median and zero otherwise. The NYSE median is the median market capitalisation of all firms in the NYSE sample. The coefficient of the lagged pseudo-conglomerate is still positive and significant, similar to column 2 of table 4. When looking at larger firms, the coefficient of the interaction term is negative at -3.182 and significant. This means that although the effect of complicated information processing is present in larger conglomerates, the effect is even stronger for smaller firms. This is consistent with the hypothesis that larger firms undergo less constraints from limits-to-arbitrage.

Finally, in column 3 and 4 the estimates for the proxies for investor attention are given. Again, dummy variables are created with the value of one if it is above the median for a firm and zero otherwise. When looking at turnover, it is evident that is a positive value but insignificant. The expectation was that if turnover was higher, the more attention there was for that stock and thus less effect of investor inattention. The coefficient is therefore in the wrong direction, as a negative value is reported. The coefficient for the number of estimates is in the right direction, however also insignificant. This underlines that the complications in the processing of information is the main driver of the return effect, and not necessarily investors ignoring information or are unaware of this information.

Most of the results obtained in table 5 are similar to those found by Cohen and Lou (2012). Except for the insignificant result of the more complicated firms based on the Herfindahl-index. They also find a negative coefficient; however, their coefficient is -3.458 and is significant. My result is twice as small at -1.719 and not significant. This does coincide with the literature on replication of anomalies in later papers, as they tend to have a trend towards zero (Marquering et al., 2006) and in 50% of cases fail the 10% significance found in the original paper (Hou et al., 2020).

Now moving to the information embedded in analyst forecasts. As analysts are constrained to give forecasts for the entire firm and not individual industry segments, I expect the forecast revision of the paired pseudo-conglomerate in the previous month to predict the forecast revision of the conglomerate in this month. I test this through Fama-MacBeth predictive regressions, similarly as in table 4. The results from this test are given in table 6.

Table 6: Analyst forecasts, 1977-2021

Dep variable	F_t	
***100**	**(1)**	**(2)**
PCF_{t-1}	-0.140	0.388**
	(-0.19)	(2.21)
F_{t-1}		29.154***
		(4.88)
$INDRF_{t-1}$	15.810***	4.378*
	(2.58)	(1.86)
SIZE	0.025	0.026
	(1.06)	(1.11)
B/M	0.027	0.015
	(0.88)	(0.79)
MOM	-0.089	0.052
	(-0.89)	(1.22)
TURNOVER	-0.009	-0.008
	(-0.90)	(-0.83)
Adj R^2	0.02	0.23

Note: In this table are the Fama-MacBeth forecasting regressions of revisions in analyst earnings forecasts. The mean revision of forecasts is the dependent variable in both columns (F). The independent variables are the lagged mean revision of forecasts for the pseudo-conglomerate (PCF) in column 1, while in column 2 the lagged conglomerate mean revision and the lagged mean revision of the conglomerate's primary industry (INDRF) are added. The regressions also include the controls similar to table 4 which are size, book-to-market ratio, momentum, and turnover, measured at the end of June of each year. Every calendar month, cross-sectional regressions are run. Newey-West (1983) standard errors are used for up to 12 lags. Fama-MacBeth t-statistics are reported in parentheses. *, **, *** denote statistical significance at the 10%, 5% and 1% level, respectively.

Table 6 shows somewhat mixed results. In column 1, the coefficient of the lagged pseudo-conglomerate mean revision is negative at -0.140, however the coefficient is insignificant. If the idea outlined before upholds, then I expect a positive and significant coefficient. However, the negative coefficient implies that if the mean revision for the pseudo-conglomerate goes up, the revision forecast for the conglomerate in this month should go down. This is counterintuitive to what I expect. The estimate for the revision of the primary industry forecast is positive and significant, which is in line with what I expect. If the forecast revision for the primary industry a conglomerate firm is operating in, is going up, then it makes sense that the forecast revision for the conglomerate in the next month goes up as well.

In column 2, I add an explanatory variable, namely the lagged mean revision of the conglomerate itself. The coefficient from this lagged variable on itself is highly positive and significant, meaning there is a lot of autocorrelation in the analyst forecast. The coefficient from the past pseudo-conglomerate mean revision is now positive and significant at the 5% level, albeit a small value. Meaning that standalone revisions can positively predict the forecasts for conglomerates. The forecasting power of the mean revision for the primary industry of a conglomerate has decreased due to the addition of the lagged revision of the conglomerate but stays positive and significant at the 10% level.

Most of the results obtained in table 6 are in line with the results from Cohen and Lou (2012), except for the negative and insignificant result in column 1. My results are a lot less pronounced as well, because Cohen and Lou find an estimate of 5.370 for the predictability of the lagged PCF on the forecast revision for the conglomerate, while I only find 0.388. The same is true for the INDRF estimate, they find 9.651, while I get 4.378. One of the reasons this can happen is that they used the IBES detail database, while I used the IBES summary database. In Black's work (1993), he accentuates that although the summary files indicate a timely consensus, it is possible that the consensus includes outdated forecasts that have not been revised after a significant information event.

5.3 Robustness

In this section, I want to address a concern brought forward by Black in 1993. He is concerned with 'data mining' and when a researcher picks what to do and the way to do it based on what others have done with the data. Therefore, I also look at a different specification of the complicated processing portfolios and the excess returns coming from the long-short strategy. Instead of using the Carhart (1997) momentum and Pastor and Stambaugh (2003) liquidity factors, I use the five factors from Fama and French (2015). In this paper I follow a similar methodology to Cohen and Lou, who published their paper in 2012 and thus could not do the inclusion of the factors described in Fama and French (2015). As a follow-up to their three-factor model (1993), they introduced two new factors that appeared to influence stock returns as well, namely a profitability factor and an investment factor called RMW and CMA respectively. The method is the same as that for table 2, just with the RMW and CMA factors instead of the momentum and liquidity factor. The results of the complicated processing portfolios are reported in table 7.

Table 7: Complicated Processing Portfolios Fama and French, 1977-2021

Decile	Excess Returns	1-Factor Alpha	3-Factor Alpha	5-Factor Alpha
Panel A: Equal weights				
1	1.38%	0.52%	0.38%	0.30%
	(4.86)	(3.27)	(2.96)	(2.11)
2	0.8%	0.04%	-0.05%	-0.11%
	(3.44)	(0.38)	(-0.58)	(-1.10)
3	1.00%	0.21%	0.10%	0.04%
	(4.17)	(1.94)	(1.27)	(0.39)
4	1.00%	0.25%	0.14%	0.06%
	(4.36)	(2.33)	(1.84)	(0.77)
5	1.05%	0.30%	0.18%	0.06%
	(4.58)	(2.77)	(2.47)	(0.86)
6	1.11%	0.38%	0.28%	0.16%
	(4.98)	(3.69)	(4.01)	(2.21)
7	1.18%	0.45%	0.34%	0.21%
	(5.26)	(4.25)	(4.58)	(2.74)
8	1.38%	0.66%	0.56%	0.45%
	(6.23)	(6.17)	(7.57)	(6.18)
9	1.42%	0.68%	0.59%	0.48%
	(6.13)	(5.95)	(6.83)	(5.53)
10	1.98%	1.21%	1.12%	1.01%
	(7.90)	(8.85)	(10.11)	(8.57)
L/S	0.60%***	0.69%***	0.74%***	0.71%***
	(3.53)	(3.94)	(4.29)	(3.76)
Panel B: Value weights				
1	0.93%	0.21%	0.19%	0.10%
	(3.87)	(1.52)	(1.37)	(0.59)
2	1.01%	0.28%	0.27%	0.21%
	(4.44)	(2.49)	(2.34)	(1.68)
3	1.13%	0.42%	0.39%	0.34%
	(5.13)	(4.05)	(3.79)	(3.08)
4	1.21%	0.53%	0.49%	0.43%
	(5.68)	(4.91)	(4.74)	(4.06)
5	1.20%	0.43%	0.39%	0.26%
	(5.25)	(4.00)	(3.75)	(2.46)
6	1.31%	0.64%	0.60%	0.48%
	(6.27)	(6.09)	(5.79)	(4.49)
7	1.14%	0.47%	0.44%	0.31%
	(5.49)	(4.73)	(4.58)	(3.31)
8	1.22%	0.55%	0.53%	0.41%
	(5.84)	(5.50)	(5.44)	(3.96)
9	1.36%	0.67%	0.65%	0.56%
	(6.02)	(5.38)	(5.27)	(3.96)
10	1.53%	0.82%	0.80%	0.61%
	(6.47)	(5.96)	(5.78)	(4.12)
L/S	0.61%***	0.61%***	0.60%***	0.51%**
	(2.96)	(2.96)	(2.88)	(2.12)

Note: This table presents the excess returns of calendar-time portfolios. At the start of every month, all conglomerate stocks are arranged in increasing order based on the past month's return of their associated pseudo-conglomerates. These ranked stocks are then distributed among one of ten decile portfolios. All stocks are equal (value) weighted and are rebalanced monthly to maintain these equal (value) weights. This table considers stocks priced over $5 at the beginning of the formation period. Alpha represents the intercept from the regressions analysis of the monthly excess return from the strategy. Included are the Fama and French (2015) factor loadings which include SMB, HML, RMW and CMA. L/S refers to the alpha of a no-cost portfolio of conglomerates which goes long on firms with the highest 10% pseudo-conglomerate returns in the past month, while shorting conglomerates with the bottom 10% of pseudo-conglomerate returns in the past month. Returns are expressed as monthly percentages. T-statistics are presented below the coefficient estimates. Values are rounded to two decimal places. *, **, *** denote statistical significance at the 10%, 5% and 1% level, respectively.

In panel A of table 7, the results of the complicated processing portfolios are given for each decile and the strategy of going long in past winners and shorting past losers, based on equal weights. Looking at the last column of the table, there is still a positive and significant excess return reported for the L/S strategy using the five factors from Fama and French (2015). There is even an increase of 0.11 percentage points compared to the excess returns column. There is a slight decrease moving from the three-factor model to the five-factor model, indicating that those extra factors did indeed cover some of the return predictability. Another similarity to table 2 is that the excess returns from the decile 2 portfolio seems to be lower again than the decile 1 excess returns.

In panel B, the results are reported based on value weights. Again, by utilising the long-short strategy with the included five-factors still yield us an excess return of 51 basis points per month on average. While this result is still significant at the 5% level, it is the lowest of all models, meaning that the five factors do incorporate a portion of the effect. Similarly to table 2, by implementing the strategy proposed by Siganos (2007) and Foltice and Langer (2015), excess returns can be larger. The excess returns from decile 1 are all positive meaning that in the end the L/S strategy is decreasing our possible returns. The factor loadings for the Fama and French (2015) five-factor model are in the appendix and will be left out in this section for brevity.

The next important part to validate the results and see if it is a general effect of the stock market, is to see whether the effect upholds throughout the entire time period by dividing the sample into sub-periods. I create sub-samples from 1977-1988, 1989-1999, 2000-2010 and 2011-2021. I choose these time periods as they all represent 132 months of data, except for the first period which has 138 months of data. By dividing the sample in more sub-samples, it can be more informative as to when the effect is present in the market and when the effect is more present. Additionally, it can also show what happens when the market undergoes certain global or financial conditions that either improve or worsen the financial markets.

This sub-period analysis is first done for table 2, the excess returns from the L/S strategy from the complicated processing portfolios. The excess returns for the decile portfolios for the given sub-periods and its corresponding factor loadings are reported in the appendix. The results are given in table 8.

Table 8: Complicated Processing Portfolio over sub-periods

Decile	**Excess Returns**	**1-Factor Alpha**	**3-Factor Alpha**	**4-Factor Alpha**	**5-Factor Alpha**
Panel A: Equal weights					
A.1: 1977 - 1988					
L/S	1.16%*** (4.33)	1.15%*** (4.26)	1.19%*** (3.98)	0.89%*** (3.40)	0.92%*** (2.71)
A.2: 1989 - 1999					
L/S	1.45%*** (4.69)	1.36%*** (4.36)	1.45%*** (4.35)	1.01%*** (3.27)	1.05%*** (3.17)
A.3: 2000 - 2010					
L/S	-0.35% (-0.79)	-0.36% (-0.81)	-0.31% (-0.75)	-0.31% (-0.81)	-0.05% (-0.14)
A.4: 2011 – 2021					
L/S	0.12% (0.43)	0.47% (1.63)	0.42% (1.48)	0.47% (1.47)	0.42% (1.38)

Panel B: Value weights

B.1: 1977 - 1988					
L/S	1.27%***	1.25%***	1.42%***	1.07%**	0.74%
	(3.23)	(3.12)	(3.21)	(2.59)	(1.52)
B.2: 1989 – 1999					
L/S	1.08%***	1.00%***	1.12%***	0.58%	0.65%
	(2.90)	(2.79)	(2.96)	(1.49)	(1.60)
B.3: 2000 – 2010					
L/S	-0.42%	-0.42%	-0.63%	-0.63%	-0.14%
	(-0.81)	(-0.80)	(-1.30)	(-1.34)	(-0.29)
B.4: 2011 – 2021					
L/S	0.46%	0.66%**	0.63%*	0.60%*	0.59%*
	(1.51)	(1.99)	(1.95)	(1.79)	(1.72)

Note: This table presents the excess returns of calendar-time portfolios. At the start of every month, all conglomerate stocks are arranged in increasing order based on the past month's return of their associated pseudo-conglomerates. These ranked stocks are then distributed among one of ten decile portfolios. All stocks are equal (value) weighted and are rebalanced monthly to maintain these equal (value) weights. This table considers stocks priced over $5 at the beginning of the formation period. Alpha represents the intercept from the regressions analysis of the monthly excess return from the strategy. Included are the Fama and French (1993) factor loadings, the Carhart (1997) momentum factor and the liquidity factor from Stambaugh and Pastor (2003). L/S refers to the alpha of a no-cost portfolio of conglomerates which goes long on firms with the highest 10% pseudo-conglomerate returns in the past month, while shorting conglomerates with the bottom 10% of pseudo-conglomerate returns in the past month. Returns are expressed as monthly percentages. T-statistics are presented below the coefficient estimates. Values are rounded to two decimal places. *, **, *** denote statistical significance at the 10%, 5% and 1% level, respectively.

Interestingly from table 8 is that the effect from table 2 which is positive and significant over the entire time-period, does not seem to be the case when looking at different sub-periods. For both panel A and panel B, it is evident that mostly the first two time periods (from 1977-1988 and 1989-1999) seem to show a positive and significant effect for the L/S strategy, while the other two time periods do not have significant results. The positive and significant results are also larger than the estimates found in table 2, meaning that those lower estimates come from the last two time periods. There are some positive and significant (mostly 10%) in 2011-2021 for the value weighted portfolio strategy. This indicates that the maybe the economic situations in the 2000-2010 period were mainly the reason for lower estimates or insignificance.

There are even negative excess returns from the L/S strategy in the 2000-2010 time period. During this time period there were macro-economic events happening such as the dot-com bubble in 2001 in which a lot of tech startups were going bankrupt and there was a lot of optimistic thinking about firm valuations, but also the global financial crisis of 2008 which was tough on the stock market and endured for some years after that. It has been apparent from the literature that buying past winners is not always the best strategy. Daniel and Moskowitz (2016) found that momentum strategies can experience constant negative returns during some periods of so-called panic states, following economic downturns and high market volatility. They found that during three months in 2009 returns of the loser portfolio were 169% while only being 8% for the winner portfolio. By implementing a dynamic momentum strategy based on forecasts of the mean of momentum and bear market indicators, they doubled their alpha compared to a static momentum strategy.

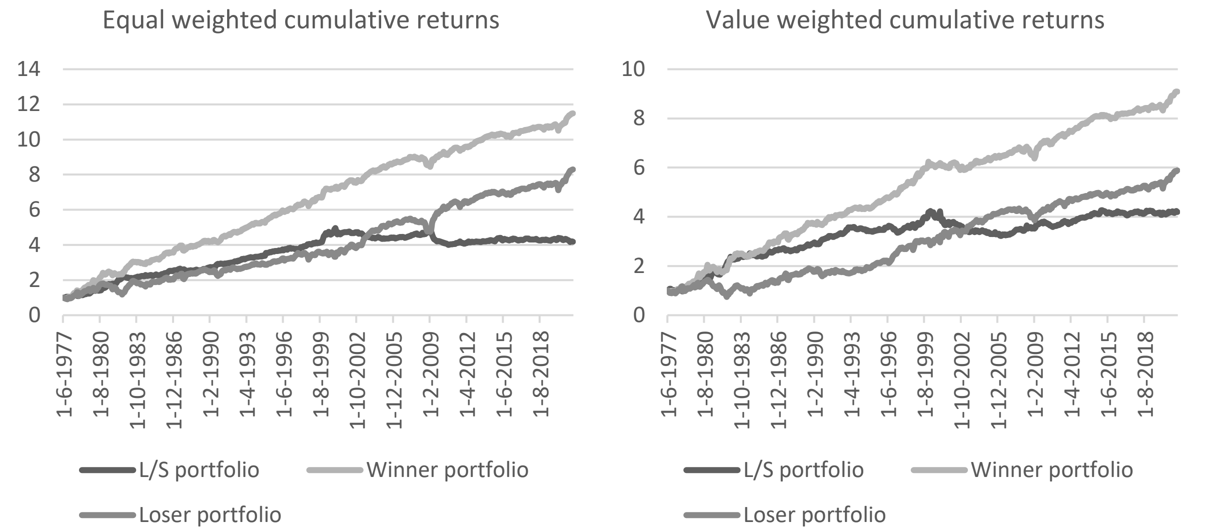

Figure 2: Cumulative returns 1977-2021

These results are underlined by figure 2. The graphs show the cumulative returns over the period 1977-2021, both equal and value weighted. The winner portfolio in both equal and value weighted cumulative returns has a small dip around the global financial crisis of 2008 but is apart from that increasing steadily throughout the period. The consistent outperformance of the winner portfolio, regardless of the weighting mechanism, suggests the resilience and strength of the stocks within this category. The loser portfolio seems to underperform compared to the winner portfolio especially in the beginning but seems to catch up after 2000 to even outperform the L/S portfolio. While the L/S portfolio appears to even outperform the winner portfolio around the early 1980’s with value weights, it loses its upward trajectory and seems to be running almost flat for the last 7 years in both equal and value weighted scenarios. These graphs visually represent the insignificant excess returns during those time periods that we’ve seen above in table 8.

Next, I examine the sub-periods for the Fama-MacBeth forecasting regressions as in table 4. These results are shown in table 9. Only the coefficients of the independent variables are displayed for brevity. As the predictability of past pseudo-conglomerate returns on conglomerate returns this month are of most interest, the sub-period analysis is only done for columns 1 and 2 of table 4.

Table 9: Complicated Processing Returns, cross-sectional regressions sub-periods

Dep variable *100	RET_t (1)	(2)
Panel A: 1977-1988		
$PCRET_{t-1}$	11.055***	8.871**
	(5.23)	(3.20)
RET_{t-1}		-7.315***
		(-11.62)
$INDRET_{t-1}$		14.255***
		(2.96)
Panel B: 1989 - 1999		
$PCRET_{t-1}$	11.678***	9.793***
	(7.83)	(3.48)
RET_{t-1}		-3.337***
		(-6.48)
$INDRET_{t-1}$		6.065*
		(1.79)
Panel C: 2000 - 2010		
$PCRET_{t-1}$	4.065**	5.876***
	(2.35)	(3.71)
RET_{t-1}		-2.883***
		(-4.36)
$INDRET_{t-1}$		-3.530
		(-0.83)
Panel D: 2011 - 2021		
$PCRET_{t-1}$	1.773*	0.970
	(1.75)	(0.63)
RET_{t-1}		-2.700***
		(-3.77)
$INDRET_{t-1}$		4.048
		(1.65)

Note: In this table the Fama-MacBeth forecasting regressions of stock returns are reported per sub-period. The monthly return of the conglomerate is the dependent variable. The independent variables are the pseudo-conglomerate returns of the previous month (PCRET), the lagged return of the conglomerate (RET) and the lagged return of the industry portfolio of the conglomerates primary industry (INDRET). All regressions include control variables which are measured at the end of June each year. These variables include size, book-to-market, momentum, and turnover, which are not mentioned for brevity. Every calendar month, cross-sectional regressions are run. Newey-West (1983) standard errors are used for up to 12 lags. Fama-MacBeth t-statistics are reported in parentheses. *, **, *** denote statistical significance at the 10%, 5% and 1% level, respectively

Interestingly, it appears that the effect of complicated information processing is again most evident in the late 1900's with it slowly decaying over time. This means that the results that we found in table 4 are mainly significant because of the first two-subperiods, therefore suggesting that it might not be a general effect in the market but just during that time. The coefficient in panel D is not even significant at the 5% level anymore in column 1, and lost all of its significance in column 2. An explanation as to why the results seem to decrease might also be the implementation of the XRBL mandate in the US. Research has found that stock prices are significantly more informative after the introduction of XRBL through information flow and dispersion, as stated by Huang et al. (2021).

Finally, I also do the sub-period analysis for the analyst forecast revision as in table 6. In this case, the controls are the same as in table 6 but are again left out here for brevity. The results are reported in table 10. The only positive and significant coefficients for the lagged pseudo-conglomerate revision (PCF_{t-1}) are actually in the 2000-2010 period, while the lagged revision forecast (F_{t-1}) is almost the least during that period. This coincides with the findings from Sidhu and Tan (2011) if you keep in mind that the mean revision for pseudo-conglomerate is insignificant for other periods. They find that forecasting errors were greater during the global financial crisis and analysts were quick to adjust their forecasts and may have even over-compensated for concerns of the effects of the economic crisis. Furthermore, the primary industry forecast revision seems to be the largest and significant during the first and last sub-period. The Dot Com bubble and the Global Financial Crisis might suggest that conglomerate-specific factors played a more dominant role than industry trends in shaping analyst' forecasts.

Table 10: Analysts forecasts, over sub-periods

Dep variable	F_t	
*100	(1)	(2)
Panel A: 1977-1988		
PCF_{t-1}	-1.435	0.736
	(-0.52)	(1.55)
F_{t-1}		18.638***
		(2.81)
$INDRF_{t-1}$	22.744*	11.160
	(1.89)	(1.36)
Panel B: 1989 - 1999		
PCF_{t-1}	0.264	-0.056
	(0.82)	(-0.55)
F_{t-1}		31.390*
		(1.84)
$INDRF_{t-1}$	7.539	-1.535
	(1.07)	(-1.06)
Panel C: 2000 - 2010		
PCF_{t-1}	0.741*	0.807**
	(1.95)	(2.01)
F_{t-1}		18.754***
		(2.93)
$INDRF_{t-1}$	1.827	1.715
	(1.16)	(1.02)
Panel D: 2011 - 2021		
PCF_{t-1}	-0.072	0.047
	(-0.12)	(0.25)
F_{t-1}		48.311***
		(3.69)
$INDRF_{t-1}$	30.815	5.866***
	(1.64)	(2.68)

Note: In this table are the Fama-MacBeth forecasting regressions of revisions in analyst earnings forecasts over sub-periods. The mean revision of forecasts is the dependent variable in both columns (F). The independent variables are the lagged mean revision of forecasts for the pseudo-conglomerate (PCF) in column 1, while in column 2 the lagged conglomerate mean revision and the lagged mean revision of the conglomerate's primary industry (INDRF) are added. The regressions also include the controls similar to table 4 which are size, book-to-market ratio, momentum, and turnover, measured at the end of June of each year. Every calendar month, cross-sectional regressions are run. Newey-West (1983) standard errors are used for up to 12 lags. Fama-MacBeth t-statistics are reported in parentheses. *, **, *** denote statistical significance at the 10%, 5% and 1% level, respectively.

6. Conclusion and Discussion

With this study, I examine the impact of complicated information processing on stock prices through the use of easy-to-analyse and complex firms. I investigate this for listed US firms in the period from 1977-2021. I use a similar approach as Cohen and Lou (2012), meaning that the same piece of information requires a more complex analysis for conglomerates that operate in more than one industry segment than for stand-alone firms who operate in one industry segment. This requires me to create decile portfolios with conglomerate stocks based on the past returns from its paired stand-alone firm through the overlapping industries in which they operate. This is referred to as the pseudo-conglomerate return. The strategy of going long in the portfolio of past winners and shorting the portfolio of past losers earns an average positive and significant monthly excess return of 60 basis points, which is equal to an annualised return of 7.44%. Adding controls from the Fama and French (1995) three factors, Carhart (1997) momentum and the Pastor and Stambaugh (2003) liquidity factor do not appear to reduce the estimate findings significantly.

The Fama-MacBeth forecasting regressions to indicate the relationship more clearly between the lagged pseudo-conglomerate returns and the conglomerate returns, seem to give evidence that the easy-to-analyse stand-alone firms incorporate industry information quicker than the conglomerates and their returns can therefore predict the future prices of these conglomerates. This phenomenon is supported by the works of Cohen and Lou (2012), but also by Dong et al. (2016) who find that high information processing costs are more distinct in complicated firms like conglomerate with opaque financial reporting.

Specifically looking at the complexity of the firms itself through the measure of the Herfindahl-index calculated for within a firm, I find that the effect is indeed larger for more complicated firms however the result is only significant at the 13% level. The results show a presence of limits-to-arbitrage in the sample based on the market capitalisation of firms compared to that of the NYSE median. I find that the return predictability through pseudo-conglomerates is higher for smaller firms than for larger firms, which is consistent with larger firms having less constraints from limits-to-arbitrage. The results also favour the outcome of complications in information processing being the major driver and not investors ignoring or being unaware of information as the number of analysts covering the firm and the turnover of the firm did not significantly impact the results.

Lee and So (2017) argued that analyst coverage contains invaluable insights into the expected returns. The results from table 6 show that a revision in the mean forecast for the paired pseudo-conglomerate does not significantly impact the mean forecast revision of the conglomerate until I add the lagged forecast revision of the conglomerate itself. The mean revision of the primary industry appears to consistently provide insight into the forecast revision of the conglomerate itself.

Finally, the results from the complicated processing portfolios L/S strategy are not changed when we alter the risk factors used in the model. By using the Fama and French (2015) five-factor model instead of the momentum and liquidity factor, results remain the same, indicating that more recent risk factors do not seem to incorporate the effect of complicated information processing between firms. When splitting the sample into four sub-periods, there is a difference in results. The excess returns and the regression analysis seems to indicate that this effect of pseudo-conglomerate return predictability is most present during the late 1900's and less so in the early 21st century. Explanations are that the two crises during the 2000's probably have influenced the results to an extent. However, it might also reveal that this effect is not a common characteristic of the US stock market but more so present during particular periods.

This study finds a few different results than Cohen and Lou (2012). Reasons for this include the extension of the period to 2021 instead of 2009. On the other hand, there can also be more subtle differences between the datasets used. The use of the IBES summary database instead of the IBES detail database can allow for variation in the results found, but also the filtering of data might have included removal of firms that were not removed in the original dataset. Next to this, Hou et al. (2020) have also found that around 50% of the anomalies from the literature cannot be replicated, and if they can be replicated, there is a significant decay in post-publication return predictability (McLean and Pontiff, 2016). An important issue this paper did not touch upon is the inclusion of transaction costs. All excess returns that were reported in this study were before transaction costs. Actually trading on the basis of the L/S strategy can in fact turn out to be costly as the portfolios are rebalanced every month to maintain equal (value) weights. Cohen and Lou (2012) do include a section in which they include transaction costs. They find that up to a portfolio of $10 million, transaction costs are modest, however they tend to increase when the portfolio is up to $50 million.

With this research come a few implications. First, the findings underline the significance of complicated information processing in understanding return predictability. The risk factors that I include in these tests do not seem to have an impact on the effect found. Besides this, I find some temporal variations. The patterns that emerged, specifically the differences between the late 1900's and the early 2000's, appear to suggest that certain stock market characteristics might be context or period-specific and not a general phenomenon which is usually argued in scientific papers. Lastly, the replicability of certain papers is questioned. Black (1993) already spoke of data mining and the risk of p-hacking is a serious problem according to Chordia, Goyal and Saretto (2017).

As this paper focuses on the technical aspects of information processing, it might be interesting to explore the behavioural side, as to the difference between how individual and institutional investors act when given simple or complex information. Additionally, in order to combat data mining or questionable replicability, the sample can be expanded beyond the US. Future studies could replicate this research in other global markets like the EU. Finally, with the rise of the XBRL language and other technological innovations in data analytics tools, it might be interesting to see the impacts of these advanced tools like machine learning on complicated information processing in the future.

7. References

Ali, U., & Hirshleifer, D. (2020). Shared analyst coverage: Unifying momentum spillover effects. *Journal of Financial Economics*, *136*(3), 649-675.

Amihud, Y., & Mendelson, H. (1986). Asset pricing and the bid-ask spread. *Journal of financial Economics*, *17*(2), 223-249.

Barber, B. M., & Odean, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *The review of financial studies*, *21*(2), 785-818.

Barberis, N. (2018). Psychology-based models of asset prices and trading volume. In *Handbook of behavioral economics: applications and foundations 1* (Vol. 1, pp. 79-175). North-Holland.

Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. *Journal of financial economics*, *49*(3), 307-343.

Barinov, A., Park, S. S., & Yıldızhan, Ç. (2022). Firm complexity and post-earnings announcement drift. *Review of Accounting Studies*, 1-53.

Bernard, V. L., & Thomas, J. K. (1989). Post-earnings-announcement drift: delayed price response or risk premium?. *Journal of Accounting research*, *27*, 1-36.

Black, F. (1993). Estimating expected return. *Financial analysts journal*, *49*(5), 36-38.

Blankespoor, E., deHaan, E., & Marinovic, I. (2020). Disclosure processing costs, investors' information choice, and equity market outcomes: A review. *Journal of Accounting and Economics*, *70*(2-3), 101344.

Brennan, M. J., Chordia, T., & Subrahmanyam, A. (1998). Alternative factor specifications, security characteristics, and the cross-section of expected stock returns. *Journal of financial Economics*, *49*(3), 345-373.

Cao, J., Chordia, T., & Lin, C. (2016). Alliances and return predictability. *Journal of Financial and Quantitative Analysis*, *51*(5), 1689-1717.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of finance*, *52*(1), 57 82.

Chan, K., & Hameed, A. (2006). Stock price synchronicity and analyst coverage in emerging markets. *Journal of Financial Economics*, *80*(1), 115-147.

Chordia, T., Goyal, A., & Saretto, A. (2017). In Defense of Market Efficiency: Evidence from Two Million Strategies.

Chordia, T., Roll, R., & Subrahmanyam, A. (2000). Commonality in liquidity. *Journal of financial economics*, *56*(1), 3-28.

Chordia, T., Roll, R., & Subrahmanyam, A. (2001). Market liquidity and trading activity. *The journal of finance*, *56*(2), 501-530.

Cohen, L., & Frazzini, A. (2008). Economic links and predictable returns. *The Journal of Finance*, *63*(4), 1977-2011.

Cohen, L., & Lou, D. (2012). Complicated firms. *Journal of financial economics*, *104*(2), 383-400.

Daniel, K., & Moskowitz, T. J. (2016). Momentum crashes. *Journal of Financial economics*, *122*(2), 221 247.

Das, S., Guo, R. J., & Zhang, H. (2006). Analysts' selective coverage and subsequent performance of newly public firms. *The Journal of Finance*, *61*(3), 1159-1185.

DellaVigna, S., & Pollet, J. M. (2009). Investor inattention and Friday earnings announcements. *The journal of finance*, *64*(2), 709-749.

Dolde, W., & Mishra, D. R. (2002). Firm complexity and FX derivatives use. *Available at SSRN 302813*.

Dong, Y., Li, O. Z., Lin, Y., & Ni, C. (2016). Does information-processing cost affect firm-specific information acquisition? Evidence from XBRL adoption. *Journal of Financial and Quantitative Analysis*, *51*(2), 435-462.

Doukas, J. A., Kim, C., & Pantzalis, C. (2005). The two faces of analyst coverage. *Financial Management*, *34*(2), 99-125.

Duffie, D. (2010). Presidential address: Asset price dynamics with slow-moving capital. *The Journal of finance*, *65*(4), 1237-1267.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, *25*(2), 383-417.

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *the Journal of Finance*, *47*(2), 427-465.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, *33*(1), 3-56.

Fama, E. F., & French, K. R. (1995). Size and book-to-market factors in earnings and returns. *The journal of finance*, *50*(1), 131-155.

Fama, E. F., & French, K. R. (1996). Multifactor explanations of asset pricing anomalies. *The journal of finance*, *51*(1), 55-84.

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of financial economics*, *116*(1), 1-22.

Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of political economy*, *81*(3), 607-636.

Frankel, R., Kothari, S. P., & Weber, J. (2006). Determinants of the informativeness of analyst research. *Journal of Accounting and Economics*, *41*(1-2), 29-54.

Foltice, B., & Langer, T. (2015). Profitable momentum trading strategies for individual investors. *Financial Markets and Portfolio Management*, *29*, 85-113.

Herfindahl, O. C. (1997). *Concentration in the steel industry*. Columbia University.

Hirschman, A. O. (1980). *National power and the structure of foreign trade* (Vol. 105). Univ of California Press.

Hou, K., Xiong, W., & Peng, L. (2006). R2 and price inefficiency. *Fisher College of Business Working Paper*, (2006-03), 007.

Hou, K., Xue, C., & Zhang, L. (2020). Replicating anomalies. *The Review of financial studies*, *33*(5), 2019-2133.

Huang, Y., Shan, Y. G., & Yang, J. W. (2021). Information processing costs and stock price informativeness: Evidence from the XBRL mandate. *Australian Journal of Management*, *46*(1), 110-131.

Jegadeesh, N. (1990). Evidence of predictable behavior of security returns. *The Journal of finance*, *45*(3), 881-898.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of finance*, *48*(1), 65-91.

Kelly, B., & Ljungqvist, A. (2012). Testing asymmetric-information asset pricing models. *The Review of Financial Studies*, *25*(5), 1366-1413.

Kim, J. B., Li, B., & Liu, Z. (2019). Information-processing costs and breadth of ownership. *Contemporary Accounting Research*, *36*(4), 2408-2436.

Loh, R. K. (2010). Investor inattention and the underreaction to stock recommendations. *Financial management*, *39*(3), 1223-1252.

Lee, C. M., & So, E. C. (2017). Uncovering expected returns: Information in analyst coverage proxies. *Journal of Financial Economics*, *124*(2), 331-348.

Lee, C. M., Sun, S. T., Wang, R., & Zhang, R. (2019). Technological links and predictable returns. *Journal of Financial Economics*, *132*(3), 76-96.

Lintner, J. (1965). Security prices, risk, and maximal gains from diversification. *The journal of finance*, *20*(4), 587-615.

Lys, T., & Sohn, S. (1990). The association between revisions of financial analysts' earnings forecasts and security-price changes. *Journal of Accounting and Economics*, *13*(4), 341-363.

Markowitz, H. M. (1959). Portfolio Selection: Efficient Diversification of Investments. *Yale University Press.* http://www.jstor.org/stable/j.ctt1bh4c8h

Marquering, W., Nisser, J., & Valla, T. (2006). Disappearing anomalies: a dynamic analysis of the persistence of anomalies. *Applied financial economics*, *16*(4), 291-302.

McLean, R. D., & Pontiff, J. (2016). Does academic research destroy stock return predictability?. *The Journal of Finance*, *71*(1), 5-32.

McNichols, M., & O'Brien, P. C. (1997). Self-selection and analyst coverage. *Journal of accounting research*, *35*, 167-199.

Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information.

Miller, B. P. (2010). The effects of reporting complexity on small and large investor trading. *The Accounting Review*, *85*(6), 2107-2143.

Moskowitz, T. J., & Grinblatt, M. (1999). Do industries explain momentum?. *The Journal of finance*, *54*(4), 1249-1290.

Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica: Journal of the econometric society*, 768-783.

Müller, S. (2019). Economic links and cross-predictability of stock returns: Evidence from characteristic based “styles”. *Review of Finance*, *23*(2), 363-395.

Newey, W. K., & West, K. D. (1987). Hypothesis testing with efficient method of moments estimation. *International Economic Review*, 777-787.

Pástor, Ľ., & Stambaugh, R. F. (2003). Liquidity risk and expected stock returns. *Journal of Political economy*, *111*(3), 642-685.

Peng, L., Titman, S., Yönaç, M., & Zhou, D. (2022). Social Ties, Comovements, and Predictable Returns. *Comovements, and Predictable Returns (July 29, 2022)*.

Scherbina, A. (2008). Suppressed negative information and future underperformance. *Review of Finance*, *12*(3), 533-565.

Scherbina, A., & Schlusche, B. (2015). Economic linkages inferred from news stories and the predictability of stock returns. *Available at SSRN 2363436*.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The journal of finance*, *19*(3), 425-442.

Shiller, R. J. (1981). Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?. *American Economic Review, American Economic Association, 71*(3), 421-436.

Shleifer, A., & Vishny, R. W. (1997). The limits of arbitrage. *The Journal of finance*, *52*(1), 35-55.

Sidhu, B., & Tan, H. C. (2011). The performance of equity analysts during the global financial crisis. *Australian Accounting Review*, *21*(1), 32-43.

Siganos, A. (2007). Momentum returns and size of winner and loser portfolios. *Applied Financial Economics*, *17*(9), 701-708.

Thaler, R. H. (1999). Mental accounting matters. *Journal of Behavioral decision making*, *12*(3), 183-206.

Van Nieuwerburgh, S., & Veldkamp, L. (2009). Information immobility and the home bias puzzle. *The Journal of Finance*, *64*(3), 1187-1215.

Veldkamp, L. L. (2006). Information markets and the comovement of asset prices. *The Review of Economic Studies*, *73*(3), 823-845.

Wheale, P. R., & Amin, L. H. (2003). Bursting the dot. com" bubble': a case study in investor behaviour. *Technology Analysis & Strategic Management*, *15*(1), 117-136.

You, H., & Zhang, X. J. (2009). Financial reporting complexity and investor underreaction to 10-K information. *Review of Accounting studies*, *14*, 559-586.

Appendix

A. Five-factor model with momentum and liquidity

Table A.1: Complicated processing portfolios, 1977 – 1988

Decile	Excess Returns	1-Factor Alpha	3-Factor Alpha	4-Factor Alpha	5-Factor Alpha
Panel A: Equal weights					
1	1.00%	0.34%	0.09%	0.22%	0.13%
	(1.83)	(1.32)	(0.47)	(1.13)	(0.51)
2	0.76%	0.12%	-0.06%	0.04%	0.10%
	(1.51)	(0.63)	(-0.45)	(0.27)	(0.55)
3	0.86%	0.24%	-0.00%	0.08%	-0.20%
	(1.75)	(1.29)	(-0.03)	(0.68)	(-1.60)
4	0.88%	0.24%	-0.00%	0.06%	-0.18%
	(1.75)	(1.30)	(-0.03)	(0.49)	(-1.27)
5	1.12%	0.50%	0.30%	0.33%	0.37%
	(2.30)	(3.01)	(3.10)	(3.38)	(3.09)
6	1.00%	0.37%	0.13%	0.16%	0.38%
	(2.03)	(2.02)	(1.33)	(1.53)	(3.01)
7	1.10%	0.45%	0.19%	0.19%	0.41%
	(2.15)	(2.31)	(1.93)	(1.83)	(3.43)
8	1.27%	0.65%	0.38%	0.33%	0.31%
	(2.59)	(3.54)	(3.42)	(3.10)	(2.17)
9	1.37%	0.73%	0.43%	0.39%	0.35%
	(2.74)	(4.07)	(4.59)	(4.18)	(2.77)
10	2.16%	1.49%	1.28%	1.11%	1.05%
	(4.06)	(6.85)	(6.92)	(6.80)	(5.33)
Panel B: Value weights					
1	0.49%	-0.16%	-0.21%	-0.10%	-0.16%
	(0.95)	(-0.44)	(-0.75)	(-0.37)	(-0.47)
2	0.76%	0.18%	0.27%	0.41%	0.39%
	(1.63)	(0.93)	(1.36)	(2.14)	(1.70)
3	0.81%	0.24%	0.28%	0.36%	0.27%
	(1.81)	(1.49)	(1.71)	(2.25)	(1.28)
4	0.96%	0.37%	0.39%	0.46%	0.39%
	(2.07)	(2.31)	(2.21)	(2.50)	(1.74)
5	0.97%	0.39%	0.49%	0.52%	0.75%
	(2.18)	(2.95)	(3.31)	(3.56)	(4.60)
6	0.97%	0.35%	0.42%	0.43%	0.60%
	(2.03)	(2.22)	(2.27)	(2.20)	(2.59)
7	1.14%	0.55%	0.57%	0.58%	0.86%
	(2.46)	(3.43)	(3.55)	(3.33)	(4.33)
8	1.30%	0.72%	0.70%	0.64%	0.80%
	(2.90)	(4.83)	(4.52)	(4.19)	(4.24)
9	1.50%	0.86%	0.89%	0.80%	0.78%
	(3.03)	(5.56)	(4.75)	(5.14)	(3.47)
10	1.76%	1.13%	1.21%	0.96%	0.58%
	(3.41)	(4.77)	(4.70)	(4.23)	(1.97)

Table A.2: Complicated Processing Portfolios, 1989 – 1999

Decile	Excess Returns	1-Factor Alpha	3-Factor Alpha	4-Factor Alpha	5-Factor Alpha
Panel A: Equal weights					
1	0.90%	-0.19%	-0.05%	0.21%	0.16%
	(2.14)	(-0.72)	(-0.27)	(1.04)	(0.78)
2	0.48%	-0.56%	-0.49%	-0.36%	-0.34%
	(1.27)	(-2.61)	(-3.98)	(-2.94)	(-2.72)
3	0.64%	-0.38%	-0.31%	-0.06%	-0.03%
	(1.68)	(-1.80)	(-1.99)	(-0.37)	(-0.20)
4	0.71%	-0.27%	-0.19%	0.09%	0.13%
	(1.96)	(-1.30)	(-1.41)	(0.71)	(1.09)
5	0.66%	-0.28%	-0.20%	0.01%	0.01%
	(1.88)	(-1.32)	(-1.62)	(0.06)	(0.12)
6	0.92%	-0.06%	0.02%	0.20%	0.25%
	(2.56)	(-0.35)	(0.18)	(1.51)	(1.79)
7	1.06%	0.10%	0.18%	0.27%	0.29%
	(3.10)	(0.56)	(1.54)	(2.25)	(2.44)
8	1.44%	0.48%	0.54%	0.59%	0.63%
	(4.11)	(2.46)	(4.28)	(4.45)	(4.66)
9	1.50%	0.57%	0.72%	0.69%	0.73%
	(4.33)	(2.94)	(4.27)	(3.95)	(4.05)
10	2.34%	1.18%	1.40%	1.22%	1.20%
	(5.14)	(4.44)	(5.85)	(4.87)	(4.23)
Panel B: Value weights					
1	1.05%	-0.01%	-0.09%	0.17%	0.12%
	(2.65)	(-0.06)	(-0.40)	(0.65)	(0.47)
2	1.21%	0.14%	0.01%	0.09%	0.13%
	(3.11)	(0.64)	(0.05)	(0.43)	(0.58)
3	1.35%	0.24%	0.16%	0.40%	0.36%
	(3.38)	(1.18)	(0.77)	(1.77)	(1.41)
4	1.16%	0.21%	0.15%	0.33%	0.39%
	(3.20)	(0.93)	(0.66)	(1.32)	(1.62)
5	1.35%	0.34%	0.32%	0.50%	0.43%
	(3.56)	(1.44)	(1.49)	(2.22)	(2.00)
6	1.54%	0.52%	0.47%	0.57%	0.63%
	(4.12)	(2.66)	(2.30)	(2.66)	(2.93)
7	1.40%	0.41%	0.31%	0.45%	0.42%
	(4.00)	(2.44)	(1.91)	(2.50)	(2.08)
8	1.63%	0.58%	0.49%	0.51%	0.52%
	(4.25)	(2.76)	(2.29)	(2.28)	(2.21)
9	1.65%	0.66%	0.72%	0.33%	0.30%
	(4.29)	(3.02)	(2.90)	(1.35)	(1.14)
10	2.14%	0.99%	1.03%	0.74%	0.76%
	(4.93)	(3.75)	(3.72)	(2.68)	(2.52)

Table A.3: Complicated Processing Portfolios, 2000 – 2010

Decile	Excess Returns	1-Factor Alpha	3-Factor Alpha	4-Factor Alpha	5-Factor Alpha
Panel A: Equal weights					
1	2.13%	2.12%	1.61%	1.62%	1.64%
	(3.07)	(5.36)	(4.74)	(5.32)	(5.10)
2	0.95%	0.95%	0.65%	0.65%	0.67%
	(1.73)	(3.66)	(2.91)	(3.01)	(2.91)
3	1.18%	1.18%	0.81%	0.81%	0.81%
	(2.22)	(5.35)	(4.83)	(4.91)	(5.15)
4	1.10%	1.10%	0.76%	0.76%	0.74%
	(2.23)	(5.53)	(4.74)	(4.80)	(4.52)
5	1.14%	1.14%	0.72%	0.72%	0.92%
	(2.32)	(5.14)	(4.06)	(4.11)	(5.20)
6	1.12%	1.12%	0.71%	0.72%	0.90%
	(2.39)	(5.32)	(4.09)	(4.14)	(5.57)
7	1.19%	1.19%	0.73%	0.73%	0.84%
	(2.47)	(4.92)	(4.00)	(4.18)	(4.45)
8	1.51%	1.51%	1.06%	1.06%	1.13%
	(3.17)	(6.14)	(6.16)	(6.37)	(5.48)
9	1.42%	1.42%	0.92%	0.92%	1.10%
	(2.63)	(4.87)	(4.60)	(4.56)	(5.50)
10	1.77%	1.76%	1.31%	1.31%	1.59%
	(3.26)	(5.73)	(5.30)	(5.28)	(6.24)
Panel B: Value weights					
1	1.13%	1.13%	1.27%	1.27%	0.90%
	(2.07)	(3.07)	(3.49)	(3.62)	(2.36)
2	0.59%	0.59%	0.66%	0.67%	0.53%
	(1.10)	(2.00)	(2.24)	(2.28)	(1.65)
3	0.69%	0.69%	0.71%	0.71%	0.29%
	(1.35)	(2.58)	(2.80)	(2.79)	(1.15)
4	1.16%	1.16%	1.19%	1.18%	1.15%
	(2.53)	(4.56)	(5.34)	(5.30)	(4.64)
5	0.84%	0.84%	0.83%	0.83%	0.89%
	(1.82)	(3.10)	(3.26)	(3.23)	(3.43)
6	1.09%	1.09%	1.03%	1.03%	0.97%
	(2.86)	(5.26)	(5.42)	(5.40)	(5.02)
7	0.65%	0.65%	0.59%	0.59%	0.60%
	(1.49)	(2.58)	(2.62)	(2.61)	(2.70)
8	0.67%	0.67%	0.68%	0.68%	0.90%
	(1.53)	(2.80)	(3.17)	(3.15)	(3.63)
9	0.91%	0.90%	0.71%	0.71%	0.88%
	(1.76)	(2.71)	(2.46)	(2.47)	(2.92)
10	0.71%	0.71%	0.63%	0.63%	0.76%
	(1.36)	(2.05)	(2.09)	(2.07)	(2.38)

Table A.4: Complicated Processing Portfolios, 2011 - 2021

Decile	Excess Returns	1-Factor Alpha	3-Factor Alpha	4-Factor Alpha	5-Factor Alpha
Panel A: Equal weights					
1	1.53%	-0.18%	0.19%	0.25%	0.28%
	(2.63)	(-0.58)	(0.87)	(1.16)	(1.28)
2	1.16%	-0.36%	-0.10%	-0.07%	-0.04%
	(2.37)	(-1.74)	(-0.73)	(-0.49)	(-0.29)
3	1.23%	-0.15%	0.11%	0.13%	0.09%
	(2.74)	(-0.76)	(0.83)	(0.93)	(0.70)
4	1.32%	-0.07%	0.24%	0.25%	0.27%
	(2.88)	(-0.28)	(1.55)	(1.64)	(1.82)
5	1.28%	-0.19%	0.14%	0.15%	0.19%
	(2.62)	(-0.73)	(1.01)	(1.03)	(1.33)
6	1.41%	0.04%	0.30%	0.31%	0.29%
	(3.16)	(0.22)	(2.25)	(2.28)	(2.12)
7	1.39%	0.02%	0.26%	0.28%	0.24%
	(3.15)	(0.12)	(2.44)	(2.54)	(2.29)
8	1.29%	-0.06%	0.15%	0.16%	0.20%
	(2.95)	(-0.37)	(1.23)	(1.35)	(1.70)
9	1.39%	0.06%	0.34%	0.37%	0.34%
	(3.14)	(0.35)	(2.52)	(2.74)	(2.59)
10	1.65%	0.28%	0.60%	0.68%	0.70%
	(3.48)	(1.14)	(3.20)	(3.74)	(3.65)
Panel B: Value weights					
1	1.06%	-0.25%	-0.17%	-0.09%	-0.10%
	(2.41)	(-1.24)	(-0.86)	(-0.45)	(-0.47)
2	1.50%	0.29%	0.34%	0.33%	0.33%
	(3.76)	(1.71)	(1.97)	(1.84)	(1.79)
3	1.70%	0.57%	0.61%	0.61%	0.53%
	(4.48)	(3.18)	(3.62)	(3.44)	(3.20)
4	1.57%	0.37%	0.43%	0.40%	0.48%
	(3.92)	(1.92)	(2.12)	(2.00)	(2.41)
5	1.29%	0.07%	0.21%	0.24%	0.20%
	(3.23)	(0.40)	(1.23)	(1.45)	(1.25)
6	1.68%	0.39%	0.47%	0.46%	0.49%
	(3.93)	(1.88)	(2.75)	(2.63)	(2.63)
7	1.38%	0.17%	0.19%	0.17%	0.14%
	(3.47)	(1.01)	(1.10)	(1.01)	(0.84)
8	1.29%	0.09%	0.12%	0.16%	0.14%
	(3.25)	(0.59)	(0.77)	(0.94)	(0.86)
9	1.39%	0.27%	0.28%	0.27%	0.23%
	(3.54)	(1.35)	(1.31)	(1.21)	(1.06)
10	1.51%	0.41%	0.45%	0.51%	0.50%
	(3.81)	(1.63)	(1.82)	(2.05)	(1.97)

Table A.5: Complicated Processing Portfolios, factor loadings 1977 - 1988

Decile	Excess Returns	5-Factor Alpha	MKT	SMB	HML	MOM	LIQ
Panel A: Equal weights							
1	1.00%	0.13%	1.072	0.869	-0.007	-0.166	-0.026
	(1.83)	(0.51)	(21.34)	(9.29)	(-0.07)	(-2.64)	(-0.66)
2	0.76%	0.10%	1.017	0.678	-0.037	-0.135	0.018
	(1.51)	(0.55)	(26.62)	(10.51)	(-0.54)	(-2.95)	(0.62)
3	0.86%	-0.20%	1.056	0.776	0.076	-0.105	-0.078
	(1.75)	(-1.60)	(35.57)	(18.40)	(1.66)	(-3.55)	(-1.60)
4	0.88%	-0.18%	1.063	0.785	0.079	-0.078	-0.067
	(1.75)	(-1.27)	(38.96)	(18.13)	(1.82)	(-2.85)	(-1.27)
5	1.12%	0.37%	1.008	0.648	0.018	-0.048	0.009
	(2.30)	(3.09)	(39.41)	(13.16)	(0.41)	(-1.88)	(0.46)
6	1.00%	0.38%	0.984	0.697	0.034	-0.039	0.060
	(2.03)	(3.01)	(44.12)	(15.30)	(0.86)	(-1.51)	(2.94)
7	1.10%	0.41%	1.000	0.755	0.047	-0.005	0.060
	(2.15)	(3.43)	(28.70)	(17.33)	(0.87)	(-0.13)	(2.50)
8	1.27%	0.31%	1.025	0.702	0.157	0.056	-0.008
	(2.59)	(2.17)	(33.45)	(13.73)	(2.99)	(1.48)	(-0.32)
9	1.37%	0.35%	1.063	0.704	0.195	0.054	-0.010
	(2.74)	(2.77)	(35.29)	(16.08)	(4.38)	(1.68)	(-0.48)
10	2.16%	1.05%	1.062	0.646	0.123	0.212	-0.017
	(4.06)	(5.33)	(20.37)	(9.28)	(1.79)	(4.07)	(-0.46)
L/S	1.16%***	0.92%***	-0.010	-0.224*	0.130	0.378***	0.008
	(4.33)	(2.71)	(-0.14)	(-1.85)	(0.98)	(4.19)	(0.16)
Panel B: Value weights							
1	0.49%	-0.16%	1.097	0.172	0.061	-0.138	-0.017
	(0.95)	(-0.47)	(17.36)	(1.22)	(0.51)	(-1.54)	(-0.34)
2	0.76%	0.39%	1.006	0.033	-0.210	-0.180	-0.005
	(1.63)	(1.70)	(20.53)	(0.36)	(-2.90)	(-3.09)	(-0.10)
3	0.81%	0.27%	1.043	-0.056	-0.056	-0.104	-0.027
	(1.81)	(1.28)	(23.50)	(-0.83)	(-0.72)	(-2.05)	(-0.70)
4	0.96%	0.39%	1.063	0.023	-0.053	-0.089	-0.020
	(2.07)	(1.74)	(20.45)	(0.41)	(-0.60)	(-1.72)	(-0.53)
5	0.97%	0.75%	0.956	-0.092	-0.179	-0.031	0.065
	(2.18)	(4.60)	(28.36)	(-1.67)	(-2.59)	(-0.65)	(2.93)
6	0.97%	0.60%	1.040	-0.040	-0.128	-0.009	0.047
	(2.03)	(2.59)	(21.60)	(-0.46)	(-1.27)	(-0.12)	(1.40)
7	1.14%	0.86%	0.985	-0.000	-0.094	-0.018	0.078
	(2.46)	(4.33)	(22.32)	(-0.00)	(-0.83)	(-0.36)	(2.59)
8	1.30%	0.80%	0.988	-0.008	0.035	0.069	0.044
	(2.90)	(4.24)	(20.17)	(-0.11)	(0.40)	(1.17)	(1.18)
9	1.50%	0.78%	1.146	-0.163	0.059	0.116	-0.009
	(3.03)	(3.47)	(33.40)	(-1.81)	(0.71)	(1.68)	(-0.31)
10	1.76%	0.58%	1.147	-0.162	0.083	0.321	-0.107
	(3.41)	(1.97)	(15.80)	(-1.63)	(0.99)	(4.03)	(-2.38)
L/S	1.27%***	0.74%	0.049	-0.334*	0.022	0.458***	-0.091
	(3.23)	(1.52)	(0.45)	(-1.66)	(0.14)	(3.40)	(-1.17)

Table A.6: Complicated Processing Portfolios, factor loadings 1989 - 1999

Decile	Excess Returns	5-Factor Alpha	MKT	SMB	HML	MOM	LIQ
Panel A: Equal weights							
1	0.90%	0.16%	1.012	0.647	0.242	-0.226	-0.043
	(2.14)	(0.78)	(20.46)	(8.78)	(2.86)	(-3.46)	(-1.07)
2	0.48%	-0.34%	1.007	0.544	0.382	-0.115	0.013
	(1.27)	(-2.72)	(29.41)	(12.71)	(6.39)	(-3.17)	(0.59)
3	0.64%	-0.03%	0.981	0.501	0.311	-0.212	0.022
	(1.68)	(-0.20)	(25.32)	(7.06)	(4.16)	(-2.83)	(0.62)
4	0.71%	0.13%	0.942	0.511	0.289	-0.241	0.037
	(1.96)	(1.09)	(33.45)	(9.70)	(5.63)	(-5.96)	(1.26)
5	0.66%	0.01%	0.914	0.548	0.341	-0.181	0.006
	(1.88)	(0.12)	(29.95)	(12.68)	(5.33)	(-4.48)	(0.26)
6	0.92%	0.25%	0.934	0.533	0.292	-0.149	0.042
	(2.56)	(1.79)	(29.32)	(11.43)	(4.69)	(-3.01)	(1.16)
7	1.06%	0.29%	0.914	0.482	0.277	-0.079	0.015
	(3.10)	(2.44)	(26.32)	(10.08)	(5.23)	(-1.89)	(0.44)
8	1.44%	0.63%	0.936	0.520	0.369	-0.040	0.034
	(4.11)	(4.66)	(31.74)	(11.03)	(7.26)	(-0.97)	(1.49)
9	1.50%	0.73%	0.817	0.574	0.128	0.031	0.030
	(4.33)	(4.05)	(13.31)	(8.93)	(1.84)	(0.52)	(1.34)
10	2.34%	1.20%	1.001	0.723	0.050	0.157	-0.012
	(5.14)	(4.23)	(14.88)	(6.80)	(0.47)	(1.35)	(-0.17)
L/S	1.45%***	1.05%***	-0.011	0.077	-0.192	0.383***	0.031
	(4.69)	(3.17)	(-0.12)	(0.57)	(-1.34)	(2.86)	(0.59)
Panel B: Value weights							
1	1.05%	0.12%	1.022	-0.136	0.096	-0.222	-0.043
	(2.65)	(0.47)	(15.80)	(-1.34)	(0.90)	(-2.24)	(-0.93)
2	1.21%	0.13%	1.075	-0.117	0.318	-0.068	0.030
	(3.11)	(0.58)	(16.69)	(-1.13)	(2.95)	(-1.02)	(0.60)
3	1.35%	0.36%	1.061	-0.180	0.090	-0.217	-0.037
	(3.38)	(1.41)	(16.35)	(-2.17)	(0.98)	(-2.26)	(-0.58)
4	1.16%	0.39%	0.905	-0.058	0.105	-0.146	0.050
	(3.20)	(1.62)	(14.33)	(-0.63)	(1.07)	(-1.80)	(0.76)
5	1.35%	0.43%	0.987	0.153	0.228	-0.159	-0.061
	(3.56)	(2.00)	(17.98)	(1.74)	(2.30)	(-1.60)	(-1.28)
6	1.54%	0.63%	0.975	0.026	0.211	-0.079	0.054
	(4.12)	(2.93)	(14.39)	(0.24)	(2.14)	(-0.71)	(0.99)
7	1.40%	0.42%	0.969	-0.186	0.163	-0.125	-0.021
	(4.00)	(2.08)	(20.68)	(-3.19)	(2.23)	(-1.93)	(-0.49)
8	1.63%	0.52%	1.002	-0.173	0.134	-0.015	0.015
	(4.25)	(2.21)	(17.24)	(-2.23)	(1.27)	(-0.19)	(0.29)
9	1.65%	0.30%	0.862	0.171	0.012	0.337	-0.028
	(4.29)	(1.14)	(12.42)	(1.50)	(0.11)	(2.59)	(-0.53)
10	2.14%	0.76%	1.017	0.139	0.016	0.249	0.014
	(4.93)	(2.52)	(11.94)	(1.14)	(0.14)	(2.02)	(0.24)
L/S	1.08%***	0.65%	-0.005	0.276*	-0.080	0.471***	0.057
	(2.90)	(1.60)	(-0.05)	(1.73)	(-0.47)	(2.70)	(0.76)

Table A.7: Complicated Processing Portfolios, factor loadings 2000 - 2010

Decile	Excess Returns	5-Factor Alpha	MKT	SMB	HML	MOM	LIQ
Panel A: Equal weights							
1	2.13%	1.64%	1.047	0.676	0.222	-0.338	-0.007
	(3.07)	(5.10)	(10.86)	(4.27)	(1.49)	(-3.21)	(0.16)
2	0.95%	0.67%	1.026	0.396	0.106	-0.087	0.006
	(1.73)	(2.91)	(14.16)	(2.91)	(0.93)	(-1.35)	(0.17)
3	1.18%	0.81%	1.059	0.359	0.263	-0.057	-0.001
	(2.22)	(5.15)	(19.21)	(3.45)	(2.99)	(-1.10)	(-0.02)
4	1.10%	0.74%	1.006	0.311	0.268	-0.037	-0.004
	(2.23)	(4.52)	(19.56)	(4.89)	(3.89)	(-0.94)	(-0.16)
5	1.14%	0.92%	0.923	0.423	0.285	-0.060	0.047
	(2.32)	(5.20)	(16.73)	(5.01)	(4.05)	(-1.58)	(2.04)
6	1.12%	0.90%	0.874	0.446	0.232	-0.052	0.046
	(2.39)	(5.57)	(16.73)	(8.03)	(4.63)	(-1.75)	(1.48)
7	1.19%	0.84%	0.833	0.490	0.291	-0.123	0.027
	(2.47)	(4.45)	(16.15)	(5.34)	(4.92)	(-2.83)	(1.09)
8	1.51%	1.13%	0.816	0.543	0.206	-0.095	0.017
	(3.17)	(5.48)	(15.02)	(7.84)	(2.90)	(-2.26)	(0.60)
9	1.42%	1.10%	0.917	0.590	0.248	-0.081	0.045
	(2.63)	(5.50)	(13.37)	(4.78)	(2.28)	(-0.90)	(1.38)
10	1.77%	1.59%	0.897	0.540	0.218	-0.072	0.068
	(3.26)	(6.24)	(12.84)	(7.02)	(1.93)	(-1.16)	(1.62)
L/S	-0.35%	-0.05%	-0.149	-0.136	-0.004	0.266**	0.062
	(-0.79)	(-0.14)	(-1.21)	(-0.75)	(-0.02)	(2.27)	(1.09)
Panel B: Value weights							
1	1.13%	0.90%	0.906	-0.103	-0.090	-0.163	-0.088
	(2.07)	(2.36)	(8.35)	(-0.59)	(-0.53)	(-1.11)	(-1.58)
2	0.59%	0.53%	1.038	-0.089	-0.009	-0.083	-0.033
	(1.10)	(1.65)	(11.45)	(-0.57)	(-0.06)	(-0.96)	(-0.63)
3	0.69%	0.29%	1.142	-0.249	0.260	0.020	-0.102
	(1.35)	(1.15)	(14.78)	(-2.42)	(2.75)	(0.36)	(-2.24)
4	1.16%	1.15%	1.004	-0.291	0.267	0.072	-0.007
	(2.53)	(4.64)	(18.11)	(-4.30)	(3.30)	(1.99)	(-0.24)
5	0.84%	0.89%	0.949	-0.180	0.209	0.064	0.015
	(1.82)	(3.43)	(11.46)	(-1.30)	(1.71)	(1.45)	(0.51)
6	1.09%	0.97%	0.784	-0.081	0.214	-0.006	-0.014
	(2.86)	(5.02)	(15.33)	(-1.14)	(3.13)	(-0.15)	(-0.48)
7	0.65%	0.60%	0.828	-0.097	0.228	-0.070	0.003
	(1.49)	(2.70)	(13.01)	(-1.06)	(3.04)	(-1.13)	(0.09)
8	0.67%	0.90%	0.846	-0.092	0.088	-0.035	0.053
	(1.53)	(3.63)	(13.27)	(-0.88)	(0.93)	(-0.52)	(1.40)
9	0.91%	0.88%	0.908	0.160	0.149	0.037	0.041
	(1.76)	(2.92)	(10.11)	(0.85)	(0.86)	(0.37)	(1.04)
10	0.71%	0.76%	0.932	-0.148	0.303	-0.030	0.031
	(1.36)	(2.38)	(13.26)	(-1.26)	(3.15)	(-0.30)	(0.73)
L/S	-0.42%	-0.14%	0.026	-0.045	0.394*	0.133	0.119
	(-0.81)	(-0.29)	(0.18)	(-0.20)	(1.80)	(0.63)	(1.56)

Table A.8: Complicated Processing Portfolios, factor loadings 2011 - 2021

Decile	Excess Returns	5-Factor Alpha	MKT	SMB	HML	MOM	LIQ
Panel A: Equal weights							
1	1.53%	0.28%	1.168	0.516	0.308	-0.207	0.033
	(2.63)	(1.28)	(16.55)	(4.21)	(3.29)	(-2.02)	(0.79)
2	1.16%	-0.04%	1.089	0.362	0.240	-0.104	0.030
	(2.37)	(-0.29)	(28.20)	(6.20)	(4.15)	(-2.59)	(1.15)
3	1.23%	0.09%	0.996	0.450	0.238	-0.037	-0.037
	(2.74)	(0.70)	(23.51)	(7.53)	(4.07)	(-0.82)	(-1.35)
4	1.32%	0.27%	0.963	0.490	0.280	-0.040	0.027
	(2.88)	(1.82)	(21.85)	(6.91)	(5.06)	(-0.92)	(0.91)
5	1.28%	0.19%	1.005	0.558	0.303	-0.025	0.039
	(2.62)	(1.33)	(20.57)	(9.88)	(4.99)	(-0.55)	(1.26)
6	1.41%	0.29%	0.991	0.433	0.235	-0.033	-0.025
	(3.16)	(2.12)	(29.35)	(7.02)	(4.48)	(-0.73)	(-0.98)
7	1.39%	0.24%	0.990	0.473	0.186	-0.032	-0.035
	(3.15)	(2.29)	(30.03)	(10.03)	(3.59)	(-0.92)	(-1.56)
8	1.29%	0.20%	0.976	0.390	0.134	-0.056	0.040
	(2.95)	(1.70)	(28.85)	(7.04)	(2.02)	(-1.28)	(1.56)
9	1.39%	0.34%	0.918	0.522	0.169	-0.110	-0.031
	(3.14)	(2.59)	(22.96)	(7.55)	(2.42)	(-1.58)	(-1.03)
10	1.65%	0.70%	0.874	0.594	0.126	-0.248	0.017
	(3.48)	(3.65)	(17.84)	(7.83)	(2.05)	(-3.62)	(0.48)
L/S	0.12%	0.42%	-0.295***	0.078	-0.182	-0.041	-0.016
	(0.43)	(1.38)	(-3.43)	(0.53)	(-1.64)	(-0.33)	(-0.28)
Panel B: Value weights							
1	1.06%	-0.10%	0.982	0.060	-0.046	-0.270	-0.007
	(2.41)	(-0.47)	(17.72)	(0.65)	(-0.55)	(-3.38)	(-0.19)
2	1.50%	0.33%	1.034	-0.162	0.246	-0.038	-0.001
	(3.76)	(1.79)	(20.43)	(-2.36)	(3.28)	(0.60)	(-0.05)
3	1.70%	0.53%	0.975	-0.084	0.199	0.034	-0.089
	(4.48)	(3.20)	(19.17)	(-1.14)	(2.35)	(0.46)	(-2.14)
4	1.57%	0.48%	0.992	-0.067	0.198	0.077	0.083
	(3.92)	(2.41)	(19.14)	(-0.78)	(2.65)	(1.19)	(2.43)
5	1.29%	0.20%	0.953	0.025	0.227	-0.099	-0.041
	(3.23)	(1.25)	(22.64)	(0.39)	(3.90)	(-2.02)	(-1.61)
6	1.68%	0.49%	1.071	-0.140	0.255	0.013	0.033
	(3.93)	(2.63)	(15.08)	(-1.87)	(2.88)	(0.24)	(0.79)
7	1.38%	0.14%	1.032	-0.054	0.120	0.065	-0.031
	(3.47)	(0.84)	(20.60)	(-0.64)	(1.57)	(0.88)	(-1.07)
8	1.29%	0.14%	0.962	-0.007	0.011	-0.110	-0.013
	(3.25)	(0.86)	(20.69)	(-0.09)	(0.12)	(-1.33)	(-0.38)
9	1.39%	0.23%	0.957	-0.036	0.102	0.049	-0.050
	(3.54)	(1.06)	(14.97)	(-0.32)	(0.95)	(0.46)	(-1.03)
10	1.51%	0.50%	0.869	-0.065	0.031	-0.181	-0.014
	(3.81)	(1.97)	(12.62)	(-0.69)	(0.31)	(-2.23)	(-0.25)
L/S	0.46%	0.59%*	-0.112	-0.125	0.077	0.089	-0.007
	(1.51)	(1.72)	(-1.24)	(-0.96)	(0.59)	(0.70)	(-0.11)

B. Fama and French five-factor model

Table B.1: Complicated Processing Portfolios, factor loadings 1977 – 2021, Fama and French

Decile	Excess Returns	5-Factor Alpha	MKT	SMB	HML	RMW	CMA
Panel A: Equal weights							
1	1.38%	0.30%	1.156	0.746	0.290	0.190	0.001
	(4.86)	(2.11)	(27.91)	(10.05)	(3.26)	(1.88)	(0.01)
2	0.8%	-0.11%	1.057	0.558	0.201	0.163	-0.039
	(3.44)	(-1.10)	(39.85)	(11.58)	(3.40)	(2.07)	(-0.48)
3	1.00%	0.04%	1.034	0.553	0.245	0.174	-0.032
	(4.17)	(0.39)	(43.48)	(12.63)	(5.35)	(2.43)	(-0.47)
4	1.00%	0.06%	1.018	0.560	0.244	0.189	-0.012
	(4.36)	(0.77)	(48.57)	(14.46)	(6.02)	(3.78)	(-0.19)
5	1.05%	0.06%	1.019	0.612	0.216	0.227	0.106
	(4.58)	(0.86)	(47.90)	(16.74)	(4.31)	(5.25)	(1.66)
6	1.11%	0.16%	0.997	0.592	0.144	0.213	0.117
	(4.98)	(2.21)	(49.05)	(19.35)	(3.42)	(5.78)	(1.87)
7	1.18%	0.21%	1.008	0.592	0.132	0.212	0.174
	(5.26)	(2.74)	(45.28)	(15.79)	(2.94)	(3.85)	(2.89)
8	1.38%	0.45%	0.977	0.585	0.127	0.176	0.145
	(6.23)	(6.18)	(47.71)	(16.80)	(3.17)	(3.81)	(2.49)
9	1.42%	0.48%	1.004	0.610	0.094	0.074	0.264
	(6.13)	(5.53)	(39.36)	(13.29)	(1.99)	(0.99)	(3.33)
10	1.98%	1.01%	1.045	0.653	0.064	0.149	0.182
	(7.90)	(8.57)	(39.98)	(14.03)	(1.11)	(1.90)	(1.99)
L/S	0.60%***	0.71%***	-0.111**	-0.093	-0.227**	-0.042	0.180
	(3.53)	(3.76)	(-2.42)	(-1.07)	(-2.36)	(-0.29)	(1.13)
Panel B: Value weights							
1	0.93%	0.10%	1.036	0.037	0.006	0.181	0.090
	(3.87)	(0.59)	(28.26)	(0.54)	(0.08)	(1.64)	(0.62)
2	1.01%	0.21%	1.059	-0.056	0.043	0.094	0.078
	(4.44)	(1.68)	(32.70)	(-0.92)	(0.62)	(1.05)	(0.67)
3	1.13%	0.34%	1.059	-0.128	0.132	0.104	0.033
	(5.13)	(3.08)	(35.81)	(-2.45)	(2.05)	(1.43)	(0.36)
4	1.21%	0.43%	1.005	-0.064	0.123	0.112	0.077
	(5.68)	(4.06)	(39.02)	(-1.32)	(2.28)	(1.67)	(0.86)
5	1.20%	0.26%	1.003	0.061	0.085	0.224	0.131
	(5.25)	(2.46)	(35.36)	(1.04)	(1.39)	(2.46)	(1.41)
6	1.31%	0.48%	1.002	0.035	0.077	0.234	0.108
	(6.27)	(4.49)	(31.80)	(0.80)	(1.14)	(3.68)	(1.13)
7	1.14%	0.31%	1.007	-0.033	0.052	0.209	0.147
	(5.49)	(3.31)	(40.17)	(-0.73)	(1.11)	(3.62)	(1.99)
8	1.22%	0.41%	1.005	0.031	-0.006	0.163	0.186
	(5.84)	(3.96)	(39.67)	(-0.72)	(-0.11)	(2.35)	(2.05)
9	1.36%	0.56%	1.008	0.060	-0.095	0.036	0.318
	(6.02)	(3.96)	(28.64)	(0.83)	(-1.33)	(0.25)	(2.69)
10	1.53%	0.61%	1.081	-0.022	-0.046	0.275	0.304
	(6.47)	(4.12)	(30.48)	(-0.38)	(-0.64)	(3.14)	(2.69)
L/S	0.61%***	0.51%***	0.045	-0.059	-0.052	0.094	0.214
	(2.96)	(2.12)	(0.83)	(-0.60)	(-0.48)	(0.64)	(1.07)

Table B.2: Complicated Processing Portfolio over sub-periods, Fama and French

Decile	Excess Returns	1-Factor Alpha	3-Factor Alpha	5-Factor Alpha
Panel A: Equal weights				
A.1: 1977 - 1988				
L/S	1.16%***	1.15%***	1.19%***	1.33%***
	(4.33)	(4.26)	(3.98)	(3.84)
A.2: 1989 - 1999				
L/S	1.45%***	1.36%***	1.45%***	1.64%***
	(4.69)	(4.36)	(4.35)	(4.98)
A.3: 2000 - 2010				
L/S	-0.35%	-0.36%	-0.31%	-0.47%
	(-0.79)	(-0.81)	(-0.75)	(-1.03)
A.4: 2011 – 2021				
L/S	0.12%	0.47%	0.42%	0.38%
	(0.43)	(1.63)	(1.48)	(1.35)
Panel B: Value weights				
B.1: 1977 - 1988				
L/S	1.27%***	1.25%***	1.42%***	1.44%***
	(3.23)	(3.12)	(3.21)	(2.76)
B.2: 1989 – 1999				
L/S	1.08%***	1.00%***	1.12%***	1.27%***
	(2.90)	(2.79)	(2.96)	(3.01)
B.3: 2000 – 2010				
L/S	-0.42%	-0.42%	-0.63%	-0.91%
	(-0.81)	(-0.80)	(-1.30)	(-1.64)
B.4: 2011 – 2021				
L/S	0.46%	0.66%**	0.63%*	0.59%*
	(1.51)	(1.99)	(1.95)	(1.86)

Table B.3: Complicated Processing Portfolios, Fama and French, 1977 – 1989

Decile	Excess Returns	1-Factor Alpha	3-Factor Alpha	5-Factor Alpha
Panel A: Equal weights				
1	1.00%	0.34%	0.09%	0.03%
	(1.83)	(1.32)	(0.47)	(0.12)
2	0.76%	0.12%	-0.06%	-0.04%
	(1.51)	(0.63)	(-0.45)	(-0.20)
3	0.86%	0.24%	-0.00%	0.03%
	(1.75)	(1.29)	(-0.03)	(0.23)
4	0.88%	0.24%	-0.00%	0.04%
	(1.75)	(1.30)	(-0.03)	(0.30)
5	1.12%	0.50%	0.30%	0.38%
	(2.30)	(3.01)	(3.10)	(3.28)
6	1.00%	0.37%	0.13%	0.13%
	(2.03)	(2.02)	(1.33)	(1.15)
7	1.10%	0.45%	0.19%	0.19%
	(2.15)	(2.31)	(1.93)	(1.55)
8	1.27%	0.65%	0.38%	0.43%
	(2.59)	(3.54)	(3.42)	(3.44)
9	1.37%	0.73%	0.43%	0.44%
	(2.74)	(4.07)	(4.59)	(4.05)
10	2.16%	1.49%	1.28%	1.35%
	(4.06)	(6.85)	(6.92)	(6.37)
Panel B: Value weights				
1	0.49%	-0.16%	-0.21%	-0.33%
	(0.95)	(-0.44)	(-0.75)	(-0.95)
2	0.76%	0.18%	0.27%	0.41%
	(1.63)	(0.93)	(1.36)	(1.78)
3	0.81%	0.24%	0.28%	0.37%
	(1.81)	(1.49)	(1.71)	(1.96)
4	0.96%	0.37%	0.39%	0.55%
	(2.07)	(2.31)	(2.21)	(3.05)
5	0.97%	0.39%	0.49%	0.52%
	(2.18)	(2.95)	(3.31)	(3.10)
6	0.97%	0.35%	0.42%	0.51%
	(2.03)	(2.22)	(2.27)	(2.37)
7	1.14%	0.55%	0.57%	0.54%
	(2.46)	(3.43)	(3.55)	(2.64)
8	1.30%	0.72%	0.70%	0.68%
	(2.90)	(4.83)	(4.52)	(3.84)
9	1.50%	0.86%	0.89%	0.93%
	(3.03)	(5.56)	(4.75)	(4.77)
10	1.76%	1.13%	1.21%	1.11%
	(3.41)	(4.77)	(4.70)	(3.95)

Table B.4: Complicated Processing Portfolios, Fama and French, 1989 – 1999

Decile	Excess Returns	1-Factor Alpha	3-Factor Alpha	5-Factor Alpha
Panel A: Equal weights				
1	0.90%	-0.19%	-0.05%	-0.15%
	(2.14)	(-0.72)	(-0.27)	(-0.75)
2	0.48%	-0.56%	-0.49%	-0.51%
	(1.27)	(-2.61)	(-3.98)	(-3.97)
3	0.64%	-0.38%	-0.31%	-0.44%
	(1.68)	(-1.80)	(-1.99)	(-2.54)
4	0.71%	-0.27%	-0.19%	-0.29%
	(1.96)	(-1.30)	(-1.41)	(-2.06)
5	0.66%	-0.28%	-0.20%	-0.33%
	(1.88)	(-1.32)	(-1.62)	(-2.66)
6	0.92%	-0.06%	0.02%	-0.07%
	(2.56)	(-0.35)	(0.18)	(-0.58)
7	1.06%	0.10%	0.18%	0.13%
	(3.10)	(0.56)	(1.54)	(1.03)
8	1.44%	0.48%	0.54%	0.50%
	(4.11)	(2.46)	(4.28)	(3.88)
9	1.50%	0.57%	0.72%	0.73%
	(4.33)	(2.94)	(4.27)	(4.21)
10	2.34%	1.18%	1.40%	1.50%
	(5.14)	(4.44)	(5.85)	(6.58)
Panel B: Value weights				
1	1.05%	-0.01%	-0.09%	-0.20%
	(2.65)	(-0.06)	(-0.40)	(-0.91)
2	1.21%	0.14%	0.01%	-0.00%
	(3.11)	(0.64)	(0.05)	(-0.02)
3	1.35%	0.24%	0.16%	0.03%
	(3.38)	(1.18)	(0.77)	(0.13)
4	1.16%	0.21%	0.15%	0.09%
	(3.20)	(0.93)	(0.66)	(0.35)
5	1.35%	0.34%	0.32%	0.17%
	(3.56)	(1.44)	(1.49)	(0.77)
6	1.54%	0.52%	0.47%	0.42%
	(4.12)	(2.66)	(2.30)	(1.84)
7	1.40%	0.41%	0.31%	0.28%
	(4.00)	(2.44)	(1.91)	(1.56)
8	1.63%	0.58%	0.49%	0.48%
	(4.25)	(2.76)	(2.29)	(1.82)
9	1.65%	0.66%	0.72%	0.74%
	(4.29)	(3.02)	(2.90)	(2.53)
10	2.14%	0.99%	1.03%	1.06%
	(4.93)	(3.75)	(3.72)	(3.61)

Table B.5: Complicated Processing Portfolios, Fama and French, 2000 – 2010

Decile	Excess Returns	1-Factor Alpha	3-Factor Alpha	5-Factor Alpha
Panel A: Equal weights				
1	2.13%	2.12%	1.61%	1.51%
	(3.07)	(5.36)	(4.74)	(4.13)
2	0.95%	0.95%	0.65%	0.44%
	(1.73)	(3.66)	(2.91)	(1.77)
3	1.18%	1.18%	0.81%	0.70%
	(2.22)	(5.35)	(4.83)	(3.80)
4	1.10%	1.10%	0.76%	0.71%
	(2.23)	(5.53)	(4.74)	(4.21)
5	1.14%	1.14%	0.72%	0.50%
	(2.32)	(5.14)	(4.06)	(2.80)
6	1.12%	1.12%	0.71%	0.53%
	(2.39)	(5.32)	(4.09)	(3.07)
7	1.19%	1.19%	0.73%	0.54%
	(2.47)	(4.92)	(4.00)	(2.85)
8	1.51%	1.51%	1.06%	0.89%
	(3.17)	(6.14)	(6.16)	(4.70)
9	1.42%	1.42%	0.92%	0.78%
	(2.63)	(4.87)	(4.60)	(3.49)
10	1.77%	1.76%	1.31%	1.04%
	(3.26)	(5.73)	(5.30)	(3.97)
Panel B: Value weights				
1	1.13%	1.13%	1.27%	1.18%
	(2.07)	(3.07)	(3.49)	(2.86)
2	0.59%	0.59%	0.66%	0.40%
	(1.10)	(2.00)	(2.24)	(1.22)
3	0.69%	0.69%	0.71%	0.66%
	(1.35)	(2.58)	(2.80)	(2.46)
4	1.16%	1.16%	1.19%	1.20%
	(2.53)	(4.56)	(5.34)	(5.22)
5	0.84%	0.84%	0.83%	0.61%
	(1.82)	(3.10)	(3.26)	(2.15)
6	1.09%	1.09%	1.03%	0.91%
	(2.86)	(5.26)	(5.42)	(4.60)
7	0.65%	0.65%	0.59%	0.41%
	(1.49)	(2.58)	(2.62)	(1.87)
8	0.67%	0.67%	0.68%	0.53%
	(1.53)	(2.80)	(3.17)	(2.15)
9	0.91%	0.90%	0.71%	0.62%
	(1.76)	(2.71)	(2.46)	(1.83)
10	0.71%	0.71%	0.63%	0.26%
	(1.36)	(2.05)	(2.09)	(0.78)

Table B.6: Complicated Processing Portfolios, Fama and French, 2011 – 2021

Decile	Excess Returns	1-Factor Alpha	3-Factor Alpha	5-Factor Alpha
Panel A: Equal weights				
1	1.53%	-0.18%	0.19%	0.15%
	(2.63)	(-0.58)	(0.87)	(0.69)
2	1.16%	-0.36%	-0.10%	-0.07%
	(2.37)	(-1.74)	(-0.73)	(-0.53)
3	1.23%	-0.15%	0.11%	0.15%
	(2.74)	(-0.76)	(0.83)	(1.12)
4	1.32%	-0.07%	0.24%	0.25%
	(2.88)	(-0.28)	(1.55)	(1.60)
5	1.28%	-0.19%	0.14%	0.15%
	(2.62)	(-0.73)	(1.01)	(1.13)
6	1.41%	0.04%	0.30%	0.31%
	(3.16)	(0.22)	(2.25)	(2.29)
7	1.39%	0.02%	0.26%	0.27%
	(3.15)	(0.12)	(2.44)	(2.47)
8	1.29%	-0.06%	0.15%	0.13%
	(2.95)	(-0.37)	(1.23)	(1.09)
9	1.39%	0.06%	0.34%	0.30%
	(3.14)	(0.35)	(2.52)	(2.28)
10	1.65%	0.28%	0.60%	0.54%
	(3.48)	(1.14)	(3.20)	(2.82)
Panel B: Value weights				
1	1.06%	-0.25%	-0.17%	-0.22%
	(2.41)	(-1.24)	(-0.86)	(-1.09)
2	1.50%	0.29%	0.34%	0.35%
	(3.76)	(1.71)	(1.97)	(1.97)
3	1.70%	0.57%	0.61%	0.63%
	(4.48)	(3.18)	(3.62)	(3.81)
4	1.57%	0.37%	0.43%	0.47%
	(3.92)	(1.92)	(2.12)	(2.34)
5	1.29%	0.07%	0.21%	0.20%
	(3.23)	(0.40)	(1.23)	(1.18)
6	1.68%	0.39%	0.47%	0.47%
	(3.93)	(1.88)	(2.75)	(2.87)
7	1.38%	0.17%	0.19%	0.19%
	(3.47)	(1.01)	(1.10)	(1.11)
8	1.29%	0.09%	0.12%	0.07%
	(3.25)	(0.59)	(0.77)	(0.45)
9	1.39%	0.27%	0.28%	0.23%
	(3.54)	(1.35)	(1.31)	(1.12)
10	1.51%	0.41%	0.45%	0.37%
	(3.81)	(1.63)	(1.82)	(1.52)

Table B.7: Complicated Processing Portfolios Fama and French, factor loadings 1977 - 1988

Decile	Excess Returns	5-Factor Alpha	MKT	SMB	HML	RMW	CMA
Panel A: Equal weights							
1	1.00%	0.03%	1.044	0.819	-0.128	0.036	0.334
	(1.83)	(0.12)	(20.12)	(8.83)	(-0.94)	(0.23)	(1.72)
2	0.76%	-0.04%	1.009	0.652	-0.078	-0.073	0.118
	(1.51)	(-0.20)	(26.66)	(10.61)	(-0.70)	(-0.60)	(0.92)
3	0.86%	0.03%	1.000	0.696	-0.011	-0.079	0.120
	(1.75)	(0.23)	(25.78)	(14.31)	(-0.14)	(-0.87)	(1.34)
4	0.88%	0.04%	1.016	0.705	-0.080	-0.114	0.229
	(1.75)	(0.30)	(30.26)	(13.75)	(-1.18)	(-1.26)	(2.44)
5	1.12%	0.38%	1.001	0.624	-0.026	-0.140	0.025
	(2.30)	(3.28)	(41.96)	(13.88)	(-0.42)	(-1.75)	(0.27)
6	1.00%	0.13%	1.012	0.713	-0.043	-0.032	0.192
	(2.03)	(1.15)	(36.84)	(16.13)	(-0.64)	(-0.42)	(1.86)
7	1.10%	0.19%	1.031	0.781	0.003	-0.023	0.114
	(2.15)	(1.55)	(33.37)	(15.56)	(0.04)	(-0.24)	(1.24)
8	1.27%	0.43%	1.024	0.697	0.138	-0.073	-0.048
	(2.59)	(3.44)	(37.43)	(12.90)	(2.02)	(-0.84)	(-0.48)
9	1.37%	0.44%	1.063	0.692	0.090	-0.044	0.135
	(2.74)	(4.05)	(43.47)	(15.63)	(1.32)	(-0.52)	(1.16)
10	2.16%	1.35%	1.072	0.635	-0.069	-0.152	0.138
	(4.06)	(6.37)	(22.03)	(8.33)	(-0.59)	(-1.17)	(0.93)
L/S	1.16%***	1.33%***	0.028	-0.184	0.059	-0.189	-0.196
	(4.33)	(3.84)	(0.36)	(-1.47)	(0.30)	(-0.88)	(-0.76)
Panel B: Value weights							
1	0.49%	-0.33%	1.081	0.143	-0.046	0.122	0.355
	(0.95)	(-0.95)	(17.71)	(1.06)	(-0.25)	(0.51)	(1.31)
2	0.76%	0.41%	0.973	-0.043	-0.330	-0.261	0.157
	(1.63)	(1.78)	(22.51)	(-0.53)	(-2.20)	(-1.57)	(0.69)
3	0.81%	0.37%	1.010	-0.127	-0.218	-0.189	0.228
	(1.81)	(1.96)	(24.14)	(-1.94)	(-1.94)	(-1.47)	(1.56)
4	0.96%	0.55%	1.031	-0.073	-0.344	-0.339	0.373
	(2.07)	(3.05)	(24.24)	(-1.19)	(-3.39)	(-2.66)	(2.52)
5	0.97%	0.52%	0.985	-0.058	-0.113	-0.027	-0.082
	(2.18)	(3.10)	(30.40)	(-1.01)	(-1.15)	(-0.25)	(-0.68)
6	0.97%	0.51%	1.057	-0.039	-0.172	-0.157	0.018
	(2.03)	(2.37)	(20.04)	(-0.45)	(-1.32)	(-1.20)	(0.10)
7	1.14%	0.54%	1.027	0.052	-0.037	0.045	-0.016
	(2.46)	(2.64)	(22.80)	(0.72)	(-0.20)	(0.32)	(-0.10)
8	1.30%	0.68%	1.020	0.014	-0.076	-0.020	0.192
	(2.90)	(3.84)	(26.00)	(0.21)	(-0.60)	(-0.15)	(1.41)
9	1.50%	0.93%	1.151	-0.166	-0.022	-0.078	0.032
	(3.03)	(4.77)	(32.10)	(-1.99)	(-0.20)	(-0.67)	(0.19)
10	1.76%	1.11%	1.134	-0.197	-0.292	0.058	0.508
	(3.41)	(3.95)	(18.19)	(-1.80)	(-1.59)	(0.31)	(1.80)
L/S	1.27%***	1.44%***	0.053	-0.341*	-0.246	-0.063	0.154
	(3.23)	(2.76)	(0.55)	(-1.70)	(-0.79)	(-0.18)	(0.33)

Table B.8: Complicated Processing Portfolios Fama and French, factor loadings 1989 - 1999

Decile	Excess Returns	5-Factor Alpha	MKT	SMB	HML	RMW	CMA
Panel A: Equal weights							
1	0.90%	-0.15%	1.076	0.737	0.163	0.054	0.346
	(2.14)	(-0.75)	(18.43)	(9.04)	(1.24)	(0.35)	(2.07)
2	0.48%	-0.51%	1.022	0.585	0.406	0.033	0.050
	(1.27)	(-3.97)	(25.76)	(11.37)	(4.87)	(0.37)	(0.43)
3	0.64%	-0.44%	1.047	0.629	0.275	0.289	0.251
	(1.68)	(-2.54)	(20.95)	(8.11)	(2.75)	(1.74)	(2.19)
4	0.71%	-0.29%	0.991	0.639	0.317	0.274	0.144
	(1.96)	(-2.06)	(24.53)	(9.82)	(3.97)	(2.75)	(1.15)
5	0.66%	-0.33%	0.983	0.666	0.271	0.269	0.292
	(1.88)	(-2.66)	(24.62)	(15.27)	(3.02)	(3.27)	(2.52)
6	0.92%	-0.07%	0.999	0.614	0.223	0.117	0.287
	(2.56)	(-0.58)	(25.52)	(11.24)	(3.14)	(1.30)	(2.49)
7	1.06%	0.13%	0.945	0.521	0.244	0.048	0.141
	(3.10)	(1.03)	(24.12)	(9.96)	(3.39)	(0.59)	(1.28)
8	1.44%	0.50%	0.963	0.545	0.339	0.032	0.110
	(4.11)	(3.88)	(26.12)	(11.15)	(4.42)	(0.43)	(0.99)
9	1.50%	0.73%	0.839	0.534	0.055	-0.210	0.155
	(4.33)	(4.21)	(12.67)	(10.01)	(0.55)	(-1.84)	(1.02)
10	2.34%	1.50%	0.991	0.569	-0.055	-0.615	0.124
	(5.14)	(6.58)	(15.33)	(5.44)	(-0.42)	(-3.51)	(0.79)
L/S	1.45%***	1.64%***	-0.085	-0.168	-0.218	-0.669**	-0.222
	(4.69)	(4.98)	(-0.91)	(-1.45)	(-1.01)	(-2.37)	(-0.92)
Panel B: Value weights							
1	1.05%	-0.20%	1.096	-0.038	-0.007	0.088	0.392
	(2.65)	(-0.91)	(12.75)	(-0.42)	(-0.05)	(0.54)	(1.80)
2	1.21%	-0.00%	1.093	-0.099	0.314	-0.030	0.081
	(3.11)	(-0.02)	(15.12)	(-0.89)	(2.58)	(-0.19)	(0.43)
3	1.35%	0.03%	1.124	-0.051	0.034	0.288	0.275
	(3.38)	(0.13)	(17.43)	(-0.53)	(0.28)	(1.58)	(1.72)
4	1.16%	0.09%	0.951	0.013	0.087	0.099	0.179
	(3.20)	(0.35)	(13.83)	(0.11)	(0.66)	(0.49)	(1.07)
5	1.35%	0.17%	1.087	0.235	0.016	0.042	0.568
	(3.56)	(0.77)	(17.08)	(2.40)	(0.12)	(0.22)	(3.06)
6	1.54%	0.42%	1.023	0.057	0.149	-0.025	0.224
	(4.12)	(1.84)	(14.89)	(0.63)	(1.03)	(-0.12)	(1.11)
7	1.40%	0.28%	0.987	-0.144	0.166	0.016	0.094
	(4.00)	(1.56)	(18.69)	(-2.13)	(1.55)	(0.13)	(0.55)
8	1.63%	0.48%	0.994	-0.142	0.182	0.164	-0.096
	(4.25)	(1.82)	(12.38)	(-1.53)	(1.30)	(0.87)	(-0.31)
9	1.65%	0.74%	0.869	0.030	-0.203	-0.312	0.181
	(4.29)	(2.53)	(10.50)	(0.28)	(-1.23)	(-0.90)	(0.94)
10	2.14%	1.06%	1.019	0.022	-0.121	-0.290	0.103
	(4.93)	(3.61)	(12.43)	(0.19)	(-0.71)	(-1.00)	(0.45)
L/S	1.08%***	1.27%***	-0.077	0.061	-0.113	-0.378	-0.289
	(2.90)	(3.01)	(-0.62)	(0.44)	(-0.44)	(-0.96)	(-0.85)

Table B.9: Complicated Processing Portfolios Fama and French, factor loadings 2000 - 2010

Decile	Excess Returns	5-Factor Alpha	MKT	SMB	HML	RMW	CMA
Panel A: Equal weights							
1	2.13%	1.51%	1.308	0.720	0.293	0.327	-0.353
	(3.07)	(4.13)	(11.19)	(4.05)	(1.63)	(1.64)	(-1.49)
2	0.95%	0.44%	1.175	0.530	0.011	0.374	-0.125
	(1.73)	(1.77)	(15.06)	(5.63)	(0.10)	(2.84)	(-0.79)
3	1.18%	0.70%	1.146	0.461	0.231	0.253	-0.171
	(2.22)	(3.80)	(19.92)	(6.05)	(3.04)	(2.46)	(-1.56)
4	1.10%	0.71%	1.053	0.378	0.266	0.155	-0.151
	(2.23)	(4.21)	(22.30)	(5.46)	(3.50)	(1.85)	(-1.33)
5	1.14%	0.50%	1.075	0.496	0.132	0.274	0.134
	(2.32)	(2.80)	(19.98)	(6.29)	(1.61)	(3.57)	(1.29)
6	1.12%	0.53%	1.008	0.518	0.103	0.248	0.094
	(2.39)	(3.07)	(24.55)	(7.57)	(1.34)	(4.37)	(0.89)
7	1.19%	0.54%	1.003	0.494	0.152	0.190	0.239
	(2.47)	(2.85)	(16.16)	(5.24)	(1.77)	(1.78)	(2.02)
8	1.51%	0.89%	0.957	0.567	0.086	0.194	0.156
	(3.17)	(4.70)	(16.04)	(7.28)	(1.03)	(2.17)	(1.55)
9	1.42%	0.78%	1.043	0.560	0.129	0.083	0.275
	(2.63)	(3.49)	(13.30)	(4.76)	(1.44)	(0.60)	(1.93)
10	1.77%	1.04%	1.088	0.621	0.019	0.322	0.209
	(3.26)	(3.97)	(14.51)	(6.39)	(0.14)	(2.47)	(1.12)
L/S	-0.35%	-0.47%	-0.220*	-0.099	-0.273	-0.004	0.562*
	(-0.79)	(-1.03)	(-1.89)	(-0.51)	(-1.36)	(-0.02)	(1.81)
Panel B: Value weights							
1	1.13%	1.18%	1.014	-0.051	-0.082	0.229	-0.200
	(2.07)	(2.86)	(12.45)	(-0.33)	(-0.55)	(1.30)	(-0.64)
2	0.59%	0.40%	1.195	0.011	-0.194	0.359	0.119
	(1.10)	(1.22)	(13.20)	(0.08)	(-1.36)	(2.34)	(0.53)
3	0.69%	0.66%	1.110	-0.196	0.230	0.099	-0.065
	(1.35)	(2.46)	(14.00)	(-1.72)	(2.12)	(0.82)	(-0.39)
4	1.16%	1.20%	0.948	-0.291	0.252	-0.053	0.057
	(2.53)	(5.22)	(14.33)	(-3.16)	(2.14)	(-0.40)	(0.39)
5	0.84%	0.61%	1.012	-0.081	0.011	0.241	0.208
	(1.82)	(2.15)	(11.47)	(-0.63)	(0.09)	(1.44)	(1.29)
6	1.09%	0.91%	0.833	-0.077	0.098	0.079	0.218
	(2.86)	(4.60)	(14.81)	(-1.08)	(1.06)	(0.79)	(1.69)
7	0.65%	0.41%	0.950	-0.088	0.080	0.161	0.250
	(1.49)	(1.87)	(14.06)	(-0.83)	(0.95)	(1.45)	(1.90)
8	0.67%	0.53%	0.957	-0.055	-0.033	0.166	0.158
	(1.53)	(2.15)	(13.67)	(-0.56)	(-0.33)	(1.33)	(1.21)
9	0.91%	0.62%	0.938	0.122	0.020	-0.028	0.358
	(1.76)	(1.83)	(9.70)	(0.78)	(0.15)	(-0.13)	(1.59)
10	0.71%	0.26%	1.126	-0.019	0.009	0.431	0.306
	(1.36)	(0.78)	(12.29)	(-0.15)	(0.06)	(2.58)	(1.68)
L/S	-0.42%	-0.91%	0.112	0.032	0.091	0.202	0.507
	(-0.81)	(-1.64)	(0.86)	(0.14)	(0.42)	(0.85)	(1.37)

Table B.10: Complicated Processing Portfolios Fama and French, factor loadings 2011 - 2021

Decile	Excess Returns	5-Factor Alpha	MKT	SMB	HML	RMW	CMA
Panel A: Equal weights							
1	1.53%	0.15%	1.236	0.587	0.355	0.104	0.159
	(2.63)	(0.69)	(20.72)	(4.20)	(3.25)	(0.88)	(1.01)
2	1.16%	-0.07%	1.111	0.358	0.345	-0.060	-0.129
	(2.37)	(-0.53)	(29.14)	(4.92)	(4.53)	(-0.71)	(-1.10)
3	1.23%	0.15%	0.985	0.412	0.315	-0.089	-0.165
	(2.74)	(1.12)	(25.12)	(6.77)	(5.44)	(-1.06)	(-1.62)
4	1.32%	0.25%	0.968	0.506	0.332	0.025	-0.104
	(2.88)	(1.60)	(21.32)	(6.82)	(4.76)	(0.27)	(-0.86)
5	1.28%	0.15%	1.010	0.570	0.348	0.008	-0.094
	(2.62)	(1.13)	(24.23)	(8.51)	(4.45)	(0.08)	(-0.81)
6	1.41%	0.31%	0.979	0.443	0.289	0.044	-0.153
	(3.16)	(2.29)	(28.38)	(7.14)	(3.63)	(0.56)	(-1.20)
7	1.39%	0.27%	0.983	0.470	0.223	0.008	-0.086
	(3.15)	(2.47)	(29.78)	(9.83)	(3.51)	(0.10)	(-0.94)
8	1.29%	0.13%	0.990	0.450	0.157	0.136	-0.030
	(2.95)	(1.09)	(27.23)	(7.14)	(2.58)	(1.70)	(-0.31)
9	1.39%	0.30%	0.957	0.555	0.143	0.061	0.228
	(3.14)	(2.28)	(21.07)	(7.45)	(1.89)	(0.64)	(1.77)
10	1.65%	0.54%	0.962	0.698	0.124	0.181	0.343
	(3.48)	(2.82)	(21.96)	(8.54)	(1.91)	(1.26)	(2.98)
L/S	0.12%	0.38%	-0.273***	0.111	-0.230*	0.078	0.185
	(0.43)	(1.35)	(-3.82)	(0.69)	(-1.78)	(0.43)	(0.96)
Panel B: Value weights							
1	1.06%	-0.22%	1.048	0.173	0.028	0.235	0.094
	(2.41)	(-1.09)	(19.96)	(1.54)	(0.30)	(1.85)	(0.66)
2	1.50%	0.35%	1.024	-0.191	0.243	-0.068	-0.026
	(3.76)	(1.97)	(20.04)	(-2.42)	(2.30)	(-0.66)	(-0.15)
3	1.70%	0.63%	0.956	-0.156	0.185	-0.145	0.022
	(4.48)	(3.81)	(18.17)	(-1.84)	(1.62)	(-1.22)	(0.12)
4	1.57%	0.47%	0.970	-0.088	0.251	-0.070	-0.236
	(3.92)	(2.34)	(19.08)	(-1.02)	(2.70)	(-0.70)	(-1.41)
5	1.29%	0.20%	0.966	0.045	0.270	0.052	-0.017
	(3.23)	(1.18)	(21.38)	(0.60)	(3.40)	(0.55)	(1.18)
6	1.68%	0.47%	1.043	-0.072	0.306	0.207	-0.274
	(3.93)	(2.87)	(16.59)	(-0.88)	(2.52)	(1.10)	(-1.66)
7	1.38%	0.19%	1.003	-0.070	0.104	0.005	-0.075
	(3.47)	(1.11)	(21.87)	(-0.72)	(1.44)	(0.04)	(1.11)
8	1.29%	0.07%	1.018	0.031	-0.054	0.052	0.368
	(3.25)	(0.45)	(17.53)	(0.41)	(-0.62)	(0.47)	(2.69)
9	1.39%	0.23%	0.962	-0.018	-0.041	0.078	0.331
	(3.54)	(1.12)	(13.40)	(-0.15)	(-0.49)	(0.72)	(1.83)
10	1.51%	0.37%	0.957	0.001	-0.058	0.096	0.545
	(3.81)	(1.52)	(14.96)	(0.01)	(-0.56)	(0.54)	(2.97)
L/S	0.46%	0.59%*	-0.091	-0.172	-0.086	-0.139	0.451*
	(1.51)	(1.86)	(-1.07)	(-1.15)	(-0.62)	(-0.66)	(1.91)